Exploiting physiological changes during the flow experience for assessing virtual-reality game design

Lazaros Michailidis

Submitted in fulfilment of the requirements for the degree of

Doctor of Engineering in Digital Media

Faculty of Media and Communication

Academic supervisors	Dr Emili Balaguer-Ballester
	Dr Xun He
Industry supervisor	Jesus Lucas Barcias

February 2021

Copyright Statement

Publications and Conference Presentations

Journal articles

- Michailidis, L., Balaguer-Ballester, E., & He, X. (2018). Flow and immersion in video games: The aftermath of a conceptual challenge. *Frontiers in psychology, 9*.

  This mini review addresses the common aspects of flow and immersion, but also highlights how flow and presence are two distinct concepts. It is argued that flow and immersion appear to refer to the same state from both a conceptual as well as a physiological standpoint, whereas presence seems to occur temporally before flow or immersion, and it is elicited by different neural patterns from those found for flow.

- Michailidis, L., Barcias, J. L., Charles, F., He, X., & Balaguer-Ballester, E. (In preparation). Towards an optimal EEG-channel set for detecting flow in virtual reality games.

Posters

- Michailidis, L., Barcias, J. L., Balaguer-Ballester, E., & He, X. (2018). Bringing Tower Defense into the physiological study of flow in virtual reality. In *EuroVR Conference 2018*, London, UK.

  This work was submitted as a poster to EuroVR 2018 conference, demonstrating preliminary statistical results from the second study carried out in this project. It argues that the Tower Defense genre is compliant to traditional experimental paradigms, offering a naturally easy way to annotate physiological data based on its design architecture.

- Michailidis, L., Barcias, J. L., Charles, F., He, X., & Balaguer-Ballester, E. (2019). Combining Personality and Physiology to Investigate the Flow Experience in Virtual Reality Games. In International Conference on Human-Computer Interaction (pp. 45-52). Springer, Cham.

  This work was submitted to the Human-Computer Interaction (HCII) conference that took place in Orlando, Florida, United States. It was submitted in the form of an extended abstract for a poster presentation. The motivation was to demonstrate that the self-reported flow experience can be classified with heart rate variability parameters. In addition, we employed personality dimensions from the Big Five Inventory scale and demonstrated that the classification performance of high- and low-flow episodes increased considerably when personality was embedded in the training stage.

Presentations

- Michailidis, L., Barcias, J. L., Charles, F., He, X., & Balaguer-Ballester, E. (2019, July). Attentional control during the flow experience in virtual reality games. In *The Second Neuroadaptive Technology Conference* (p. 51).

  This work was presented at the Neuroadaptive 2019 conference in Liverpool, UK, demonstrating the potential of eye blink parameters as a surrogate measure for the flow experience. Using Machine Learning, we were able to demonstrate that the self-reported flow experience presented blink patterns that corresponded to high- and low-flow episodes.

List of Abbreviations

ANS	Autonomic nervous system
BDT	Bagged ensemble of CART decision trees (classification algorithm)
CNS	Central nervous system
EBV	Eye blink variability
ECG/EKG	Electrocardiography
EEG	Electroencephalography
EOG	Electrooculography
FAA	Frontal alpha asymmetry
FCM	Fuzzy C-means (unsupervised learner)
HMD	Head-mounted display
HR	Heart rate
HRV	Heart rate variability
kNN/k-NN	K-Nearest Neighbors (classification algorithm)
MRMR	Minimum redundancy maximum relevance algorithm for feature ranking and selection
PNS	Peripheral nervous system
SMOTE	Synthetic Minority Oversampling Technique (algorithm for data oversampling)
SVM	Support Vector Machine (classification algorithm)
TD	Tower Defense (game genre)
TD-VR	A virtual reality game developed for the purposes of this project in PlayStation VR
VR	Virtual reality

Contents

List of Figures

List of Tables

Abstract

Immersive experiences are considered the principal attraction of video games. Achieving a healthy balance between the game's demands and the user's skills is a particularly challenging goal. However, it is a coveted outcome, as it gives rise to the flow experience – a mental state of deep concentration and game engagement. When this balance fractures, the player may experience considerable disinclination to continue playing, which may be a product of anxiety or boredom. Thus, being able to predict manifestations of these psychological states in video game players is essential for understanding player motivation and designing better games.

To this end, we build on earlier work to evaluate flow dynamics from a physiological perspective using a custom video game. Although advancements in this area are growing, there has been little consideration given to the interpersonal characteristics that may influence the expression of the flow experience. In this thesis, two angles are introduced that remain poorly understood. First, the investigation is contextualized in the virtual reality domain, a technology that putatively amplifies affective experiences, yet is still insufficiently addressed in the flow literature. Second, a novel analysis setup is proposed, whereby the recorded physiological responses and psychometric self-ratings are combined to assess the effectiveness of our game's design in a series of experiments.

The analysis workflow employed heart rate and eye blink variability, and electroencephalography (EEG) as objective assessment measures of the game's impact, and self-reports as subjective assessment measures. These inputs were submitted to a clustering method, cross-referencing the membership of the observations with self-report ratings of the players they originated from. Next, this information was used to effectively inform specialized decoders of the flow state from the physiological responses. This approach successfully enabled classifiers to operate at high accuracy rates in all our studies. Furthermore, we addressed the compression of medium-resolution EEG sensors to a minimal set required to decode flow. Overall, our findings suggest that the approaches employed in this thesis have wide applicability and potential for improving game designing practices.

Keywords: flow experience, anxiety, boredom, virtual reality, personality, electrocardiography, electroencephalography, electrooculography, HRV, fuzzy clustering, video games, immersion

Acknowledgements

My PhD journey has been long and arduous. However, the positives outweighed the temporary inconvenience it caused, as well as the mountainous work that had to be done. A network of supportive individuals made that possible and helped me cultivate resilience. First and foremost, I wish to thank the Center for Digital Entertainment and Bournemouth University for this brilliant opportunity. Without them, this work would not have been carried out. With them, working for Sony Interactive Entertainment, a dream job to many, became a reality. The lovely staff, including Michael Board, Daniel Cox, and Zoe Leonard, were always happy to and never failed to provide their help; they made me feel welcome and facilitated the conditions of my industry placement. I hold them all very dear. Next, my sincere appreciation to Dr Emili Balaguer-Ballester and Dr Xun He, my academic advisors, for keeping me inspired, motivated, guided, and aspired to better my research conduct. They treated me as an equal rather than "just another student" and allowed freedom of thought and decision-making; for that, I am indebted.

My dear friend, Ifigeneia Mavridou, without whom I would have probably lost my sanity. Our daily chats and her consolation are not traits one should take for granted in their friendships. Jesus Lucas Barcias, my industry advisor, was pivotal to this work. He knows very well how much I have appreciated his support and his ability to negate my doubts with nothing more but logical arguments. He has put a lot of trust and hard work in my name; thanks to him, I am now officially a full-time employee in Sony. My gratitude is ineffable. Last, but not least, I wish to thank Dr Fred Charles, Dr Charlie Hargood and Dr Ellen Seiss, Tom Collingwood, Jeremy Hogan, and Fabio Capello for their guidance and feedback – their contribution helped shape this work.

1 Introduction

1.1 Background

Video games are among the contemporary entertainment tools that are largely expanding. In the recent decades, the digitization of physical games makes gaming more accessible than ever before. What makes video games particularly attractive is their ability to instill immersive experiences. The present thesis focuses on flow, a mental state of deep concentration and loss of contact with the surrounding world, that is typically experienced during game playing. Video game designers acknowledge the importance of flow, by providing content that hooks the players for long periods. Flow is commonly viewed as the result of an optimal level of challenge relative to the player's skills. It has been associated with affective components, whereby this mental phenomenon becomes highly pleasant. Hence, flow can also be defined with the widely adopted dichotomy in the context of emotions, namely arousal and valence (e.g., Lang, 1995). Importantly, an engaging video game stimulates willingness to replay the game. The games industry is essentially a business; thus, replay intention equates to customer loyalty, which further taps into revenue stabilization and increase.

Game designers are committed to delivering entertainment. Through an array of monetization techniques, ranging from one-time purchase and subscriptions to microtransactions, they have commercialized digital entertainment. A look at the numbers paints a clear picture. The games industry is a multibillion-dollar business with a staggering estimate of $152.1 billion revenue for the year 2019 (Newzoo, 2019). Several years are spent into the development cycle of a video game (Teipen, 2016); with an average work schedule of 35 hours per week per employee (Young, 2018), this would translate to hundreds of thousands of hours expended in the development of a single video game. Video game players play on average 13 up to 23 hours per week (Mills et al., 2018; Yee, 2006).

Thousands or even millions of players can coexist in an online virtual world at the same time (Fernandes et al., 2018). The popular multiplayer game World of Warcraft by Blizzard

Entertainment (Irvine, California, United States) has had over a hundred million registered accounts (Sarkar, 2014), paid subscriptions from more than 12 million players in 2010 (Molina, 2011) and 10 million active players (Ojeda et al., 2019). Finally, the cumulative playtime of games bought on the popular Steam platform amounts to 1.11 million years (O'Neill et al., 2016). These gargantuan quantities are only a glimpse at the success of video games.

What makes the games industry so lucrative is that it satisfies people's needs (Przybylski et al., 2010; Ryan et al., 2006). Video games propitiate the ego, by indulging players into the pursuit of recognition, power, sense of belonging and self-esteem, which are primitive needs of the human species. These egocentric needs may not be easily attainable in real life; yet, video games provide the platform for actualizing them, by utilizing a plethora of game design techniques. People demonstrate the occasional need to escape reality, and thus engage in video games (Calleja, 2010), even though that very need may prove pathological (Warmelink et al., 2009).

However, video game playing is not evolving into a habit merely as a tool to subdue existing problems (see e.g., Plante et al., 2019). Rather, it is chiefly sought because it offers pleasure, a positive experience that underlies fundamental aspects of human psychology (Kahneman et al., 1999; Russell, 2003b). This is the emergence of homo ludens, where the individual enters a fictional reality deprived of the real-life consequences (Huizinga, 1950; Soetaert et al., 2011). Games have existed since ancient times and thus entertainment may not merely be a luxurious affordance, but instrumental to human optimal functioning (e.g., Ruckenstein, 1991).

Nowadays, video games have been injected in the pop culture through traditional computers, mobile devices, consoles, virtual reality, augmented reality and mixed reality. PlayStation, one of the top consoles in the video games scene, boasting nearly 37 million PlayStation Plus subscribers, sold approximately 103 million PlayStation 4 units (Sony Interactive Entertainment, 2019), 4.2 million units of PlayStation VR (Shuman, PlayStation Blog, 2019), as well as 12.2 million games since December 2017 worldwide (Gilyadov, IGN, 2017).

Virtual reality (VR), in particular, has seen its commercial release in the previous decade. Contrary to traditional displays, such as computer or mobile screens, VR can produce richer experiences (e.g., Pallavicini et al., 2019) and compelling virtual worlds. As technology advances, we should expect higher graphical fidelity and computational power that are able to produce highly realistic virtual environments. In addition, new sensory modalities are currently being investigated, including haptics (e.g., Farooq et al., 2019) and olfactory stimulation (e.g., Braun, 2019), with which we can further augment the vividness of VR. These developments call for a better understanding of the variables factoring in the experience of virtual reality users. However, the tools to make such an investigation possible are lacking.

The importance of ensuring flow's onset in video games is as fundamental for traditional displays as it is for virtual reality games. Yet, understanding how flow emerges is not only important for the individual user. The success of video games may additionally rely on the popularity they amass. With the emergence of "let's play" videos on social platforms, such as YouTube (Glas, 2015), the public interest can increase both rapidly and significantly. It might take a single person to influence prospective players (Sjöblom & Hamari, 2017). User reviews are also important, in that they may determine the arrival and purchase inclination of new players (Bond & Beale, 2009).

Questions such as "What keeps the player concentrated in the game?" or "How to prevent my player base from discontinuing our game?" may be very commonly addressed to game designers. It is not unusual for the game design to be adjusted such that it favors participation in in-game purchases (Hamari et al., 2017; Hooker et al., 2019). Even in this scenario, the flow experience becomes important, as it can enhance purchase intention of virtual “goods” (Liu & Shiue, 2014). Therefore, video games relying on loot boxes or random reward mechanisms (Nielsen & Grabarczyk, 2019) for profitability should similarly rely on forging flow episodes.

Research in video games has bridged evidence from psychology and neuroscience to elucidate the underpinnings of flow. The physiological mechanisms that underlie flow may help game designers understand or even predict the conditions under which their game will be successfully embraced before release. This process can be facilitated by the development of tools that are beneficial to the industry, without imposing the demand for theoretical or technical prowess.

Assessing when flow occurs in video game play is not only vital to the industry for the obvious reasons of revenue increase, optimization of user experience and a better understanding of the player base. It is also interesting to the academic world, as it can help design effective prevention, intervention and rehabilitation strategies for game addiction. Indeed, many studies have asseverated the link between game-induced flow and addiction (e.g., Boniwell, 2008). Likewise, it is interesting because of a growing body of evidence suggesting that enhanced performance and augmented cognitive capacity may result from the flow experience (e.g., Kozhevnikov et al., 2018; Stoll, 2018). Flow research may be more complicated when dealing with new-generation platforms, as in the case of virtual reality. This is because of its novelty, and hence less understood than traditional video games that have been commercially available for decades.

This section provided an overview of video games in the contemporary market. It can be seen that digital games stand prominently amongst other forms of media and entertainment, further illustrating that flow is a major component of a successful video game. Although there are already developments in detecting flow through physiology, they are mostly a prerogative of PC and console gaming. Thus, it was stressed that novel technologies, such as virtual reality, have not been widely studied in the specific context of flow. In the next section, we will discuss methodological concerns in the study of flow and how they can be incorporated into a virtual reality setting.

1.2 Research Problems

1.2.1 Conceptualization

The conceptualization of flow remains debated; the principal challenge in investigating this mental state stems from its multifaceted nature. Flow is usually compartmentalized into antecedent qualities, core dimensions and experiential outcomes (e.g., Hoffman & Novak, 1996; Hooker et al., 2019). This finer segmentation enables research to operationalize flow by targeting specific variables and studying their effects on the suppression or expression of game-induced flow. Furthermore, revisiting flow as a system of input – process – output, that the antecedents – dimensions – outcomes represent, falls within the longstanding computing metaphor of the mind (Varvatsoulias, 2014). Such conceptualizations can similarly be observed in other definitions, including engagement (Bonometti et al., 2020; O'Brien & Toms, 2008) and immersion (Brown & Cairns, 2004; Ermi & Mäyrä, 2005). Indeed, the inconsistent views on flow's core qualities fosters a predicament that has spawned various definitions of similar phenomenology.

For this reason, there is an abundance of methodological approaches aiming to identify how flow-like episodes occur. While assailing the problem from multiple angles is desirable, the diversity of the existing approaches complicates the landscape of flow research and hinders the standardization of measurement practices. This is not only problematic for the study of flow in general, but also for migrating the study of flow to a VR setting.

Bringing VR into the flow discourse is of relevance given the former's connection to *presence*, a term that many researchers agree it refers to a sense being in the virtual world (McCreery et al., 2013). Although presence and flow have been considered different states (Weibel & Wissmath, 2011), immersion incorporates elements from both (Brown & Cairns, 2004; McMahan, 2003). These subtleties in nomenclature may seem innocuous, but they reciprocally influence the conceptualization, operationalization, and pragmatic measurement of flow.

Proponents of immersion have argued that flow is rarer than immersion in video games (e.g., Brockmyer et al., 2009). The basis for this argument revolves around the idea that flow is an optimal and extreme state (Jennett et al., 2008). Contrarily, it is suggested that immersion is sub-optimal, therefore more suitable for video game play, considering negative experiences can also be evoked whilst playing (Jennett et al., 2008). According to Brown and Cairns (2004), immersion is a gradual state comprising three stages of increasing intensity, i.e., engagement, engrossment, and total immersion. In their work, they explained that engagement refers to the first-level entry that requires effort, time, and attention. The second stage, engrossment, progresses into sustained attention and emotional investment. Finally, total immersion is viewed as being equal to presence (Brown & Cairns, 2004).

In Michailidis, Balaguer-Ballester and He (2018), we argued that the assumed distinction between immersion and flow does not have empirical support from a methodological standpoint. Any such claims are either definitional (e.g., Cairns et al., 2014; Jennett et al., 2008), speculative over game design elements (e.g., Nacke & Lindley, 2008) or of psychometric origin (e.g., Kannegieser & Atorf, 2020; Procci & Bowers, 2011). The obvious conceptual overlap between immersion and flow in the literature (Procci & Bowers, 2011), specifically for the engagement and engrossment stages, alludes to these states sharing the same neural correlates.

However, in the case of presence, a distinguishable neural representation has been identified (e.g., Baumgartner et al., 2008; Clemente et al., 2013) that also clearly separates it from flow on a theoretical level (Weibel & Wissmath, 2011). Consequently, total immersion does not present an organic physiological connection to the previous stages of engagement and engrossment. The separation between flow and immersion is therefore not well justified in the literature.

In this work, the focus is placed on the flow experience as proposed by Csikszentmihalyi (1990), yet the problems identified in this section are not unique to the flow literature. Thus, the findings of this work may extend their support to the construct of immersion. For a more detailed review of the immersion, flow, and presence concepts, refer to Michailidis, Balaguer-Ballester and He (2018).

1.2.2 Individual characteristics

Video game characteristics can be highly variable and unique to specific video games (e.g., Nah et al., 2014). Thus, flow may become highly unpredictable across different players. In addition, personal characteristics can instigate differences in the expression of flow episodes. For that reason, flow is considered highly subjective (e.g., Webster et al., 1993), which makes it definitionally challenging to address at an objective level. The systematic approach at defining flow, that upholds its subjective nature, is typically done through self-reports (e.g., Rißler et al., 2018). At an attempt to make a more objective evaluation of flow, physiological measures are employed either as an alternative to self-reports (e.g., Berta et al., 2013) or as complementary measures (e.g., Harmat et al., 2015; Peifer et al., 2014). However, the experimentally controlled variables may not consistently account for inter-player differences, but rather depend on the design principles with which the trials were structured.

In the traditional theory of flow, anxiety and boredom are thought to antagonize flow, since they can interfere with or even replace flow (e.g., Massimini & Carli, 1986). Not only can these states result from the design of the game, they can also be differentially expressed by players. While some players may be more prone to experiencing anxiety, others are more resistant to this psychosomatic state, given certain personality traits or player expertise (e.g., Augustine & Larsen, 2015; Borderie & Michinov, 2016). For example, personality may determine one's flow proneness, i.e., whether someone can access the flow experience (e.g., Ullén et al., 2012). Yet, this is rarely taken into consideration in the video games literature. Instead, flow appears to be traditionally approached as a state that anyone can attain, despite evidence suggesting otherwise (e.g., Csikszentmihalyi, 1997a).

A known example is the personality trait of neuroticism, which has been consistently shown to be unconducive to flow (e.g., Heller et al., 2015). Research employs a common paradigm wherein biometric data from different time points (observations) are fed into complex algorithms, which aim to identify patterns amongst them. These patterns are part of a classification process, where the observations are grouped together to fit an array of classes to which they are thought to belong. For example, they may group observations from experimental trials designed to elicit flow, boredom and anxiety (e.g., Berta et al., 2013). However, player characteristics can be inadvertently suppressed in the process, and thus the classification precision may be compromised due to unaccounted factors in the data – i.e., characteristics of the player in addition to game characteristics. Indeed, the game's characteristics are as important as the player characteristics – the interaction of the two result in the final player experience (e.g., Nah et al., 2014).

1.2.3 Physiological study of flow in VR

Nevertheless, numerous studies have used different psychophysiological measures to address flow in video games, but similar experimentations are scarce in the VR domain. Currently, there are lacking contributions in this area, with little direction on the integration of flow's physiological assessment in VR (see e.g., Bian et al., 2016). This is mainly because of the novelty of the VR medium in the market, but also due to the increased equipment requirements for carrying out such research. Moreover, the drawback of motion sickness, commonly paired with VR, has only recently become a prominent issue in the studies of user experience. These relatively understudied concerns evince a broad area of research. The increasing interest in brain-computer interfaces has turned the attention to solutions such as electroencephalography (EEG), which gradually become highly coveted in the field of human-computer interaction (e.g., Li et al., 2017).

Electroencephalography (EEG) is preferred for portability, high temporal precision and ability to convey information that other physiological mechanisms may not. However, in the case of VR, it may be treated with reluctance. This is because of EEG's sensitivity to interference and muscle movement (e.g., Motamedi-Fakhr et al., 2014) as well as the time required to set it up. Equipment-wise, medical-grade EEG may be inconvenient due to the high number of electrodes and wires involved, making research outcomes impractical to the industry. On the other hand, lighter alternatives such as Emotiv EPOC+ (Emotiv Inc., San Francisco, California, United States) or NeuroSky (NeuroSky Inc., San Jose, California, United States) neuroheadsets, that typically feature a limited number of electrodes, may have an insufficient spatial distribution for the purpose of detecting flow. Yet, VR headsets are fixed on the head, and thus EEG is tempting to integrate due to its reliance on sensors that are similarly fixed on the scalp.

Based on the issues mentioned in this section, the detection of the flow state in a VR setting poses considerable challenge. Evidence from non-VR studies is still helpful to guide our study in the VR domain, but more customized approaches may be needed towards the successful identification of the flow state. The interaction of game and player characteristics, both acting as antecedents to flow, should not be dismissed, as they can provide rich information about the dynamics of the player's experience. Taken together, the goals of this research are stated in the next section.

1.3 Research Goals

The motivation of this research is to detect the flow state in VR game play. The states of boredom and anxiety are also considered since they have been widely subsumed under the context of flow and video game play.

The objectives of this work can be summarized as follows:

1. The first goal is to investigate the influence of personality on the self-reported experience of flow in the context of virtual reality game play. The aim is to understand which personality traits best explain the reported experience of flow. Addressed in Chapter **4**.
2. The second goal is to identify physiological correlates of the flow state in different phases of the game based on design decisions using personality traits as an ancillary measure. In this case, we create design rules for different types of game levels, to experimentally induce flow, anxiety and boredom. The aim is to identify how the physiological responses are generated based on the design rules of each game level. Addressed in Chapter **5**.
3. A third goal is to determine electrophysiological activity that is best represented in settings where flow, anxiety and boredom are desired to be elicited by design. EEG is typically subdivided into frequency bands of neural activity originating from different areas of the cortex. In this step, we aim to identify the most relevant bands and to roughly locate the associated cortical regions. This includes anatomical inspection of laterality (left versus right hemispheric activity) and posteriority (frontal or anterior versus posterior regions). Addressed in Chapter **6**.
4. The final goal is to conduct a more scrupulous search to single out electrodes that best contribute to the detection of flow, anxiety and boredom. In doing so, we seek to identify the optimal minimum number and location of electrodes required for decoding the mental states of interest. The idea of starting with a higher EEG spatial resolution with the prospect of reducing it to the minimum optimal number of electrodes has not been extensively addressed in the context of flow. This is an important topic for games research, in that EEG conducted in a commercial setting should aim to minimize manufacturing costs. Addressed in Chapter **6**.

The goals mentioned in this section will be approached with the methodological considerations addressed in **Chapter 3: General Methodology**. We consider our findings to be beneficial to both the industry and academia, with the broader aims being:

a) Improve user experience studies by measuring flow, anxiety and boredom in a variety of game settings to pinpoint the areas of the game that are detrimental to player retention.
b) Facilitate the creation of rehabilitation strategies for game addiction.
c) Provide richer virtual reality experiences, by enabling developers to utilize real-time data and adapt their game for sustainable flow episodes.
d) Foster the effectiveness of VR training programs.

1.4 Structure of the thesis

The present manuscript documents our approach toward detecting the flow experience in virtual reality. It comprises a chapter that reviews existing work (Chapter 2), in order to familiarize the reader with the flow experience in the context of video game design. The methodological approach employed in this project is further provided in the following chapter (Chapter 3), wherein it is explained how a virtual reality game for PlayStation VR was designed to meet the research goals. In addition, information is given on the materials that were used to assess flow and personality, as well as the analysis and classification methods used to reach the studies' objectives.

The second part of this document is the pith of the project, detailing the experiments carried out to test our hypothesis. Each subsequent chapter in this part is structured with an introduction, a methodology section, their findings and a discussion that reflects and concludes on those findings, by making associations with existing literary work. The first chapter of this part (Chapter 4) concerns the initial study that was conducted to investigate the potential for flow in our custom VR game. In the second chapter (Chapter 5) the physiological measures of electrooculography and electrocardiography were used to investigate the physiological responses of the players during VR game play. This study was also complemented with a dynamic difficulty scaling mechanic to accommodate a more balanced game play to players of variable skills. Classification algorithms were used to estimate a better representation of the data. We specifically aimed to compare our design choice assumptions of when flow, anxiety and boredom were to occur versus the self-reported flow experience of the participants.

The third chapter of this part (Chapter 6) implemented electroencephalography to identify the minimum number of channels, their location and the related frequency bands for the detection of flow. Similarly, as in Chapter 5, classification was used to decode the different mental states of interest (flow, anxiety and boredom). We also explored frontal EEG alpha asymmetry, which has been given very little attention in the literature of flow. The final part of this document is a standalone chapter (Chapter 7) that draws conclusions from the work done and proposes future work to improve and extend the presented methodology.

2 Research Background

In this chapter, we will review the concept of flow and how it emerges in video game playing. We begin by documenting earlier accounts of the flow experience, which supported flow as an operationalized form of happiness, and proceed to demonstrate how flow is relevant in the domain of video games. Further, we explain game design techniques to help elicit flow episodes that maintain player motivation and foster replay intention as a prelude to our experimental approach. Finally, we specifically address video game play in virtual reality applications and the implications for flow and other mental states. We conclude by detailing different approaches to the measurement of flow that are prevalent in existing research.

2.1 Disambiguating happiness

Mapping the complexity of human behavior has been the focus of research for decades. American psychologist Abraham Harold Maslow theorized the motivation behind human behavior and formulated a pyramidal hierarchy of needs (Maslow, 1943). He posited that behavior is goal-driven and humans traverse through a range of bio-psychological needs. On the lowest level of the hierarchy reside the biological needs, such as food and water consumption, as well as reproduction, which are instinctive and uphold an evolutionary significance that warrant the organism's survival and procreation of the species. Low-level needs are mechanistic, but also straightforward: they can be modelled as a simple input-output pair, or stimulus-response, whereby the input may be illustrated as a mere thalamic signal.

However, higher-order needs are more complex, as they aim to assuage psychological drives, such as the need for safety and the need to belong (Kunc, 1992). Contrary to their lower-level neighbors, those needs are challenging to reduce to a simple input and output, in that the output is determined by multiple intermediary factors (e.g., Csikszentmihalyi, 1988). In pursuing the fulfillment of higher-order needs, an individual may attain self-actualization (Poston, 2009). Self-actualizing individuals are portrayed as being optimally functioning in both intrapersonal and interpersonal contexts (e.g., Leary, 2007; Maslow, 1974). Achieving self-actualization has been shown to present a strong relationship with happiness (Beaumont, 2009).

The hedonic model, proposed by Diener and Larsen (1993), views happiness as the ratio of positive affect to negative affect. Consequently, a dominating numerator would result in a better approximation of well-being. According to Kahneman (1999), this ratio is a marker of an “objectively happy” person. Larsen (2000) further contributed to this assertion and suggested that people have a propensity toward regulating their mood positively, thereby engaging in a variety of actions to achieve it. Consistent with the hedonic model, individuals appear to have a predilection for experiencing positive affect over negative affect. More crucially, happiness appears to be more pertinent to the frequency of experiencing positive affect, whereas the duration thereof may have a more subsidiary role (Diener et al., 2009).

Yet, albeit Diener and Larsen (1993) approached happiness through a quantifiable ratio, the challenge is that it becomes murky predicting when the person may experience positive and negative affect. Thus, recording the person's positive experiences and modeling their objective happiness is somewhat infeasible in practice, as the quantification tools for happiness are lacking. Indeed, happiness is a loose construct that is hard to examine directly. This apparent gap may be complemented via a concept known as the flow experience, which is reviewed in the next section.

2.2 The flow experience

Csikszentmihalyi (1975) operationalized a set of measurable phenomena that were identified through oral patterns in interviews (also, Csikszentmihalyi, 1988). The emergent pattern was that people reported experiencing their happiest moments during their involvement in an activity. Csikszentmihalyi coined the term “flow experience” to describe the cognitive and affective concomitants of this involvement. The importance of flow lies in its connection to satisfaction and hedonism (Peterson et al., 2005). Csikszentmihalyi (1990) postulated that flow is a state that can unlock happiness.

Flow's effects on mood improvement may be transient (e.g., Guo et al., 2012; Rodríguez-Sánchez et al., 2008, 2011) but also long-term whereby they contribute to subjective well-being and life satisfaction (Ryff, 1989; Sahoo & Sahu, 2009). To secure a happy life, one has to experience flow frequently (Csikszentmihalyi, 1997b, 1997c). These observations corroborate the formulaic conceptualization of subjective well-being proposed by Diener and Larsen (1993) that was mentioned earlier; experiencing flow frequently would result in a dominant numerator, enabling the individual to approach happiness (also, Diener et al., 2009).

2.2.1 Finding flow – qualifying the experience

Notably, the flow experience has been regarded as a psychological need per se, "the need to seek and master challenges" (Baumann & Scheffer, 2011). The need to seek challenges has also been acknowledged outside the flow context (e.g., White, 1959; Deci, 1975). Virtually any activity may elicit the flow experience consistently across different individuals as well as different activities (Csikszentmihalyi, 1988; Nakamura & Csikszentmihalyi, 2014). For example, sports, painting, working, playing games and music composition are all viable platforms to flow (Ceja & Navarro, 2011; Csikszentmihalyi, 1997a; Jackson & Csikszentmihalyi, 1999). Applying Larsen's (2000) postulate of a personal tendency to yearn for happiness, while also considering the analogy between happiness and the flow experience, individuals may demonstrate a desire to experience flow (e.g., Vallerand, 2008). Indeed, Wu and colleagues (2013) argued that flow is sought through various outlets. However, individuals may develop a dependency on flow (Csikszentmihalyi, 1990) whilst aiming to mitigate existing dissatisfaction or emotional dissonance (Wan & Chiou, 2006).

The factorial nature of flow (Cowley et al., 2008) serves to distinguish it from other mental states. In order for an experience to qualify as flow, several criteria have been proposed to represent this state. People who experience flow often report loss of self-awareness, which causes the cessation of self-directed thoughts and rumination (Carpentier et al., 2012). This has been described as *hypo-egoic self-regulation* that benefits automatic behavior and control over the activity, and has clear connections to flow (Leary et al., 2006). During flow, the person is highly concentrated in the task and the allocation of attention is directed to a clear goal (Harris et al., 2017a). A byproduct that emerges through this intense concentration is a distortion of time perception (Csikszentmihalyi, 1990), where our internal clock seems to desynchronize from our casual perception of time. The perception of time has been suggested to be contextual and malleable depending on internal circumstances, such as the emotional state of the person (Droit-Volet & Gil, 2009). As such, time perception undergoes a transformation under the effects of different emotional states. An activity that is experienced as being tedious or frustrating instils the illusion of time moving slowly, whereas if the individuals enjoy what they are doing, time feels like it passes quickly (Agarwal & Karahanna, 2000; Ozkara et al., 2017; Sackett et al., 2010). As a result, time perception distortion in video game playing can be positive, i.e., engender relaxation, or negative, in which case it can generate guilt (Wood et al., 2007).

During the activity, the individual should be given feedback about their proximity to their goals and how well they are performing (Lyons, 2015; Nakamura & Csikszentmihalyi, 2014). Performance is mediated by a merging of action and awareness, which gives rise to automatic execution of the task (Jackson et al., 2001). Although automaticity has been questioned for its proneness to errors (e.g., Toner et al., 2015), it enables an economical allocation of resources to mitigate cognitive load (see Feldon, 2007; Moors & De Houwer, 2006). Critically, the activity should be manageable with a challenge that is compatible with the individual's skills (Rani et al., 2005). Such a compatibility is regarded as the chief prerequisite for flow's onset (Csikszentmihalyi, 1990; Llorens et al., 2013; Mirlohi et al., 2011) and it gives rise to a sense of control (Ghani et al., 1991). Finally, the activity should be autotelic (Csikszentmihalyi, 1990), meaning that performing it is intrinsically rewarding without the expectation of external rewards (Csikszentmihalyi, 1997a).

2.2.2 The problem of defining flow

The nine criteria that synthesize the flow experience remain contentious with respect to their individual capacity for triggering flow. For example, Cairns and colleagues (2014) have suggested that flow is an "all-or-nothing" experience, during which the individual either fulfills all the criteria for flow to kick in, or they do not, in which case flow will not come into effect. However, this position has not been widely accepted (e.g., Arzate & Ramirez, 2017; Heo et al., 2010; Ozkara et al., 2017; Quinn, 2005). Treating flow as an overarching construct may be a rather extreme approach in that certain experiential components may not be met, without necessarily implying the absence of flow (e.g., Csikszentmihalyi, 1990). On the other hand, extracting a subset of the criteria selectively may be arbitrary and can have significant theoretical implications. For example, the selection of such a subset presupposes knowledge over which criteria constitute the necessary conditions versus the sufficient conditions for flow to trigger (e.g., Keller & Bless, 2008; Swann et al., 2012), which creates a circular problem in defining flow.

To amend the dimensionality problem of the flow construct, multiple frameworks have been devised that segregate flow's criteria into antecedents, dimensions and outcomes (e.g., Nah et al., 2014; Hoffman & Novak, 1996). The antecedents are preconditions that facilitate the onset of flow, the dimensions are the core attributes of the flow experience, and the outcomes are the results of this experience (e.g., Hoffman & Novak, 2009). However, even this approach has stirred controversy. For example, autotelicity's twofold meaning (Michailidis, Balaguer-Ballester, & He, 2018) has been interpreted as a dimension of flow (e.g., Kiili et al., 2014; Van den Hoot et al., 2018), but also as an outcome of flow (e.g., Quinn, 2005; Rodríguez-Sánchez et al., 2011), whereas others have argued that this criterion should not be used for defining flow (Engeser & Schiepe-Tiska, 2012). Likewise, the sense of control has been viewed as an antecedent (e.g., Hoffman & Novak, 2009; Kiili et al., 2012) and as a dimension (e.g., Nah et al., 2014).

Other studies have embedded additional criteria, which are not listed in the original theory of flow. Such examples include telepresence (Shin, 2006; Skadberg & Kimmel, 2004) and playability (Kiili et al., 2014). However, the associated risk with the inclusion of criteria that may be unique to specific contexts is that those frameworks may be inapplicable beyond the domain for which they were developed. Although special activity characteristics may be effective facilitators of flow, they should not be entangled with the actual experience (Csikszentmihalyi, 1997a; Quinn, 2005), as they can distract us from its structural composition. Notably, some of those criteria may be generalizable to most activities, thereby fulfilling flow's alleged plausibility of a state in multiple activities (e.g., Nakamura & Csikszentmihalyi, 2014).

These include creativity (Ghani, 1995), improved learning (Ghani, 1995; Hoffman & Novak, 2009; Kiili et al., 2012) and increased performance (Jackson et al., 2001; Stavrou et al., 2007).

From these observations, it becomes clear that there is disagreement about the ontology of flow (Rodríguez-Sánchez et al., 2008). Notwithstanding flow's ill-defined construct due to its connection to multiple experiential components (Hsu & Lu, 2004), most authors follow a generative approach instead of a reductive, which entails the addition of characteristics in flow's conceptualization. Nevertheless, as mentioned earlier, these approaches may be intended for and tailored to specific contexts. However, Rodríguez-Sánchez and colleagues (2008) identified three criteria that most commonly appear in the flow literature, i.e., deep concentration, enjoyment and autotelicity.

While it is tempting to reduce the flow construct to three dimensions, it is questionable whether it can represent the richness of this experience. Perhaps, a more promising way to identify the experiential components of flow is through a retrospective assessment of its physiological responses. Manipulating task difficulty, immediacy of feedback and clarity goals, which clearly act as antecedents to the flow experience (Fullagar et al., 2013; Rodríguez-Sánchez et al., 2008), may expose signatures of flow that can be evaluated based on existing knowledge of the underlying physiological functions. For example, activation in the zygomaticus major facial muscle, that has been linked to the emotional expression of smiling (e.g., Epstein, 1990; Schmidt et al., 2006), has been interpreted as enjoyment during flow (e.g., Nacke & Lindley, 2008).

2.2.3 The relationship between identity formation and flow

Experiencing flow during activities in which the individuals partake seems to promote habitual engagement (Chou & Ting, 2003) as a function of time and energy invested therein (Emmons, 1999). Habit formation may be moderated by the internalization of the activity within one's own identity (e.g., Mao et al., 2016). For instance, an individual who derives satisfaction from painting may eventually identify as a painter, whereas one who finds pleasure in sports may identify as an athlete (e.g., Vallerand et al., 2003).

Linguistically, identification statements take the form of "I am…" and are accompanied by a subjective complement, which express identity as a perception of oneself (McCall, 2003). Self-identification not only does it serve to ascribe a meaning to oneself (Stryker & Burke, 2000), that is inherently egocentric, it may also transform into a social identity (Mao et al., 2016). By extension, social identities may satisfy the need for belonging (Greenaway et al., 2016) as individuals increasingly adopt in-group norms and behaviors (Thorbjørnsen et al., 2007). However, not all activities are equally flexible in being socially normative. For example, media entertainment use pivots on offering enjoyment (Vorderer et al., 2004), but the specific subcase of video game play seems to have been subjected extensively to social stigma (Szablewicz, 2010; Yates & Littleton, 1999).

In this section, we reviewed the flow experience. While there are definitional problems with this construct, it becomes increasingly clear that flow manifests through an array of cognitive phenomena during an activity. The diversity and abundance of interpretations surrounding the same construct stem from a correlational perspective. This perspective does not diminish the importance of the existing literature; instead, it underlines how the activity's characteristics may increase or decrease the likelihood of experiencing flow.

The consequence, however, is that flow may heavily depend on the activity, even though Nakamura and Csikszentmihalyi (2014) suggested that it is similarly experienced across multiple contexts and individuals. We also noted that the activity may be internalized by the individual as part of their identity, thereby exposing a potential circuit at play: activity-individual-flow. Thus, for flow to be experienced, both the activity and the individual should maintain some level of cohesion and compatibility. We explore this notion in the following sections and address this within a facet of media enjoyment, i.e., video game play.

2.3 "Synthetic" worlds: The case of video games

Earlier research instigated the debate on the effects of video game play, focusing on three major points: violence and aggression, addiction and social isolation (e.g., Anderson & Ford, 1986; Ferguson, 2010; Kirsch, 2006; Setzer & Duckett, 1994; Silvern & Williamson, 1987). These longstanding concerns were inherently social and helped bring video games in the spotlight rapidly (Williams, 2003). Soon after, video games encroached on the average teenager's routine (e.g., Klimmt & Hartmann, 2006) and the extant criticism was complemented with research aiming to identify the positive side of video game play (e.g., Granic et al., 2014; Green & Bavelier, 2003; Griffiths, 2002).

2.3.1 Understanding the controversy behind game playing

In spite of the debate remaining heated to this day, researchers started espousing the notion that video games have more to offer than mere consequences (e.g., Buckley & Anderson, 2006; Emes, 1997). This stance was also reinforced by the insertion of video games in education, work, rehabilitation, training and clinical settings as tools whose purpose was not to solely entertain (e.g., Kirriemuir & McFarlane, 2003; Ritterfeld et al., 2009; Spring, 2015). Naturally, the introduction of video games in these domains meant that some form of resolution had commenced when counterevidence started surfacing (e.g., Funk, 1993; Griffiths, 2000). Yet, the interpretation of video games as simulations of reality (Calleja, 2010; Grodal, 2000) begs the question whether enjoyment derived from video games is similarly simulated or "synthetic" (e.g., Castronova, 2005). In other words, are video games able to deliver positively arousing experiences just as other activities do?

Contrary to other media forms, the appeal of video games lies in their ability to interact with and within a mediated world (Granic et al., 2014). The gratification from such an interaction is allegedly higher in real-life scenarios than in their computer-mediated counterparts (Sacco & Ismail, 2014). Individuals may resort to computer-mediated environments in order to be entertained (Bartle, 2004; Green et al., 2004) and to satisfy their psychological needs (Przybylski et al., 2009a; Tamborini et al., 2010). However, earlier research pivoted on negative outcomes of video game play and, as such, promulgated the conception that video games are inherently addictive (King et al., 2010). More recently, habitual video game play and amount of time expended was positively linked to self-reported well-being (Johannes et al., 2021).

Although video games have largely been discredited in the literature, other activities have not historically undergone similar scrutiny. As an example, studies have traditionally treated sports as an activity conducive to psychological well-being (e.g., Snyder & Spreitzer, 1974;

Steptoe & Butler, 1996). Yet even sports are not without controversy. Recent research has demonstrated that sports can also become addictive (e.g., Freimuth et al., 2011; Landolfi, 2013). At the same time, the emergence of e-sports, which refer to professional, competitive video game play with characteristics emulating sports (Hamari & Sjöblom, 2017), have allowed the parallelism between video games and sports to be made (e.g., Heere, 2018).

The criteria to qualify an activity as addictive may be mediated by several factors, rather than the mere number of hours expended within (e.g., Lemmens et al., 2009; Peters & Malesky, 2008; Wood & Griffiths, 2007). For instance, a recent account that studying can take the form of study addiction (Atroszko et al., 2015) testifies that virtually any activity could devolve into being addictive. Indeed, working, that has also been shown to engender the flow experience (e.g., Csikszentmihalyi & LeFevre, 1989), has been colloquially referred to as 'workaholism' when done excessively (e.g., Andreassen et al., 2011). All of the aforementioned activities seem to have a common denominator: the flow experience. It should be noted that the current focus on addiction is due to the wide assumption that addiction is counterproductive to happiness and well-being (e.g., Akın, 2012; Çardak, 2013; Longstreet & Brooks, 2017).

2.3.2 Putting a positive spin: playing as a "positive" addiction

A useful consideration in this case is the dualistic model of passion, i.e., harmonious and obsessive passion (Vallerand et al., 2003). Flow as an optimal state (Csikszentmihalyi, 1990) seems to be similar to harmonious passion; engaging in an activity is volitional and other activities are not sacrificed in the process (Vallerand et al., 2003). On the other hand, addiction resembles an obsessive passion, whereby the individual engages in an activity because they feel compelled to do so (Vallerand et al., 2003). Addiction may thus occur when the activity is prioritized over daily activities and the person loses control (Beard & Wolf, 2001; Hellman et al., 2013; Vallerand et al., 2003). However, if daily activities are not sacrificed in the process (activities are in harmony with each other) and the individual regulates the frequency of performing the activity, then it is likely that negative affect is suppressed and positive affect is fostered (Vallerand et al., 2003). Taken together, decreased negative affect and increased positive affect may be the determinants of happiness (e.g., Diener & Larsen, 1993; Kahneman, 1999).

The dualistic model of passion has also been empirically contextualized in video games (e.g., Lafrenière et al., 2009; Przybylski et al., 2009b, Utz et al., 2012; Wang et al., 2008). Studies have consistently found a link between video game play as a harmonious passion and life satisfaction, positive affect and psychological well-being (Lafrenière et al., 2009; Przybylski et al., 2009b). Contrarily, video game play as an obsessive passion was identified as being pertinent to negative affect during and after video game play, excessive amount of time devoted to playing (Lafrenière et al., 2009; Przybylski et al., 2009b) and low quality of real-life social interaction (Utz et al., 2012). More recently, Fuster and colleagues (2014) found that exploration, socialization and achievement were characteristic motives of video game play as a harmonious passion that has implications in real life. For example, an extraverted individual may be equally sociable online as well as offline (e.g., Tosun & Lajunen, 2009).

Based on these accounts, it may not be an overstatement to suggest that video games are potentially platforms to happiness (e.g., Green et al., 2004; Johannes et al., 2021; Ryan et al., 2006). Although Castranova (2005) dichotomized real life from virtual – or, as he proposed, "synthetic", it appears that video game play is a qualified activity for enjoyment. Blaszczynski (2008) argued that the individual might be the primary suspect in the emergence of addiction in video game play rather than the content of the games. However, the structural characteristics of video games that promote them as outlets to enjoyment should not be overlooked.

This section reviewed the origins of the controversy behind video game play. Based on earlier considerations, it appears that video games have been criticized for their content as well as their capacity in inculcating obsessive play. However, the riposte from studies highlighting beneficial effects is not to be dismissed. More so, it appears that a fine line between healthy and unhealthy game playing may be defined, though it is not well understood how this transition can occur. Nevertheless, the flow experience seems to modulate repetitive play, which further highlights its necessity in video games. These observations suggest that flow can be depended on for sustainable business models in the games industry. At a lower level, game designers bind design elements that synthesize their game and motivate their audience(s).

2.4 Video game design: A window to flow

Playing video games seems to satisfy flow's criteria (Cowley et al., 2008; Sherry, 2004). Flow's relationship with gaming action (Weibel & Wissmath, 2011) suggests that the gameplay as well as the game mechanics are at the core of triggering flow. Notably, gameplay and game mechanics are distinct concepts (Desurvire et al., 2004). Gameplay refers to the interaction between the player and the game world as well as the outcomes resulting from the enactment of said interaction (Fabricatore, 2007). For example, the goals and challenges presented to the player are a facet of gameplay (Federoff, 2002). In contrast, game mechanics pertain to the *ways* this systemic interaction is achieved.

Unlike gameplay, game mechanics are discovered gradually instead of being directly known to the player (Adams & Dormans, 2012). Furthermore, game mechanics constitute the instrumental blocks that make gameplay possible (Adams & Dormans, 2012). Albeit the design alone may be conducive to flow, it does not necessitate that every player will experience flow (Cowley et al., 2008). Hence, both the game and the player should present characteristics that are compatible with flow (e.g., Finneran & Zhang, 2003; Keller & Bless, 2008; Whitlock et al., 2014) and the interaction of the two might determine whether and how the player will experience flow (e.g., Chen, 2007; Nah et al., 2014). Thus, the structural characteristics of video games may explicate the ways in which flow is fostered.

2.4.1 Reward systems

Reward systems are a powerful tool to maintain a player's interest in the game. They are disguised abstractions of the Skinner box (Yee, 2001), which aim to hook the players for extended periods (Griffiths & Wood, 2000). Essentially, they are operationalized forms of reinforcement and may satisfy intrinsic, psychological needs in game playing (Hsu et al., 2009). Transposing Maslow's (1943) hierarchy of needs to rewarding mechanisms that abound video games, helps clarify the drive behind player commitment. Unlike the low-order needs that are biologically significant, higher-order needs are more resistant to satiation (O'Donnell & Epstein, 2019) albeit reward omission may also precipitate satiation to higher-order needs (e.g, Morgan, 1974; Ronimus et al., 2014).

Previous empirical research demonstrated that reinforcement is most effective when the rewards are obtained contingently or in specified intervals (Ferster & Skinner, 1957). Yet, compensatory rewards for overcoming a challenge may amplify player enjoyment and sense of progression in the game (Griffiths & Nuyens, 2017; O'Donovan, 2012). It is worth noting that player preference over the consistency of reward provision, i.e., variable or fixed reward provision, may vary upon the genre of the game (Kao, 2012).

2.4.1.1 Symbolic and fiscal rewards

Existing taxonomies on reward types in video games may have been developed without empirical support (Phillips, 2018) or they may be genre-dependent (e.g., Hallford & Hallford, 2001). However, two broad categories that seem to encompass the majority of reward types are *symbolic rewards* and *fiscal rewards*. Symbolic rewards in games are geared toward the self; they have a personal meaning to the player and can satisfy personal needs. Thus, they can be referred to as *ego-rewards* – an umbrella term Emerson (1962) used for status recognition. Conversely, fiscal rewards are more relevant to the progression in the game (e.g., Phillips et al., 2013) and could be better characterized as *game-centric rewards*.

Albeit game-centric rewards can be self-contained sources of satisfaction, ego-rewards may circumstantially depend upon them. For example, a player who is able to afford top-tier gear may be able to increase his or her status due to their visual appearance being suggestive of a high rank (e.g., Consalvo & Harper, 2009; Ducheneaut et al., 2006; Tosca & Klastrup, 2009). In this context, game-centric rewards may influence the desirability of a player in becoming a member of a virtual community (Chen et al., 2008), which can promote a stronger sense of belonging (Kang et al., 2009). Similarly, virtual clothing accessories, which usually serve to enhance an avatar's outer appearance (Dickey, 2007), may be used as devices to promote inclusion in social contexts and to reflect personality traits of the player's idealized self (Banakou et al., 2009).

2.4.1.2 Ego-rewards to appease intrinsic motivation?

Based on the distinction between game-centric and ego-rewards, flow's autotelicity seems to be compatible with ego-rewards, if flow is to be regarded as a personal need (e.g., Baumann

& Scheffer, 2011). Yet, unlike the elevation of in-game social status as an ego-reward, which is mostly relevant within the game domain, flow's autotelicity has an effect beyond the game's world, since enjoyment succeeds the activity (Csikszentmihalyi, 1996; Peterson et al., 2007). The observable difference, in this case, may relate to a known distinction made in the literature, namely the extrinsic and intrinsic motivation. Extrinsic motivation alludes to the expectation of a reward, i.e., the acquisition of a reward becomes the primary motivator (Cameron et al., 2001). In contrast, intrinsic motivation, which has been associated with the flow experience, portrays engagement in an activity as being inherently pleasant (Singh et al., 2005). Hence, albeit both autotelicity and social status may qualify as ego-rewards, the former is intrinsic but the latter is extrinsic (e.g., Griffiths & Nuyens, 2017). Notably, rewarding mechanisms are widely accepted to be associated with the dopaminergic system (Schultz, 2002), the same system that has been implicated in the flow experience (e.g., Weber et al., 2009).

Extrinsic rewards are allegedly less effective in stimulating motivated behavior compared to intrinsic rewards (Garris et al., 2002; Schmid, 2011). However, their distinction from intrinsic rewards advocates the existence of two categorical incentives for playing (e.g., Griffiths & Nuyens, 2017; Liu, 2017). On the one hand, intrinsic rewards may enable behavioral repetition, which is manifested through replay intention across different instances of play (e.g., Bakker, 2005; Keller et al., 2011b; Wiedemann et al., 2014). On the other hand, extrinsic rewards seem to be mostly associated with maintaining the player's motivation within the same play session, thereby prolonging playing (King & Delfabbro, 2009). Bearing in mind that no activity is purely intrinsic (Csikszentmihalyi, 1975, 1999; Hidi, 2000), the combination of intrinsic and extrinsic motivators may cater for a stronger motivation overall (Cameron et al., 2001; Deci & Ryan, 1985). However, designing gameplay with an emphasis on one over the other type of reinforcement may prompt adverse effects. For example, an individual may grow dependent on the anticipation of extrinsic rewards, thereby causing the significance of intrinsic motivation to be diminished (Csikszentmihalyi, 1975; Medler, 2011; Murayama et al., 2010). Through this process, the motivation to engage in an activity may perennially default to external causes instead of an intrinsic desire (Deci, 1971; Deci et al., 1999). However, this view has been challenged recently in the context of video games (e.g., Cruz et al., 2017). In addition, verbal reinforcement of the individual may not necessarily shift the intrinsic value to extrinsic (see Albrecht et al., 2014), which is also used in video games as a means of positive reinforcement (e.g., Sanchez et al., 2017).

2.4.1.3 Flow as an amalgam of intrinsic and extrinsic rewards

The theory of flow presupposes that playing is an intrinsic reward (Csikszentmihalyi, 2014). At the same time, certain activities within or qualities about the game may also be intrinsically motivating without entailing the provision of a reward, for instance exploration, fantasy and curiosity (Griffiths & Nuyens, 2017; Malone, 1981; Phillips et al., 2013). However, game-centric rewarding systems, and even ego-rewards, as mentioned earlier, can be extrinsic. Thus, the aforementioned observations are seemingly antithetical to flow's autotelicity. If playing is intrinsic, how do extrinsic rewards tap into flow? Cowley and colleagues (2008) subsumed reward systems under the immediacy of feedback dimension of flow, whereas Nah and colleagues (2014) did not acknowledge in-game rewards in their flow framework applied into video game playing. However, viewing game reward systems as mere feedback may be a reductive approach.

To demonstrate using an example, games featuring quests introduce new goals to the player (Ashmore & Nitsche, 2007). The reward for completing quests is typically communicated to the player before he or she embarks on the deliverable task. In addition, rewards are proportional to the difficulty of the task: a task that is harder to achieve offers a higher incentive than a challenge that is easy to tackle (Hunicke, 2005; Wang & Sun, 2011). Consequently, goal direction, feedback and skills-challenge compatibility seem to all be implicated in the design of extrinsic rewarding mechanisms (Wang & Sun, 2011). This indicates that game rewards are diffused across several facets of gameplay and are consequently entrenched in the game experience.

This section demonstrated that rewarding mechanisms in games are complementary to the self-rewarding experience that flow has to offer. Due to the widespread understanding that reward processing is strongly connected to motivated behavior (Mitchell, 1982), designing effective rewards, that players will vie for, requires understanding of what motivates players. In addition, increases in the reward magnitude have been shown to proportionally increase player enjoyment (e.g., Johnson et al., 2018). However, the relationship between reward magnitude and player enjoyment may be mediated by the consistency of the reward's provision.

For example, Kao (2012) claimed that fixed-schedule rewards with an extensive range of values might be an optimal rewarding strategy. Indeed, rewarding the player as a means of acknowledging their effort may encourage them to invest more effort in the future – a concept that ties to operant conditioning (e.g., Hilgard et al., 2013). However, in spite of the differentiation between extrinsic rewards and flow as an intrinsic reward, both seem to share the potential outcome of addiction (e.g., Chou & Ting, 2003; Csikszentmihalyi, 1990), which may have serious implications for the individual's happiness (Hull et al., 2013). While Johnson and colleagues (2018) found high rewards to be more enjoyable, their cost may be pathological reliance upon them (Hsu et al., 2009). Although addiction may be subject to individual proneness (Hsu et al., 2009), game designers should implement balanced reward schedules and magnitudes.

2.4.2 Challenge

Challenge is widely apprehended as a strong predictor of game enjoyment (Klimmt et al., 2009; Sherry, 2004). Hardy and colleagues (2014) further claimed that measuring challenge might be an indicator of the player's motivation. Though ultimately the sense of challenge is in essence a subjective perception, it is generated by the objective expenditure of mental and physical effort (Hsu et al., 2007). Thus, challenge can be both subjective and objective (Gaillard, 1993; Gastorf, 1981). For example, on a sensory-cognitive level, a task can be objectively challenging, when it overloads the working memory with information (Klingberg, 2009; Lavie, 2005; Paas et al., 2004). When such content is amply available, the individual might become susceptible to distraction, anxiety or mind wandering (Forster, 2013; Kane et al., 2007; Klarkowski et al., 2015). Thus, in the case of mental overload, the mobilization of attentional resources is compromised and individual performance on the task may exacerbate (He et al., 2011; Ziakopoulos et al., 2017).

On the other hand, physical exertion may also be overwhelming. Physical exertion in video games may be challenging not only as a result of the amount of physical energy expended, but also proportionally to the number of options available for the user to execute (e.g., Chen, 2007). It should be stressed that physical exertion, in this context, is not to be confused with exergames, i.e., video games promoting physical exercise (e.g., Yang et al., 2008). Instead, physical exertion refers to physical activity, which is distinct from exercise in that it encompasses any coordinated movement of the muscles (Caspersen et al., 1985; Oh & Yang, 2010). Traditional video games are more pertinent to physical activity, which is realized via handling an input device that establishes the game-player interaction. Moreover, the type of controller used may impinge on the perceived sense of control over the game (Limperos et al., 2011).

In addition, task novelty is likely to stimulate higher cognitive load, as it requires explicit control, whereas familiar tasks are processed in an automated and implicit way (Hodent, 2016; Speelman & Maybery, 2013). Overcoming the initial challenge of handling physical controls is key to progressing into deeper levels of immersion (Brown & Cairns, 2004). This is also reflected in flow theory, which proposes that action execution feels automatic and not consciously processed (Singer, 2002). Though automaticity has been predominantly interpreted as a dimension of the flow experience (e.g., Nah et al., 2014), Quinn (2005) attested to its importance by equating it to the flow experience. However, evidence from a few studies do not corroborate Quinn's interpretation and instead showed that other criteria may be more characteristic of flow (e.g., Klasen et al., 2012; Sillaots & Jesmin, 2016; Skadberg & Kimmel, 2004). Yet, automaticity is still important in that it may help explain the alleged individual performance advantages during flow (e.g., Stoll, 2018) as well as the diminution of the adverse effects of mental workload, such as anxiety (Cottrell & Barton, 2013).

2.4.2.1 Challenge, Performance and Enjoyment

In order to optimize performance and to amplify enjoyment, a balance between the player's skills and the game's demands should be met, which in turn provides fertile ground for flow to trigger (Csikszentmihalyi, 1990). The Four Channel Flow Model (Massimini & Carli, 1986) identifies four experiential outcomes resulting from compatibility or incompatibility of the skills-challenge pair. When a balance is achieved, players are more likely to experience flow, whereas an incompatibility thereof may engender boredom (low challenge and high skills) or anxiety (high challenge and low skills), which result from the perception of an easy or a hard game respectively. Although anxiety and boredom are traditionally interpreted as the polar opposites of flow and as mutually exclusive (e.g., Csikszentmihalyi, 1990; Jackson et al., 1998), skills-challenge incompatibility alone does not guarantee absence of flow, but rather a weak flow experience (Fong et al., 2015).

It should be stressed that possibly only one study has been unable to empirically show that a fit between skills and challenge gives rise to flow (e.g., Løvoll & Vittersø, 2014). The authors of this study suggested that perfect balance might actually result in boredom rather than flow. Finally, apathy may occur when both the player's skills and the challenge are low. Taken together, albeit apathy is made of compatible conditions (low-low), which would imply a flow experience, it is instead associated with the absence of motivation to engage with the activity altogether (Stavrou, 2008). On the basis of these observations, the game play experience appears to be highly influenced by the challenge posed by the game (Mathwick & Rigdon, 2004).

2.4.2.2 Personal characteristics influence challenge perception

Notably, the same task characteristics may bestow differential mental load to different players. For example, expertise in a task may enable the individual to cope more efficiently under mental load (e.g., Benner, 1984). Indeed, if the perceived difficulty of the task is not considered in context with the player's skills, it may not be intuitively comprehensible. An earlier study argued that it is not clear whether expert and novice players experience flow similarly (e.g., Kirschner & Williams, 2014). This may be because skilled players expect a higher level of challenge compared to novice players (Berta et al., 2013; Chen, 2007; Cox et al., 2012; Iacovides et al., 2015; Klimmt & Hartmann, 2006).

Novice players in training studies using action video games have been found to have benefited from enhancement of attentional capacity whereby a higher number of visual objects are permitted for concurrent processing (cf. Hubert-Wallander et al., 2011). Likewise, increased working memory capacity observed in action video game players (McDermott et al., 2014) may entail increased resistance to distraction and mind wandering (Kane & McVay, 2012). Generally, player expertise may cultivate the ability to suppress distraction and negative affective states under mental overload (e.g., Borderie & Michinov, 2016; Bunian et al., 2018; Chisholm & Kingstone, 2012), which would diminish the perception of a high challenge.

Bandura (1982) proposed that individuals make personal judgments about their perceived likelihood of accomplishing their goals, which is described with the term self-efficacy. High-efficacy individuals tend to persist more on the challenges they face (Zimmerman, 1995), whereas low-efficacy individuals may avoid challenge altogether (Bandura, 1994). This is an important observation, which suggests that certain players may not enjoy challenge as much as others do (e.g., Klimmt & Hartmann, 2006). Paradoxically, challenge is precisely one of the principal appeals of video game play (De Schutter, 2011; Griffiths & Hunt, 1998).

Klimmt and Hartmann (2006) suggested that low efficacy could impinge on the intrinsic motivation to engage in video game playing. This same relationship has been shown for flow, where self-efficacy acts as an antecedent to flow (e.g., Rodríguez-Sánchez et al., 2011). Does this imply that low-efficacy individuals may prefer low challenge despite having high skills – a scenario that flow theory treats as boring? Presumably, as Lee and LaRose (2007) observed that, for flow to emerge, self-efficacy should match with the perceived challenge. Hence, a low-efficacy player may prefer low challenge, whereas a high-efficacy player may prefer high challenge. This corroborates the notion that highly skilled players will prefer high challenges in order to experience flow (e.g., Jin, 2012; Rodríguez-Sánchez et al., 2011). Importantly, self-efficacy has been shown to mediate the relationship between player performance and player enjoyment (Trepte & Reinecke, 2011).

Other prerequisites that may mediate one’s motivation to seek challenge have been shown to include personality traits, such as conscientiousness (Johnson et al., 2012; McMahon et al., 2012; Zeigler-Hill & Monica, 2015) and extraversion (Zeigler-Hill & Monica, 2015). Both traits have been related to the flow experience (e.g., Demerouti, 2006; Johnson et al., 2014; Tatalović Vorkapić & Gović, 2016; Ullén et al., 2012, 2016). Ullén and colleagues (2016) argued that conscientiousness is relevant to flow, as it relates to motivation, problem solving and emotional self-regulation. On the other hand, neuroticism is a trait that relates to a tendency in experiencing anxiety and negative emotions (Gray & MacNaughton, 2003), lack of motivation (Ryan & Deci, 2000b) and inactivity (D'Zurilla et al., 2011), and is therefore highly inconsistent with the state of flow. Other personality factors that presented a strong, positive relationship with flow are openness to experience, extraversion and agreeableness, which relate to well-being, an active lifestyle, and social desirability (Ullén et al., 2016).

2.4.2.3 Achieving flow by achieving balance

From the previous section, it becomes clear that video game designers need to account for individual differences and accommodate different player archetypes. Previous studies in human-computer interaction have used the umbrella term *player modeling* that acknowledges the role of individual differences in the video game experience (e.g., Charles & Black, 2004; Yannakakis et al., 2013). The purpose of player modeling is to dynamically model the player's idiosyncrasies, as they are enacted through in the game world, in order to augment game enjoyment (e.g., Yannakakis et al., 2013). Challenge adaptation seems to be a qualified instance of player modeling (e.g., Missura & Gärtner, 2009). Anecdotally, balancing the game's challenge is a particularly challenging task for game designers (Moroşan, 2019).

Hence, another option would be to dynamically adjust the game's difficulty (e.g., Van Lankveld et al., 2009; Zohaib, 2018), which has been shown to be more immersive than predefined difficulty settings (Denisova & Cairns, 2015a). Hunicke (2005) found that when a game is adjusted to the player's skills, their performance increases. Based on our previous observations, this increase in performance may substantiate a more enjoyable experience. Dynamic difficulty adjustment is substantially better than a conventional approach of static difficulty, in that it also accounts for the player's perceived difficulty (e.g., Adams, 2009; Hunicke, 2005). Furthermore, dynamic difficulty scaling may encourage long-term commitment with the game (Youssef & Cossell, 2009). Although different authors have formulated methods for classifying players on the basis of their expertise (e.g., Ip & Adams, 2002; Manero et al., 2016), it may be more effective to adjust difficulty "on the fly" rather than based on a cluster of characteristics that correspond to a class of player expertise.

To conclude, designing interesting challenges is essential to game enjoyment. "Interesting" challenges should be able to stir curiosity (e.g., Kashdan & Silvia, 2009), introduce unexpected events (e.g., Brandtzæg et al., 2018), feature fantasy elements (e.g., Kannetis et al., 2009) and be either intrinsically (e.g., a challenge combined with exploration of an innately inaccessible area) or extrinsically motivational (e.g., provision of symbolic or fiscal rewards). Yet, as previously discussed, the same challenge may not be universally pleasant to different players. Hence, challenge should be defined in relation to the player's skills. Notably, challenges may differ based on the solicited cognitive and executive demands. For example, increasing the health of the enemy units as a tactic to boost challenge is different from increasing the frequency and the number of enemies spawned at a time. The former scenario requires efficient character resources, which may entail an arduous process to acquire, while the latter requires quick reflexes and agility on the player's side. These two scenarios are not mutually exclusive, but the nature of the challenge they pose is different.

2.4.3 Flow's "competition" with other states

The skills-challenge compatibility component that has been treated as an antecedent to the flow experience (e.g., Csikszentmihalyi, 1990; Fong et al., 2015) and violations of which might cause the suspension of flow, suggests that flow remains susceptible to task difficulty alterations. Earlier research showed that the anterior cingulate cortex strongly responds to such changes in task difficulty (e.g., Barch et al., 1997; Paus et al., 1998). This region has been argued to comprise an affective subdivision, which forms projections to brain areas associated with affective processing and experience (subgenual anterior cingulate cortex; Vogt, 2005).

Thus, challenge perception seems to be associated with emotional responses on an anatomical level, which corroborates the proposed outcomes of the Four Channel Flow Model (Massimini & Carli, 1986). Apart from challenge, affective responses in video games may also result from the game's narrative (e.g., Schneider et al., 2004) and its reward systems (e.g., Mathiak et al., 2011). Affect has long been viewed as the composition of two dimensions – valence (positive/negative) and arousal (low/high) (Lang, 1995). Building on Lang's two-dimensional affective space as a guide to flow, the following can be surmised: boredom is a low-arousal negative state (Mikulas & Vodanovich, 1993; Raedeke & Stein, 1994), anxiety is a high-arousal negative state (Raedeke & Stein, 1994), whereas flow is a moderate/high-arousal positive state (Nacke & Lindley, 2008; Peifer et al., 2014; Ullén et al., 2010).

2.4.3.1 A boring game cannot be immersive at the same time

Boredom has been consistently correlated with negative affect (e.g., Gordon et al., 1997; Vodanovich, 2003; Vodanovich et al., 1991). It is considered a state that can emerge as a failure to allocate attentional resources in spite of an existing motivation to engage (e.g., Danckert et al., 2018; Eastwood et al., 2012). Hence, concentration is much more reduced during boredom than during flow (Mathwick & Rigdon, 2004). Traditionally viewed as a long-lasting state that characterizes the individual's life (e.g., Harris, 2000), boredom presents a negative relationship with life satisfaction (Farmer & Sundberg, 1986).

Researchers have posited that boredom can be conceptualized as a trait, suggesting that certain individuals may be more prone to experiencing boredom than others (e.g., Vodanovich, 2003). However, "boredom proneness" is a relevant topic when the individual experiences boredom across multiple contexts and times (Elpidorou, 2014; Watt & Hargis, 2010). This has also been addressed in the literature, which distinguishes *emotion* from *mood* (e.g., George, 2011; Warr et al., 2014). For example, boredom may occur as a response to a particular situation but can also take the form of a prolonged state over a period of time, in which case there is not a direct trigger for that state (Russell, 2003a). In video game play, we refer to boredom as a response to the interaction with a computer game, whereby various triggers in the game can facilitate the onset of this state.

Certainly, individuals with boredom proneness may be less likely to experience flow (Harris, 2000). However, in the case of video game playing, we should refer to boredom as a situational state, which makes no assumptions about the individual's general tendency to experience boredom. Such scenarios may be externally driven (see Goldberg et al., 2011) and involve task repetitiveness (Smith, 1981), prolonged and monotonous stimulation (Hill & Perkins, 1985; O'Hanlon, 1981), and, as seen in the theory of flow, low challenge (Csikszentmihalyi, 1990; Massimini & Carli, 1986).

Notably, a less popular consideration is boredom as a *high*-arousal negative state in that it acts as a self-motivated state to pursue new goals (Bench & Lench, 2013; Berlyne, 1960). Bench and Lench argued that its conventional conception as a low-arousal state originates from a misconception between boredom and apathy. As mentioned earlier, the Four Channel Flow Model (Massimini & Carli, 1986) treats boredom and apathy as two separate experiential outcomes. However, Mikulas and Vodanovich (1993) suggested that high arousal automatically suggests the absence of boredom. In a study by D'Mello and colleagues (2007) it was shown that, during instances of boredom, the users of a tutor-simulated system tended to lean back, away from the computer. On the contrary, flow episodes were characterized by a tendency to lean forward, closer to the computer. The authors concluded that boredom's postural pattern was indicative of an inclination to disengage, which can be considered a coping strategy to alleviate the effects of boredom (e.g., Smith, 1981).

These findings corroborate an earlier view that boredom is manifested as disinclination to action (Greenson, 1953), which is in direct opposition with the idea that boredom prepares for action (Bench & Lench, 2013). Goetz and colleagues (2014) identified five types of boredom. Bench and Lench's definition closely resembles the "reactant boredom" (Goetz & Frenzel, 2006), which is a combination of high arousal and negative valence that motivates the individual to "leave the situation for specific alternatives" (Goetz et al., 2014, p. 413). Contrarily, the "calibrating boredom" (openness to distraction from the current situation) and the "apathetic boredom" typify the common conception of boredom as a low arousal state (Goetz et al., 2014). Perhaps, qualifying boredom, as a high-arousal state, requires more information about the context it takes place. A possible disambiguation between apathy and boredom in the affective space is that apathy may be viewed as a neutral rather than a negative state (e.g., Shirzad & Van der Loos, 2016). Again, it should be noted that apathy and boredom are not meant here as individual traits, i.e., internally driven (Goldberg et al., 2011), but as transient states.

2.4.3.2 The detriment of anxiety could also be beneficial

Likewise, anxiety is also considered a negative affective state (e.g., Raghunathan & Pham, 1999) and a personality attribute (Spielberger et al., 1971). It is worth noting that other terms have also made appearance in the literature, namely frustration (e.g., IJsselsteijn et al., 2007; Nuñez Castellar et al., 2016), stress (e.g., Nijholt et al., 2009) and overload (e.g., Keller et al., 2011a; Klasen et al., 2012). In some cases, frustration and anxiety have been used interchangeably (e.g., Craveirinha & Roque, 2010; Jin, 2012; Knox et al., 2011). Although frustration seems to be more similar to the basic emotion of anger (Triberti, 2016), Saadé and Kira (2009) grouped frustration, anxiety, anger and confusion together as similar affective states. While the differences in the phenomenology of these terms is beyond the scope of this section, anxiety and frustration will be preferred for their common appearance in the video games literature.

The connection of boredom and anxiety to negative affect (e.g., Diener & Emmons, 1984; Fisher, 1993; Raedeke & Stein, 1994) and flow's putative connection to positive affect has popularized the idea that flow will likely be suspended when a person experiences negative affect (e.g., Jennett et al., 2008; Poels et al., 2012). However, Csikszentmihalyi (1990, 1996) mentioned that a person might not always receive satisfaction whilst engaging in an activity, but rather immediately after (also, Ghani & Deshpande, 1994). In video games, both positive and negative emotions have been identified as valid constituents of the game-playing experience (e.g., Granic et al., 2014; Kaye et al., 2018; Silpasuwanchai & Ren, 2018).

These observations challenge the idea that flow is uniquely connected to positive affect (Walker, 2010). Thus, flow could also be portrayed as a high-arousal negative emotion, which clearly overlaps with anxiety. Interestingly, flow's competition with anxiety may not be easily distinguishable. For example, a few studies have been unable to find a significant physiological difference between flow and anxiety (e.g., Keller et al., 2011a; Larche & Dixon, 2020). While Keller and colleagues did not offer a proper justification for this finding, they stressed a distinction previously made in the literature – the positive and negative stress (Selye, 1975), alluding to a resemblance between flow and positive stress.

2.4.3.3 Flow meets anxiety and boredom

Still, the aforementioned claims do not elucidate the conditions under which frustration and boredom can replace the flow experience provided that negative affect does not pose a threat to flow. Yet, Klimmt (2005) contended that players who are unable to achieve a subset of challenges in the game will eventually give in to negative emotions. Anxiety and frustration may emerge as a result of failure (Gilleade & Dix, 2004; Saadé & Kira, 2009) as well as fear of possible failure in a task (Fullagar et al., 2013; Juul, 2009; Sarason, 1984; Schüler, 2007). However, the consequences of failure may also be negligible when extrinsic rewards are provided (e.g., Miller & Hom, 1990), which is a common design strategy in video games, as discussed earlier. Thus, failure alone does not seem to adequately explain how the transition from flow to a negative state can occur. This paradox may be clarified using a distinction, similar to Selye (1975)'s, between positive and negative frustration (e.g., Allison et al., 2015; Gilleade & Dix, 2004; Miller & Mandryk, 2016; Nylund & Landfors, 2015; Roest and Bakkes, 2015). Whereas positive frustration can maintain the player's motivation, negative frustration may be destructive to flow, leading to disengagement (Miller & Mandryk, 2016).

Juul (2013)'s observations, which stem from the attribution theory (e.g., Fiske & Taylor, 1991) and applied in video games, further elaborate on this scenario. Failure may cause differential effects on the emotional experience depending on whom the player will hold responsible for their failure. For example, the bipolar pair of external and internal failure suggests that, if failure is attributed to external causes, motivation may be suspended, whereas if it were attributed to internal causes (e.g., the player holds himself or herself responsible), the person may persist in the challenge (e.g., Child, 1994; Gilleade & Dix, 2004).

What this means for flow is that external failure (e.g., the game being too demanding) would place the source of failure outside the player's control, which indeed violates flow. In that case, the player loses their sense of control and its escorting pleasure (Klimmt et al., 2007). Moreover, Van den Hoogen and colleagues (2012) provided evidence for another attribution pair as a possible explanation for flow's suspension. The authors found that repetitive failure, which is described as stable failure (e.g., Juul, 2013), is likely to be experienced as negative, whereas unstable failure might preserve motivation (e.g., Child, 1994; Van den Hoogen et al., 2012). Likewise, flow may wane if the player is prematurely interrupted during playing, in which case negative affect (e.g., annoyance) is more likely to be experienced (e.g., Freeman & Muraven, 2010). Similar observations have been made for boredom – the person experiences boredom when the source of boredom is attributed to external factors (e.g., Mikulas & Vodanovich, 1993).

Nevertheless, is it tenable to view flow as being characterized by either positive or negative valence? Cacioppo and colleagues (1999) suggested that valence might not be as binary (negative/positive) as Lang (1995) originally proposed. Instead, they mentioned that negative and positive emotions are processed by different neural systems and thus do not undergo a functional limitation on concurrent processing. For example, positive and negative emotions may be, to an extent, functionally lateralized via left and right brain circuits respectively (Kahneman et al., 2003). This is also manifested in the approach-withdrawal theory, where approach behaviors, usually as a result of positive affect are lateralized to the left, and withdrawal behaviors, usually resulting from negative affect, are lateralized to the right (Davidson, 1992). Empirically, this has been widely known to be reflected through frontal alpha asymmetry (for a review, see Briesemeister et al., 2013).

Hence, even if flow is viewed solely as a state of positive affective valence, negative affect may not necessarily compromise it. For example, in the study of Borderie and Michinov (2016), it was found that players experienced flow even under failure. The authors further asserted that frustration may even be a signature of flow. This may also be viewed from the scope of duality, i.e., the simultaneous experience of emotions with opposite valence, which supposes that mixed emotions may be experienced if the person is willing to accept such an emotional dissonance (Williams & Aaker, 2002). In the same vein, Larsen and colleagues (2001) proposed that affective experiences be interpreted as bivariate rather than bipolar entities, considering that joy and sadness may be experienced concurrently (Menninghaus et al., 2015).

The evidence on the triad of flow, anxiety and boredom may be somewhat inconclusive. Though boredom seems to be a competitive state to flow's, the same is not necessarily true for anxiety and flow. There are authors, for example, who hold that anxiety may be a necessary component in games and further suggest that games can be both engaging and frustrating (Roest & Bakkes, 2015). This draws a blurry line from the conventional interpretation that anxiety supersedes flow (e.g., Massimini & Carli, 1986) and alludes to a paradox that is perhaps unique to video games. Indeed, the plethora of video games available narrow the possibility that anxiety and flow are unconditionally incompatible. However, the differences between these states may be more pronounced across different technological mediums. An example of such a medium is virtual reality, which is purported to engender life-like experiences.

2.4.4 Video games in virtual reality

Virtual reality (VR) allows users to engage in video game play, by enabling more vivid representations of reality than traditional displays (e.g., Pallavicini et al., 2018; Van Kerrebroeck et al., 2017). It has been considered a technological innovation of the century (Perry Hobson & Williams, 1995). Amongst the distinctive features of VR is that of *presence*, which elicits the illusion of self-location displacement (Sanchez-Vives & Slater, 2005), loss of body ownership (Krekhov et al., 2018) and a feeling of being present in the mediated world (Steuer, 1992). Although the sense of presence is not bound to head-mounted displays or CAVE displays, i.e., stereoscopic projections, which are typically associated with VR (Pausch et al., 1997), presence appears to be more pronounced in VR (Carroll et al., 2019; Pallavicini et al., 2019; Peng et al., 2019). This is achieved by whelming the user's sensory system such that the information flow emulates information encoding in real life (e.g., Mania & Chalmers, 2001). The proximity of the virtual world to the physical world facilitates suspension of disbelief and the virtual world replaces the real one on a perceptual level (Rizzo & Koenig, 2017; Saunders et al., 2011).

In comparison to 2D displays, VR seems to contribute to more arousing and intense experiences (Pallavicini et al., 2018, 2019; Tan et al., 2015) that greatly enhance the overall gaming experience (Tan et al., 2015). Indeed, user responses in an immersive mediated world may approximate real-life responses (Bohil et al., 2011; Feigenson, 2006), making VR particularly interesting to study or use as an alternative approach to field experiments (Bohil et al., 2011). The connection of presence to enjoyment (Shafer et al., 2019) in conjunction with high arousal reported by Pallavicini and colleagues (2018) during VR game play suggests that VR may engender highly enjoyable experiences. Presence is an experience that is purported to precede the flow experience (Michailidis, Balaguer-Ballester, & He, 2018; Weibel & Wissmath, 2011). Thus, virtual reality may act as a launch pad to the flow experience and facilitate the transition to this state. Notably, if user experiences in VR are greatly augmented (Shelstad et al., 2017), then VR can be a promising technology to investigate flow.

The obstruction of real-life visual stimuli, that is achieved with head-mounted displays, precludes them from distracting the user. Jennett (2010) found that during immersive episodes, susceptibility to distractions is significantly reduced, but the distractor, depending on its nature, may still undergo cognitive processing. For example, Jennett employed three categories of distractors, namely relevant to the person, relevant to the game, and irrelevant. The author found that irrelevant distractors were significantly suppressed while players engaged in a game that matched their skills, but distractors relevant to the person or the game did not undergo the same suppression as the irrelevant distractors did. It was thus surmised that a special attentional filter may be active during an immersive experience, when stimuli beyond the game's environment are not automatically discarded and could lead to potential user disengagement. On the basis of this presumption, head-mounted displays engulf the user's visual field, while headphones are used for audio transmission, thus making virtual reality a medium that minimizes the existence of environmental distractions (e.g., Izkara et al., 2007). As such, VR may enable higher sustainability of experiences which may not have been otherwise possible (Baştuğ et al., 2017).

This may also justify the commercial success of the VR platform – for example, *PlayStation VR* has sold over 4.2 million units since its commercial release (Shuman, PlayStation Blog, 2019). Other major platforms include the *Oculus* series (i.e., Rift, Go and Quest; Oculus VR, Irvine, California, United States), *HTC Vive* (HTC Corporation, Xindian, New Taipei, Taiwan & Valve Corporation, Bellevue, Washington, United States), *Google Cardboard* (Google LLC, Mountain View, California, United States) and *Samsung Gear VR* (Samsung Electronics, Suwon, South Korea). More recent developments have embedded additional features within the headsets' capabilities, such as eye-tracking technology in *FOVE VR* (Fove Inc., Torrance, California, United States), *HTC Vive PRO, Varjo VR-1* (Varjo HQ, Helsinki, Finland) and *Looxid VR* (Looxid Labs, Gangnam-gu, Seoul, South Korea) and electroencephalography (EEG) sensors in Looxid VR and Neurable (Neurable Inc., Boston, Massachusetts, United States).

The surge of these technologies indicate that the user is promoted to an active participant of the virtual reality experience rather than a passive receiver of stimulation. Though there are studies wherein the EEG power is measured in the context of adaptive games (e.g., Ewing et al., 2016), similar investigations for the combination of VR, EEG and flow are scant (e.g., Monteiro et al., 2018). The commercial availability of these products provides evidence that the industry is becoming increasingly interested in outcomes that used to be primarily concerns of academic research.

These accounts suggest that virtual reality may provide an ideal setting for the measurement of immersive experiences. Because these experiences have been found to be more intense in virtual reality compared to traditional displays, they may be more memorable to the user in a similar manner to other highly arousing emotions (e.g., Barnacle et al., 2016; Bradley et al., 1992), with dedicated neural networks for affective memory (e.g., Dolcos et al., 2004). Consequently, in a retrospective setting, wherein user experience questionnaires are administered after the game, these affectively loaded experiences may offer an advantage in terms of memorability and by extension reliability of the self-reported experience. Those approaches, toward the measurement of flow, are described in the next section.

2.4.5 Game genres

Video games are often clustered into discrete categories that define their characteristics and the game mechanics that typify a particular style of gameplay. These categories, otherwise referred to as genres, are reminiscent of other media forms, such as films and literature (Wolf, 2001). Though the classification of games has been occasionally abused in the past, by distinguishing games based on aesthetics in lieu of the ergodic interactive elements (Apperley, 2006), game genres provide loose boundaries between different play styles. In the recent decades, there appears to be a strong presence of hybrid genres (Dale & Green, 2017; Doherty et al., 2018) and the implications of what gameplay those genres intend to provide are not intuitively interpretable. Calleja (2011) specifically referred to hybrid genres as games within games. The absence of a consensus on genre characteristics has also been found to greatly vary among well-known online video game databases (e.g., Faisal & Peltoniemi, 2015).

2.4.5.1 The problem with categorizing games

Although the absence of a consensus on the characteristics that define a genre is not inherently problematic, it reveals that a common vocabulary is lacking, which can impede our understanding on research findings (Doherty et al., 2018). For example, the topic of training individual skills using video games garnered ample attention from a study by Green and Bavelier in 2003 (Dale & Green, 2017). The study made somewhat bold claims, for its time, stating that action video games have the capacity to enhance visual attention (see, Green & Bavelier, 2003). However, Latham and colleagues (2013) pointed out that the term "action", which was used to describe fast-paced gameplay, typically in the form of a first-person shooter (Boot et al., 2011), is problematic. The authors argued that action gameplay comprises multiple genres, including role-playing games, racing and sport (Latham et al., 2013), all of which are conventionally regarded as distinct genres (Faisal & Peltoniemi, 2015). In a similar vein, the ongoing debate of post-game hostility has been found to be probably unique to certain genres (Dickmeis & Roe, 2019).

According to Faisal and Peltoniemi (2015), there appears to be a consensus over the distinction of the following genres: action, shooting, role-playing game, strategy, simulation, sports, racing and fighting. However, as noted earlier, the action genre is a broad umbrella term and may include some of the genres already listed by the authors. Interestingly, Apperley (2006) stated that the action genre comprises first-person shooters and third-person games. However, these "sub-genres", as he claims them to be, may denote virtual camera perspectives (see Denisova & Cairns, 2015b; Taylor, 2002), that blend with the narrative (Bryce & Rutter, 2002), and are not indicative of their gameplay content.

Although the author goes on to mention the typical characteristics of games employing these types of perspectives, he seems to be grouping them retrospectively, based on existing design practices, rather than a concrete set of rules that distinguish action games from other genres. In addition, the distinctive element of action games is generally accepted to incorporate fast-paced gameplay, which requires quick reflexes from the player’s side (Spence & Feng, 2010). This signifies the presence of quantifiable properties at play. However, just how fast should the game's pace be in order to qualify as an action game? These accounts suggest that the defining elements of each genre may be equivocal to different authors.

2.4.5.2 Distinguishing genres based on cognitive expenditure

In an attempt to bridge the apparent gap in the literature, Doherty and colleagues (2018) suggested that genres could be distinguished by the cognitive skills that they marshal. This is a sensible approach and is corroborated by Boot and colleagues (2011) who contended that expert action video game players may not necessarily benefit from the act of playing action games, but instead, they are originally drawn to these games because of their individual capacity to meet the requirements of those games. However, classifying games on the basis of cognitive demands and/or the cognitive benefits they endow (Doherty et al., 2018) poses two issues. First, the genre might become all-inclusive and insensitive to thematic cohesion among blatantly different play styles (e.g., Clearwater, 2011). Thus, this classification scheme does not view games as entertainment tools entirely and clusters virtually irrelevant games together.

For example, spatial orientation is a cognitive ability that is utilized in survival horror games as well as massively multiplayer online role-playing games (e.g., Hsu & Chen, 2009; Tajerian, 2012). Yet, exemplars from each of these genres testify that their differences may not adequately justify their conceptual union (for instance, compare Resident Evil 7: Biohazard by Capcom to World of Warcraft by Blizzard Entertainment that exemplify this case). The issue here is amplified because we apply a traditional understanding of what a game genre entails, as it happens in the case of films (e.g., Wolf, 2001). A second problem arising from Doherty et al.'s recommendations is that classifying games based on their cognitive benefits is questionable, due to the opposing evidence from studies reporting null effects from video game training (e.g., Roque & Boot, 2018; Sala et al., 2018; Van Ravenzwaaij et al., 2014).

2.4.5.3 Individual traits are still relevant

To diverge from this theoretical debate, that remains unsettled, we can follow Faisal and Peltoniemi (2015)'s view that a number of genres are commonly purported to be distinct: action, shooting, role-playing games, strategy, simulation, sports, racing and fighting.

However, shooting, sports and racing games could in principle be collapsed to action games (Latham et al., 2013). From the aforementioned accounts, a strong takeaway emerges from Boot et al. (2011)'s observation: Video game players may exhibit a propensity toward playing specific types of games, which may be in harmony with their own skills (also, Ryan et al., 2006). Essentially, this observation alludes to the competence motive (see Ryan & Deci, 2000b) and the skills-challenge compatibility of the flow experience. However, research has demonstrated that other variables might also factor in game genre preferences, such as *personality traits* (e.g., Peever et al., 2012), *expertise* (e.g., Scharkow et al., 2015), *gender* (e.g., Greenberg et al., 2010; Rehbein et al., 2016; Vermeulen & Van Looy, 2016) and *age* (Bilgihan et al., 2013; Greenberg et al., 2010). In addition, taxonomy systems that identify player archetypes based on their motivations for play (e.g., Bartle, 2006; Nacke et al., 2014) indicate that players who fall under specific archetypes may not enjoy games that fail to satisfy their intrinsic needs. For example, in order to accommodate the "killer" player dynamics (Bartle, 1996), the game should support player-versus-player (most commonly abbreviated as PvP) combat.

The effects of these variables on game genre preferences suggest that they can also strengthen the intrinsic motivation toward specific types of games (Joyner & TerKeurst, 2006). By extension, increased intrinsic motivation may facilitate the onset of the flow experience (Csikszentmihalyi, 1990). These notions may seemingly discount the importance of game design, in that individual preferences are beyond a game designer's control. However, as pointed out earlier for flow, the emerging experience during video game play results from the combination of user characteristics and the game's characteristics (e.g., Nah et al., 2014). Notably, certain game genres may bestow slightly different experiences than others (Johnson et al., 2015), but there is little evidence to suggest that specific games are better at triggering flow (e.g., Michailidis, Balaguer-Ballester, & He, 2018). In the next section, we provide a brief critical overview of the existing methodology in flow research, illustrating some of the challenges associated with its measurement.

2.4.5.4 The challenges in measuring flow in video games

Commercial games are commonly closed source software, which makes content modifications implausible. Though commercial games are usually exemplars of established game designing practices, their conformity to experimental control or the ability to store game events into log files are lacking (Järvelä et al., 2015). This makes it challenging for an experimenter to identify the effects of certain game mechanics on the user's experience. Notably, user experience research tackles with this issue via think-aloud protocols or interviews (e.g., Tan et al., 2014). However, these approaches typically introduce artificial pauses to the play sessions and do not always allow the player to engage naturally with the game (Tan & Pisan, 2012). In addition, the variety of game design features there exist in commercial games can produce a variety of experiential outcomes, which poses another challenge in comparing data among participants. As such, a customized game provides unconstrained development that can be tailored to the goals of the study (Järvelä et al., 2015).

Game genres used in research. The landscape of research in flow is far from standardized (e.g., Jackson & Marsh, 1996). In the video games domain, the vast diversity of commercial games available has sparked a similar diversity of games employed in the study of immersive experiences. For example, Jennett and colleagues (2008) used Half-Life, an FPS shooter game to investigate immersion. Likewise, Nacke and Lindley (2008) used modified maps of Half-Life 2 to induce flow and immersion. On the other hand, Cox and colleagues (2012) used a Tower Defense game from the family of strategy games, manipulating time pressure to increase perceived challenge and, by extension, immersion. Klasen and colleagues (2012) used Tactical Ops: Assault on Terror, which is another game from the FPS shooters genre, to investigate brain activity changes during flow. Yun and colleagues (2017) also employed an FPS shooter in conjunction with electroencephalography to identify the neural dynamics of flow. Finally, Chanel and colleagues (2011) used Tetris, a puzzle game, to distinguish emotions elicited by difficulty (boredom, engagement and anxiety) via electroencephalography.

These studies exemplify the variety of games used in flow research. A clearly disproportionate trend can be observed toward first-person shooters in this research domain. This may be because FPS games are amongst the most popular genres in the gamer population (Cardamone et al., 2011; Clyde & Thomas, 2008). However, as pointed out earlier, first-person shooters rely on player-controlled avatar movement, which would be challenging to implement in virtual reality without running the risk of motion sickness. While one can suppress movement or even disable it entirely in an FPS game, a viable concern from such a modification is that the final game may not be representative of the average commercial FPS model. As such, research concerns, such as external validity and novelty effect, may pose considerable limitations.

<u>Research approaches.</u> A common practice in video games research is to alter the game's challenge as part of the experimental manipulation (e.g., Denisova & Cairns, 2015). This is also a common strategy in the study of flow specifically, where a singular antecedent is isolated to study its effects on the overall flow experience. Indeed, balance of skills-challenge has been the most common method for achieving experimentally induced flow episodes (e.g., Keller & Bless, 2008; Klarkowski et al., 2015; Peifer et al., 2014; Ulrich et al., 2014). This antecedent is driven by theoretical support, which suggests that achieving balance between player skills and task demands is the chief prerequisite of the flow experience (Csikszentmihalyi, 1990; Massimini & Carli, 1986).

The flow experience has been measured quantitatively with self-reports (e.g., Nacke & Lindley, 2010), the experience sampling method (e.g., Csikszentmihalyi & Hunter, 2003; Fullagar & Kelloway, 2009), which supports both quantitative and qualitative data (Salkind, 2010), interviews (e.g., Jackson, 1996) as well as physiological measures (for a review, see Nah et al., 2017). Of these, flow has been predominantly measured through self-report scales (Lee et al., 2014), but, more recently, there have been studies investigating the physiology behind flow, which do not necessarily exclude the administration of self-reports. In the video games literature, self-reports have been adapted to three paradigms: (1) *stop-and-report* (e.g., Drachen et al., 2010; Jennett et al., 2008; Kirschner & Williams, 2014), (2) *intermittent* (e.g., Nacke & Lindley, 2010), and (3) *retrospective* (e.g., Borderie & Michinov, 2016; Rau et al., 2017).

In the *stop-and-report* approach, participants are sporadically interrupted from playing, in order to report on their most recent experience. The benefit of this approach is that the reported experience is of considerable relevance to the context of the game at the given time. Consequently, one can record the most salient responses originating from a specific set of in-game events. However, this approach may force the players out of flow by directing attention to the self, which results in a conflict with the loss of self-awareness constituent of flow (Bian et al., 2016). The *intermittent* paradigm may rely on the design of the game used in the study, in that it should comprise artificial pauses (e.g., a loading screen or advancement to the next level). Once the player completes one level, the examiner(s) asks them to report their experience. Then, they continue to the next level, and the cycle is repeated. This approach is advantageous in that it offers insight on the interplay between the particular level design and the player's experience.

Finally, in the *retrospective* paradigm, participants are asked to report their experience after the play session has finished entirely. Such an approach might be more efficient in giving a holistic impression of the player's experience during the game. Additionally, it ascertains that the participants engage naturally with the game, without the interference of research settings, which can be invasive to the player's experience (Nuñez Castellar et al., 2016; Yannakakis et al., 2013). However, it may also obscure other parameters of interest, such as the length and intensity of the flow experience during the game (Rißler et al., 2018).

Concerns on existing methods. The methodological concern that is often raised in the literature is the reliance on self-reports (Rißler et al., 2018). Although self-reports can be of complementary value to the study of flow (e.g., Peifer et al., 2014), they may be unreliable (Shearer, 2016; Yun et al., 2017). Recent accounts have also identified a strong overlap among different questionnaires, suggesting a potential redundancy (Denisova et al., 2016; Nordin et al., 2014). The drawbacks of self-report tools can be compensated with neurophysiological measures, as they offer an objective way to identify the cognitive functions associated with flow (Bian et al., 2016; Engeser, 2012).

Following Csikszentmihalyi & Nakamura's (2014) proposition that nearly any activity can trigger flow, the flow experience has been investigated in multiple contexts. Keller and colleagues (2011b) compared flow experiences between two activities, video game playing and a knowledge task. The authors posited that flow shall emerge regardless of the activity, as long as there is compatibility between individual skills and task challenge. This finding is corroboratory of Csikszentmihalyi's (1997a) suggestion that skills-challenge balance is a sufficient condition for the flow experience. It is clear from Keller et al.'s study that the authors understand that a knowledge task may not be as intrinsically motivating as video game playing: "participants’ attitude toward the knowledge question activity may differ from attitudes toward playful computer games" (p. 409). In a follow-up review by one of the authors (Landhäußer & Keller, 2012), they raised the concern that skills-challenge balance is often equated to the flow experience in the literature.

This is an important observation that calls several studies into question. For example, mental arithmetic tasks have also been used in the study of flow where the only controlled variable was the difficulty of the task (e.g., Ulrich et al., 2014, 2016). However, there is little evidence supporting the idea that mental arithmetic tasks are intrinsically motivating, i.e., people choosing routinely to engage in mental arithmetic tasks (e.g., Ryan & Deci, 2000a). Yet, even then, free choice, as a determinant of intrinsic motivation, does not necessarily lead to engagement with an activity (Katz & Assor, 2007). Keller and colleagues (2011b) surmised that skills-demands compatibility leads to intrinsic motivation and treated intrinsic motivation as a consequence to flow. According to the authors, willingness to re-engage with the activity lies at the core of intrinsic motivation (also, Patall et al., 2008). Similarly, enjoyment is also understood as a consequence to the flow experience (e.g., Csikszentmihalyi, 1990, 1996; Rodríguez-Sánchez et al., 2011).

In light of this position, enjoyment becomes separable from the activity, which is, by definition, an extrinsic incentive (see Ryan & Deci, 2000a). Yet, in the study by Rodríguez-Sánchez and colleagues (2008), intrinsic motivation and enjoyment were found to have a bidirectional relationship with each other, i.e., intrinsic motivation fosters enjoyment, but also enjoyment fosters intrinsic motivation. Likewise, Reeve (1989) showed that enjoyment fosters intrinsic motivation. Thus, it appears that the activity may be pursued in future instances for the enjoyment derivative of the flow experience. If both extrinsic and intrinsic motivation have ties to the dopaminergic system (e.g., Baldassarre, 2011), then it is unclear what qualifies enjoyment as an intrinsic incentive instead of an extrinsic. These accounts give rise to an important question: Does *any* cognitively challenging activity automatically qualify for a flow experience? Perhaps, findings from neural studies of flow employing different tasks may require more evidence to be made comparable (cf. Harmat et al., 2015; Klasen et al., 2012; Ulrich et al., 2014, 2016).

In this section, we adumbrated few of the methods toward the measurement of flow in the video games literature. Among those, the principal method involves the use of self-reports, aiming to capture the subjective essence of flow. However, the extant criticism suggests that self-reports may be somewhat insufficient to interpret the richness of the flow experience. We also demonstrated three prevalent paradigms that define the order in which questionnaires may be administered. While each approach has different advantages, none of them appears to ward off the risk of compromising the player's experience. Nevertheless, the retrospective administration of self-reports does seem to minimize the invasiveness of the experimental protocol, as it does not interrupt the players midway through the game.

2.4.6 Conclusion

Video game design encompasses a set of rules that define the ways the player interaction can be realized. This interaction alone is considered a legitimate source of satisfaction (e.g., Grodal, 2000). However, it should be noted that vicarious play, i.e., watching someone else play, can be also satisfying (Cheung & Huang, 2011; Sjöblom & Hamari, 2017; Wulf et al., 2018). Thus, there are more elements at play other than player interaction. For example, fantasy (Ferguson & Olson, 2013; Kim & Ross, 2006; Malone, 1981), curiosity (Jin, 2012; Malone, 1981) and the need to escape from reality, known as escapism (Calleja, 2010), constitute only a handful of the motives underlying video game playing.

Nevertheless, flow's onset is facilitated by the way games are designed. Its dimensions, clarity of goals, balance between player skills and game challenge and immediate and unambiguous feedback, have been regarded as antecedents to the flow experience (e.g., Chen, 2006; Quinn, 2005). These preconditions can be adjusted on a technical level: The game's difficulty, the design of a user interface to effectively communicate game progress to the player, as well as the definition of specific goals for the player to fulfill, can be established at the design and development stage.

However, promoting other flow dimensions through game design, such as automaticity or loss of self-awareness, is more challenging and may involve complex techniques, such as suspension of disbelief (Whitson et al., 2008) and an intricate narrative (Sweetser & Wyeth, 2005). Instead, substitutional design strategies may be used to facilitate the emergence of flow, such as *background music* (Sanders & Cairns, 2010; Zhang & Fu, 2015), believable characters to stimulate *character identification* (Hefner et al., 2007; Soutter & Hitchens, 2016), *time pressure* (Cox et al., 2012; Nah et al., 2014; Yildirim, 2016), and the opportunity for *social interaction* (Cowley et al., 2008; Sweetser & Wyeth, 2005).

To conclude, the building blocks of a game derive fundamental support from psychological studies. However, the ultimate experience is highly subjective and cannot always be tamed by conventional approaches. From this review, it becomes increasingly clear that individual characteristics play an important role in the onset of flow. The games used in different studies, as well as commercial games, try to mold the player to the game instead of the opposite: mold the game to the player. By aiming for personalized experiences, one can essentially create an idealized testbed to investigate flow dynamics. This context may be somewhat challenging to generalize, as it departs from the conventional research methodology that aims to minimize contingencies across participants. However, this is not limited to the game content, but also the platform on which the game operates. Novel technologies such as virtual reality are tempting to investigate, not merely because of their relatively recent commercialization, but also because they promise richer user experiences (e.g., Baştuğ et al., 2017; Tan et al., 2015).

3 General Methodology

3.1 Introduction

In this chapter, the general methodological aspects, common to all experiments, will be described and reviewed in the broader context of flow. Specific methodological details will be discussed for each experiment in Chapters 4 to 6. The building blocks of our methodological approach encompassed the development of a virtual reality game, the collection of participant data and the classification thereof.

The initial section is dedicated to the description of the game used in the study and the strategies employed to associate physiology with the flow experience during video game play. First, the reasoning behind the selection of a Tower Defense game is detailed, followed by the materials and the data preprocessing steps to reach to the solution of the classification problem. An overview of common classification algorithms is also presented, all of which have been used to predict flow episodes through electrocardiography, electrooculography and electroencephalography.

3.2 The design of the experimental game

3.2.1 Design considerations

Customized video games are not uncommon in video games research, as the experimenter can tailor the game's characteristics to the experimental goals (see **The challenges in measuring flow in video games**). We similarly proceeded with a custom virtual reality (VR) game for the purposes of the present work. As argued earlier, VR was chosen for two main reasons, its intrinsic immersive value and novelty. The idea that a technology can be intrinsically more immersive than others has been propounded by Slater (e.g., Slater & Sanchez-Vives, 2016). Because of commercial VR's head-mounted display, the user is immediately impelled to engage with the virtual environment, given that the real, surrounding world is visually inaccessible. This requires that the sensory stimulation achieved by VR be an effective substitute for the real world (Slater & Sanchez-Vives, 2016).

In the case of VR games, studies have found that players report a heightened sense of flow whilst in VR than when using desktop displays (e.g., Tan et al., 2015). Similarly, environmental, non-interactable objects are paid closer attention to when in VR than when using a 2D display (Tan et al., 2015). As per the emotional experiences, they have been shown to be more arousing than their desktop-induced equivalents (Pallavicini et al., 2018, 2019). Fortunately, the players' perception of difficulty remains similar despite the type of display used (Pallavicini et al., 2019). This is important in the study of flow, considering that game difficulty plays a pivotal role in the experimental design.

In the design of our game, we adopted the general view that skills-challenge balance, presence of feedback and clarity of goals constitute flow's antecedents as earlier research proposed (e.g., Chen, 2006; Fullagar et al., 2013; Quinn, 2005; Rodríguez-Sánchez et al., 2008). Each of these antecedents generates controllable variables, which would be translated into game difficulty, (un)ambiguous or (in)existence of user interface elements for feedback and (un)clear or absent goals. To preserve consistency with earlier studies, we focused on the skills-challenge antecedent as traditionally approached in most flow studies (e.g., Keller & Bless, 2008; Klarkowski et al., 2015; Peifer et al., 2014; Ulrich et al., 2014).

Selection criteria. First, we aimed to narrow the scope of genres that could be used as potential candidates for the design and development of the experimental game in virtual reality. These criteria were chosen such that the game in question would demonstrate qualities that are compatible with flow, but also for improving the experimental setup without compromising the quality of user experience. There were three major considerations for the selection and design of gameplay content, i.e., *usability*, *liability to individual differences*, and *virtual reality comfort*. **Usability** describes the ease of interaction with the game and the learning curve that goes into the familiarization with the game's controls and the gameplay (Pinelle et al., 2008). The advantage of balancing the learning curve is the minimization of the experimental times.

Based on prevalent usability practices, usability in this work was accounted for by the following dimensions: (a) a short learning curve characterized by gameplay simplicity (Desurvire et al., 2004), (b) a finite set of functionally distinct options the player can choose from that precludes choice overload (Besedeš et al., 2015; Carson, 2000; Vriend, 2017), (c) a finite set of deterministic actions the player can perform at a time with consistent outcomes (Al-Azawi et al., 2013; Yuan et al., 2011), (d) clear and immediate feedback (Desurvire et al., 2004), and (e) naturalness of controls, where applicable, to accommodate virtual reality idiosyncrasies (Skalski et al., 2011; Sutcliffe & Kaur, 2000).

The purpose of these criteria is to facilitate the onset of flow. We expected that a game with a short learning curve would potentiate faster access to flow, preventing the player from idling as a result of goal obscurity or unintuitive game controls. On the other hand, limiting the number of available options at a time is coveted for reducing the game's complexity and the variability of experiential outcomes. Finally, the usability criterion of feedback's clarity and immediacy ties directly to the antecedent of the flow experience. Once the game's state changes, it should be communicated directly to the player through visual or auditory elements (e.g., Jørgensen, 2008).

Liability to individual differences refers to the attractiveness of a given game to the average player. For example, the game should be accessible to players regardless of their previous game experience or genre preferences (Klarkowski et al., 2015; Shelley, 2001). In practice, neutralizing the game's content in favor of a broader audience poses a considerable challenge to its design. Perhaps this challenge justifies the recent emergence of hybrid genres in the games industry (e.g., Dale & Green, 2017; Doherty et al., 2018). The core elements of digital games include narrative, mechanics, aesthetics, interface, etc. (e.g., Ralph & Monu, 2015; Takatalo et al., 2010). From an experimental standpoint, each of these elements can be treated as an experimental variable, which adds to the complexity of interpretation in games research. These elements work in a complementary way to one another (Ang, 2006; Ralph & Monu, 2015), thereby making video games an ergodic form of media (Calleja, 2013).

Notably, storytelling is not to be dismissed as a mere piece of contextual information, as it can also define new goals for the players depending on game progression (Calleja, 2013; Qin et al., 2009). This dynamic relationship between narrative and gameplay (Ang, 2006) can delay the player's sense of progression, because the goals can be generated or presented on the fly rather than being communicated to the player a priori. As such, the experimental time may increase significantly depending on the complexity of the storyline. In addition, different narratives or modes of narrative presentation may appease to different players, whereby they widen the gap of inter-player preferences.

Video games that are destitute of narrative can still be engaging (Ang, 2006), without sacrificing the advantage of experimental control. While such an omission could potentially impinge on the game's ludic elements, research usually favors experimental control while acknowledging potential trade-offs (McMahan et al., 2011). In this project, and for the reasons highlighted above, narrative-driven games were dismissed. For example, role-playing games that are rich in storytelling (Grace, 2005; Roth et al., 2010) were not considered.

Similarly, games that tax specific resources as determinants of the players' in-game success may also be undesirable, if we consider the notion that players may prefer games that are attuned to their individual prowess (e.g., Boot et al., 2015; Granic et al., 2014; Steinkuehler & Duncan, 2008). For example, game progression in first-person shooters relies heavily on individual advantage of aiming accuracy, which typically requires time to develop (Vicencio-Moreira et al., 2014). This may be an apparent limitation for players who fail at aiming correctly, causing delayed game enjoyment (Klimmt et al., 2009).

However, in the specific case of first-person shooters, external factors such as frame rate, can also pose detrimental effects on player performance and enjoyment (Claypool et al., 2006), because these games depend on quick reflexes (Buckley et al., 2013; Sweetser et al., 2017). Thus, in the present research, the experimental game in question should collapse into a simple set of interactions, that transcend innate player advantages. For the aforementioned reasons, racing, shooting and sports games were not considered for the purposes of the study.

Finally, **virtual reality comfort** pertains to the potential for motion sickness, which may be detrimental to the player experience (e.g., Seibert, 2014; Tan et al., 2015). A prominent theory behind the etiology of motion sickness is sensory conflict, which suggests that motion sickness may trigger when motion perception (i.e., vection) is not linked to physical self-motion (Hettinger & Riccio, 1992). Symptoms include ataxia, dizziness, nausea, and vomiting (Akiduki et al., 2003; Cohen et al., 2019; Oman, 1990; Rogers et al., 2018; Sharples et al., 2008). In addition, the severity of motion sickness may increase as a function of the depicted navigation speed (So et al., 2001). Games which are reliant on player mobility include sports, action, role-playing and adventure, due to their intrinsic relation to exploration or physical activity depicted through a player-controlled avatar. As such, these genres were excluded from potential candidacy.

3.2.2 Evaluation of selection criteria

An interesting candidate that obliges with some of the aforementioned criteria are strategy games, which offer content that is compatible with the flow experience (Sweetser et al., 2012). Strategy games may be preferred by "hardcore" gamers (Gackenbach & Bown, 2011) perhaps because of the complexity in their design (Sweetser & Wyeth, 2005). This may be a limitation with respect to our consideration of usability, in that strategy games may require time to learn to play (Sweetser & Wyeth, 2005). However, a subcategory of strategy games used in earlier studies are Tower Defense (TD) games (e.g., Chesham et al., 2017; Cox et al., 2012), also rarely classified as puzzle games (e.g., Wetzel et al., 2012). These games are typically viewed from above (bird's eye view) without shifts to the local position of the in-game camera. Hence, they are easy to recreate in a VR environment, using the player's head tilt as the governing view input. Essentially, this mode of interaction maintains a camera with a static local position, thus precluding physical movement, and exempts from the occurrence of motion sickness. These aspects endeavor *virtual reality comfort*.

In terms of their usability, they are supported by a single goal that is easy to memorize and comprehend. Likewise, their mechanics are easy to learn (Brich et al., 2015). Despite some variability in existing TD games on the market, the gameplay of a TD game can be summarized as follows: A wave of enemies is spawned at fixed intervals and start from point A to an exit point B. The enemies walk along a predefined path and, if any of them escapes, a life point is removed from the player's life pool. Once the player's life pool reduces to zero, the player is defeated. Thus, the goal of the player is to eliminate as many enemies as possible before they escape through the exit point. Enemies can be eliminated via special constructions, which are objects placed by the player across the game level, usually depicting towers (hence the name Tower Defense). An example of such a game is Kingdom Rush (Ironhide Game Studio & Armor Games) (see **Figure 1**), a highly successful Tower Defense game with an "Overwhelmingly Positive" rating on Steam [1] and 9/10 on IGN [2].

Figure 1: Sample screenshot from the successful Tower Defense game Kingdom Rush. The visible humanoids on the light gray path are the enemies of the player who walk along the path to reach the flags with the shield symbols on top. The structures are scattered across the game level whose purpose is to eliminate or slow down the enemies from reaching the exit point. *Image credits*: Ironhide Game Studio and Armor Games.

[1] https://store.steampowered.com/app/246420/Kingdom_Rush/
[2] https://uk.ign.com/articles/2012/01/31/kingdom-rush-review

Typically, there are multiple types of towers the player can choose from, which add layers of complexity and strategy to the game. These layers prevent linear gameplay and allow for multiple ways toward winning or losing the game. However, every subsequent enemy wave is more powerful than its preceding, therefore the player must balance their damage output. This is achieved by upgrading existing towers or by creating pockets of towers across the map's areas that work optimally together. Enemies are also of different types, with diverse characteristics ranging from movement type (e.g., ground or aerial) to a unique arsenal of abilities. Those abilities may make the enemies vulnerable to certain towers, but more resistant to others. Finally, bosses are also typical in TD games. Bosses are very powerful enemies that take much longer to eliminate than a common enemy. Because of this discrepancy, they reward generously if eliminated.

Importantly, TD games have been deemed a viable testbed for several research applications, including player experience (Avery et al., 2011). In addition, they are usually organized in rounds, with finite length, and clear boundaries demarcated by the beginning and end of each round. Therefore, they conveniently emulate experimental block design as in the case of traditional experiments. A problem that has been pointed out in the literature is that games are dynamic and complex, making it virtually impossible to preserve control over the stimuli presented on the screen (Yan & El-Nasr, 2006). Hence, researchers may resort to clustering techniques (Yannakakis et al., 2013) or coding game events via annotation (e.g., Klasen et al., 2012) to narrow the scope of their search for behavioral outcomes of interest. Tower Defense games compensate for this limitation through their structured, round-based gameplay that is exceptionally common to this genre.

The simplicity of TD mechanics makes them additionally accessible to novice players, supporting the *liability to individual differences* criterion. In a study by Chesham and colleagues (2017), TD games were rated among the highest casual video games in terms of "ownership". Ownership is a construct derived from a questionnaire developed by Calvillo-Gámez et al. (2015) that denotes player self-attribution of responsibility for the events in the game. In other words, ownership relates to the way the player takes advantage of the action affordance to reach a goal as well as the utilization of in-game rewards (Chesham et al., 2017).

As explained earlier, in-game failure that is attributed to the self may not necessarily lead to loss of player motivation (Child, 1994; Juul, 2013; Nylund & Landfors, 2015) and flow. Naturally, irrespective of the game, the affective and behavioral consequences of failure attribution can vary considerably across individuals (e.g., Singh, 2013). Yet, the incremental difficulty that characterizes TD games allows the player to gradually develop tolerance to frustration. Nylund and Landfors (2015) discussed that building tolerance can hinder players from experiencing the negative effects of frustration, whereby they preserve their motivation to overcome a challenge. This is an important advantage, in that frustration may precipitate postural disturbance (e.g., Perron, 2005), resulting in signal artifacts, or even quitting the game (e.g., Wauck et al., 2017), both being undesired outcomes.

What makes the TD genre even more interesting as a testbed is that it presents a particularity with respect to difficulty management. As opposed to existing methodological approaches of flow induction, wherein the condition of boredom is traditionally approached as a low-difficulty condition, boredom in TD is somewhat trickier to replicate using the same principle. Because of the linear difficulty increase that TD games typically implement at each round, abrupt suspension of the natural difficulty scaling may be surprising to the player and expose the experimental manipulation. Ergo, a different approach must be considered that is perhaps specific to this genre.

Overall, the TD genre appeared to be a promising testbed for the investigation of flow in virtual reality. Notably, this genre is not as popular as other genres, such as first-person shooters. Nevertheless, TD games satisfy flow's antecedents, i.e., presence of feedback with diegetic user interfaces, clarity of goals with their simple gameplay, and skills-challenge balance with modifiable attributes of the game actors and structures. In addition, they facilitate data annotation based on modifications made across the gameplay experience conditions (i.e., flow, anxiety and boredom).

3.2.3 Tower Defense VR for PlayStation VR

The game developed in this study, Tower Defense VR (TD-VR), was, to our knowledge, one of the first attempts to ever port a TD game in a VR platform. TD-VR game was developed in the Unity 3D (Unity Technologies) game engine for the PlayStation VR platform. The game mechanics that were adopted for TD-VR were chosen such that the resulting game did not deviate considerably from the average commercial Tower Defense game. The player was virtually placed in front of a game table which was allotted for the game action. The remaining virtual space was a room emulating a bedroom. Thus, the virtual scene was composed of two main elements, the game table and the surrounding virtual space. The surrounding space was unrelated to the game progress but was created as a complementary means of contextualizing the activity within the 3D space.

The virtual camera was responsive to the player's head rotation, but not their physical position. Thus, the players were restricted from moving into the virtual room. As a result, the overall physical motion was minimized since the game action took place in a static space in front of the player. The advantage of this approach is the minimization of motion sickness (Kim et al., 2020; Whittinghill et al., 2015). The camera's field of view was adjusted to encompass most of the game table's contents in its viewing frustum when gazing at the center of the table. Thanks to this approach, the player could naturally maintain a constant posture throughout the game, thereby reducing the presence of motion artifacts in the physiological data. These steps were taken to ensure that the criterion of virtual reality comfort was met.

The features of TD-VR were made to match the average TD game. However, to keep the experimental variables to a minimum, only a few options were given to the player. The game had only one type of common enemies, with their appearance emulating a medieval footman, and boss enemies, who were significantly harder to defeat. The experimentally manipulated variables, as a function of difficulty scaling over time, were the <u>health</u> of the enemies and the <u>reward</u> that was bestowed to the player after an enemy was eliminated. The reward magnitude was directly associated with difficulty, as it was used for building and upgrading towers, which would in turn increase the overall damage output of the player (see also, Cox et al., 2012). These two variables are commonly used, either exclusively or partly, to differentiate levels of difficulty in commercial TD games. Three tower types were available, each having three upgrade levels. All towers were given the same virtual price for building and/or upgrading, irrespective of their unique abilities. This was done to prevent players from associating high building/upgrading costs with more power, which could have led to biased preference toward specific tower types. The power of the towers was regulated through attack frequency (also known as attack speed), attack power (the number of health points removed from the enemy on attack – also known as damage dealt) and attack range (distance needed for the tower to pick up its target).

The virtual currency for tower transactions was given to the player in two ways, (a) by eliminating an enemy and (b) every 8 seconds of game time. The monetary reward increased proportionally to the game stage, i.e., the more powerful the enemies, the higher the gold reward. The gold reward that was bestowed intermittently was equivalent to the amount granted from eliminating a common enemy from the most recent game round. This amount was not given during the inter-round breaks - only during the actual game rounds. The aim for these rewards was to compensate the player for their effort with the additional intermittent bonus. In that sense, we employed a fixed-reward schedule, as commonly practiced in TD games. The aforementioned details were common to all subsequent studies. Specific implementation details are mentioned in the *Materials: Test Game* section of each corresponding study.

In short, the TD-VR game aims to produce varying levels of skills-challenge balance along the course of the game. This is done by manipulating two variables that drive the action: the health of the enemies and the monetary compensation for eliminating an enemy. To optimally capture the connection of personality to flow as per the research goals, it was deemed that the game should be denuded of highly idiosyncratic content that would appease to a niche group of players. The resulting flow experience is still naturally linked to the game and its particularities, but we attempted at abating the potency of the game's content in favoring specific personality types. For example, the presence of aggressive/violent content may lead to inflated enjoyment scores in individuals high in neuroticism or those low in agreeableness (Chory & Goodboy, 2011).

Next, TD games have a convenient design principle with artificial pauses between rounds and/or game levels. This makes it easier to conceptually separate the three design considerations of flow, anxiety and boredom, allowing for distinct content. The onset of each game round offered a natural way to demarcate the physiological recording epochs that were considered most valuable, thereby facilitating the annotation process. The principal advantage of the TD-VR regarding the biometrics acquisition was the flexibility in using a statically placed camera without influencing the gameplay content. This has a two-fold benefit, i.e., the virtual elimination of motion sickness and a relatively stable body posture. In shooter games, for instance, the player may physically steer toward the direction they are moving in the virtual world (Bianchi-Berthouze et al., 2006), and thus risk signal contamination.

3.3 Questionnaires

3.3.1 Overview

The questionnaires used in the study were aimed at capturing the most salient features of the flow experience and individual personality. Naturally, because the flow experience has mainly been studied through questionnaires (Lee et al., 2014), there are several assessment scales available. Examples of general flow questionnaires (not bound to a specific activity or task) include the Flow State Scale – 2 (Jackson & Eklund, 2002) and the Flow State Questionnaire of the Positive Psychology Lab (PPL-FSQ; Magyaródi et al., 2013). Flow assessment in video game play includes the Game Engagement Questionnaire (Brockmyer et al., 2009), a 6-item questionnaire developed by Choi and Kim (2004), EGameFlow (Fu et al., 2009) and the Game Experience Questionnaire (IJsselsteijn et al., 2008).

However, the psychometric properties of those scales incorporate notions that diverge from the original conceptualization of flow. For example, the Game Experience Questionnaire (IJsselsteijn et al., 2008) entangles the definitions of flow and imaginative immersion, by using dedicated statements related to the story of the game. As such, this questionnaire is inapplicable to the game used in this work. Omitting the unrelated items to preserve the flow construct of the questionnaire results in only five items that aim to capture the dimensions of concentration and loss of time perception. We posited that these two dimensions do not sufficiently represent the flow experience. A similar observation can be made for Choi and

Kim's (2004) short scale, where the statements translate to the flow dimensions of autotelicity, loss of self-reflection, sense of control and concentration.

Finally, the Game Engagement Questionnaire (Brockmyer et al., 2009) distributes statements across multiple constructs, i.e., flow, absorption, immersion and presence. Though automaticity has been conceptualized as a dimension of flow (e.g., Csikszentmihalyi, 1990), the authors have assigned it under both constructs of flow and presence. Flow, immersion and presence have occasionally overlapping elements and definitions in the literature (Michailidis, Balaguer-Ballester, & He, 2018). However, the classification of the statements in the Game Engagement Questionnaire seems to arbitrarily choose the most characteristic constituents of flow. Given these issues, the Flow State Scale – 2 (Jackson & Eklund, 2002) was chosen to capture each flow dimension separately, as it adheres to the original flow theory by Csikszentmihalyi (1990) without enmeshing the phraseology used in the scale items toward a specific activity.

Though PPL-FSQ (Magyaródi et al., 2013) is also activity-independent, the factor loadings exposed only two factors able to explain more than 50% of total variance, which were interpreted as *skills-challenge balance* and *absorption*. Thus, PPL-FSQ was likewise regarded as being inadequate. It should be stressed that these issues do not necessarily reflect our support for a holistic flow model, where all dimensions should be simultaneously present for qualifying the flow experience. Instead, it was deemed necessary that all dimensions be captured regardless of the interpreted necessity different studies have reported for individual dimensions. Below are descriptions of the questionnaires used in the present work.

3.3.2 Flow State Scale – 2

The Flow State Scale – 2 (FSS-2) was developed by Jackson and Eklund (2002). The questionnaire is used to assess the flow experience in an activity that has already happened. This is different from the Dispositional Flow Scale (DFS-2), which is used for assessing the individual's tendency to experience flow in a certain activity (Jackson et al., 1998). The long version of FSS-2 has two counterparts, one used to assess physical activity (FSS-2 – Physical) and the other for any kind of activity (FSS-2 – General). The questionnaire comprises 36 items of statements about the activity that are to be rated on a Likert-type scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). For all nine flow dimensions, there are four questions dedicated to each. This questionnaire is commercial, and the appropriate number of copies were licensed for administration by the official distributor Mind Garden. The questionnaire was administered immediately after the end of the game session.

The use of DFS-2 remains controversial regarding its applicability in video games research (cf. Procci et al., 2012; Hamari & Koivisto, 2014; Wang et al., 2009). Procci and colleagues (2012) used the DFS-2 questionnaire to capture the flow experience in experienced players and found that the questionnaire did not generalize well in the context of video game play. The authors identified a potential overlap between the dimensions, with their model maintaining seven out of the original nine factors. In contrast, Wang and colleagues (2009) successfully identified all nine factors that converge into a global flow index.

For FSS-2, Sweetser and Johnson (2019) identified high consistency with other flow questionnaires, namely GameFlow (Sweetser & Wyeth, 2005) and the Player Experience of Need Satisfaction (PENS; Ryan et al., 2006), suggesting a uniformity in flow's conceptualization across different player experience measures. The authors observed a notable discrepancy in the average ratings between the dimension of skills-challenge balance of FSS-2 and GameFlow's equivalent dimension of Challenge. The authors considered the skills-challenge balance dimension of FSS-2 to translate poorly to the video games context, even though the same disparity was observed between GameFlow and PENS. However, their conclusion was based on descriptive statistical summaries rather than an inferential analysis.

Overall, the FSS-2 questionnaire was chosen for its ability to capture all nine dimensions originally proposed by Csikszentmihalyi (1975, 1990) and for being agnostic to the activity and domain within which it is administered.

3.3.3 Big Five Inventory

The Big Five Inventory (BFI; John & Srivastava, 1999) is a scale used for measuring personality. It consists of 44 items which are rated from 1 (Strongly Disagree) to 5 (Strongly Agree) based on how much the respondents feel that each statement represents them. The questionnaire comprises five sub-constructs, namely Agreeableness, Conscientiousness, Extraversion, Neuroticism and Openness. These personality dimensions have been thought to be efficiently capturing the human personality (Gosling et al., 2003). According to John and Srivastava (1999), agreeableness captures qualities such as altruism and trust, conscientiousness targets goal orientation and self-discipline and extraversion captures sociability and positive emotionality. On the other hand, Neuroticism relates to anxiety and negative emotionality, while Openness pertains to open-mindedness and creativity (John & Srivastava, 1999).

The Big Five personality traits have also been adopted in video games research (e.g., Bean & Groth-Marnat, 2016; Nagle et al., 2016). For example, Nagle and colleagues (2016) found that conscientiousness positively predicted the length of play individuals invested in a game with dynamic difficulty adjustment. Similarly, agreeableness may predict the frequency and length of play for mobile games (Seok & DaCosta, 2015). In another study, the five dimensions were used to predict genre preferences (e.g., Braun et al., 2016; Johnson et al., 2012; Peever et al., 2012). These findings reflect only a fraction of the application of Big Five in games research. Thus, the BFI is a particularly interesting tool in the study of game-induced flow.

3.4 Data preprocessing pipeline

3.4.1 Measurement data

To approach the research questions of this work (see Chapter 1, **Research Goals**), several sources of information were combined: (a) the game's design in line with the research hypotheses, (b) subjective ratings on a battery of psychometric tests, and (c) objective markers of physiological origin.

Multimodal measurement of complex constructs, such as flow, has been found to be more effective of an approach than individual methods (Martey et al., 2014; Robinson et al., 2018). Physiological tracking offers a way to decode mental and affective states continuously. This is in contrast with questionnaires that capture subjective perceptions at discrete times and asynchronously to the states they are intended to measure (Holzwarth et al., 2021). Through a more holistic lens, our aim was to capture the flow experience from three different perspectives and to identify their common ingredients.

As objective markers, a plethora of biometrics have been examined in the literature. We opted for heart rate variability (HRV), eye blink variability (EBV) and electroencephalography (EEG) all of which belong to the family of electrograms (Reilly & Lee, 2010). These metrics can potentially be extracted from various locations over the head, making them convenient for implementation in a flow-aware VR technology. Heart rate variability is the most widely used metric in the objective evaluation of the flow experience (Knierim et al., 2018). It can capture information pertinent to physiological arousal and mental workload and it is therefore useful in distinguishing varying levels of flow, anxiety and boredom (for more information, see Chapter 2, **Flow's "competition" with other states**).

Eye blink variability is less popular, albeit emerging, in the study of flow (e.g., Hild et al., 2020; Min et al., 2018; Rau et al., 2017). The metrics were computed through electrooculography (EOG). Ocular measures are also commonly obtained via a web camera (e.g., Cho, 2021; Ngxande et al., 2017). An advantage of this method is its robustness against motion artifacts (Li et al., 2018). However, an external camera was impractical for a VR setting, in which the head-mounted display (HMD) covers the user's eyes.

On the other hand, integrating a camera inside the PlayStation VR HMD would require a bespoke approach, and was hence considered less reliable than EOG. Even though EBV has a considerable informational overlap with HRV in terms of physiological arousal, affective valence and mental load (e.g., Maffei & Angrilli, 2019; Barathi et al., 2020), EBV is also sensitive to distractibility, attentional focus, and visual load (Magliacano et al., 2020; Recarte et al., 2008). All the above have been linked to different facets and degrees of flow, anxiety and/or boredom.

Finally, EEG has been used in many studies involving the measurement of complex states of affect and cognitive processes (Babiker et al., 2021). These states correlate with oscillations that emerge from neuronal communication of variable frequency and functional significance (Klimesch et al., 2005). Packed with its own shortcomings, EEG is still widely adopted for human-computer applications into a new wave of applied research known as brain-computer interfaces (Lotte et al., 2007). These interfaces see increasing applications in the domain of video gaming (Kerous et al., 2018) for user state recognition and brain-controlled gaming (Nijholt, 2008). Hence, they are aligned with the purposes of the present work.

The theta and alpha rhythms, for example, have shown sensitivity to task difficulty manipulations (see review by Antonenko et al., 2010). This makes EEG desirable in the measurement of flow, when one considers the prevalence of difficulty manipulation as the most widely employed methodological approach toward flow induction (e.g., Peifer et al., 2014). In a similar vein, attention is central to the concept of flow and EEG has been used extensively to measure it (Cisler et al., 2019). It is also worth noting that EEG has versatile dimensionality based on the number of channels used during the recordings, which can potentially improve the amount of information obtained.

For the detection of mental workload, eye-related features are the third most widely used metrics after cardiac measures and EEG (Tao et al., 2019). However, in terms of their ability to recognize different levels of mental workload, Hogervorst and colleagues (2014) found that their performance ranks as follows: (1) EEG, (2) EBV and (3) HRV. More specific details about these metrics are provided in their corresponding chapters (**Chapter 5** for HRV and EBV; **Chapter 6** for EEG).

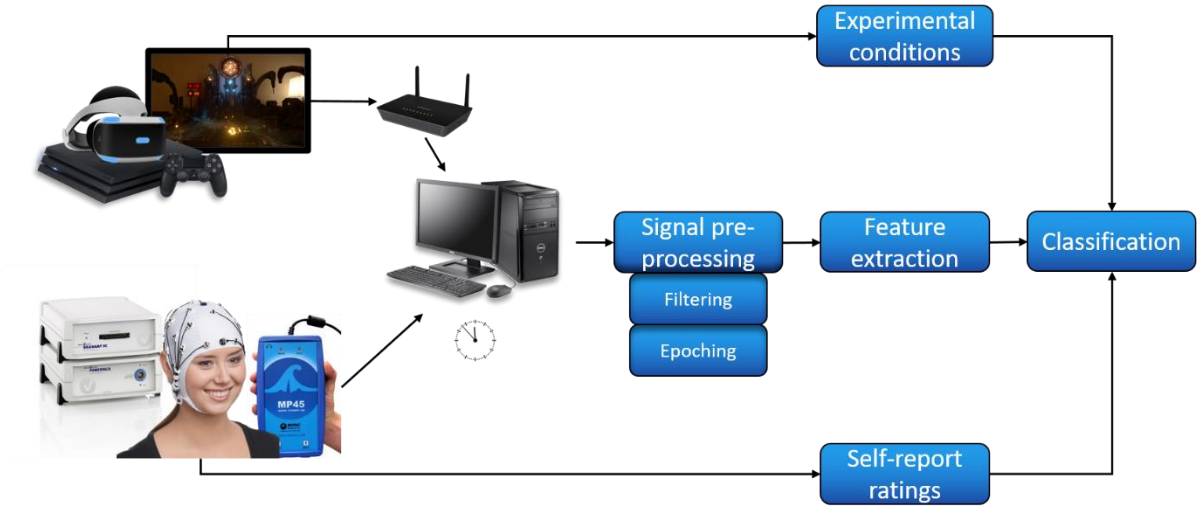


Figure 2: Semiotic diagram of the data processing pipeline used in this work.[3]

A standard approach of data collection and analysis was employed (**Figure 2**). While the game was being played on PlayStation VR, the participant's physiological responses were recorded in parallel. The game comprised three design levels aimed to induce flow, anxiety and boredom. As such, the timing of those levels was integral to understanding the effects of the game design onto the player's physiology. To synchronize the two sources, the game sent out triggers of events that signaled the onset of each game round.

The data were later submitted to a series of conventional pre-processing steps, including filtering, baseline correction, and epoching. The signal in each epoch (window) was transformed into discrete information by means of feature extraction. Then, the two definitions of subjective flow (based on the self-reports) and game-design flow were used at class-level to train and cross-validate the objective data (physiology) in order to estimate their separability. Subjective flow was tested at two levels, low- and high- flow, therefore a binary configuration. On the other hand, the game-design flow was a three-way configuration, i.e., flow, anxiety and boredom.

[3] Images of depicted devices Courtesy of BIOPAC Systems, Inc. and Brain Products GmbH.

3.4.2 Data acquisition

The devices used to capture the participant's physiological changes were connected to a desktop computer (receiver) for the acquisition of the physiological data (tags 3, 4 and 5 respectively in **Figure 3**). The internal clocks of PlayStation 4 and the desktop were asynchronous to each other. Hence, a workaround was made to pinpoint the timing of the events in the game relative to the timing of the physiological data. PlayStation 4 and the desktop receiver were connected to a router to enable non-reciprocal communication via messages (triggers).

Internally, there were four events that were emitted by the Unity software from PlayStation 4 to a specific port and IP address: *Game_Started*, *Game_Ended*, *Round_Started* and *Round_Ended*. On the desktop, a custom application developed in Visual Studio 2017 (C#/.NET) served as a listener to the same port for incoming messages. Both the desktop application and the Unity software have had access to a library (.DLL file) that held an enumeration of all event names. The messages were encoded within the Unity software into bytes and later deserialized on the receiving desktop application to be appended in a file stream.

Figure 3: Visual simulation of the experimental setup using virtual reality in Unity 3D. (**1**) PlayStation 4, (**2**) PlayStation Camera, (**3**) ECG and EOG devices, (**4**) EEG amplifier, (**5**) desktop receiver of physiological data, (**6**) router physically connected to PlayStation 4 and the desktop receiver, (**7**) monitor displaying the visual output from PlayStation 4. The image has been intentionally left dark to approximate the lighting conditions of the testing room.

All incoming messages were stored into .csv files, with each event listed in order of reception on one column and the time it was received on a second column. Date time was dictated by the receiving computer in Coordinated Universal Time (UTC), where the physiological data were also acquired. This approach helped demarcate both ends of each game round and the overall game session, after each event time was subtracted from the start of the recording time. The raw physiological data were exported from the signal recording software and converted into MATLAB-compatible files. These data were used to segment the continuous signals into 30 data structures (as many as the game rounds), and then further partitioned into shorter windows. The same procedure was employed for the baseline recordings, except they were not linked to any game events nor were they partitioned.

Figure 4: Visual simulation of the experiment setup for virtual reality data acquisition in Unity 3D. (**5**) Desktop receiver of physiological data, (**6**) router physically connected to PlayStation 4 and the desktop receiver, (**7**) monitor displaying the visual output from PlayStation 4, (**8**) monitor displaying the physiological data acquired in real time.

3.4.3 Electrooculography Preprocessing and Features

To detect eye blinks, the raw data were first centered around zero (**a**). Then, a notch filter was applied to remove 50Hz power interference (**b**). The resulting waveform was band-pass filtered with a third-order Butterworth filter of 1 to 10Hz (**c**). To correct severe baseline drifting, a 20^{th} order polynomial curve was fit into the data and then subtracted from (c) (**d**). In order to maintain the morphology of longer blinks, a moving mean filter was applied with a window length of 100 samples (**e**). The signal was later first-order differentiated (**f**). To remove very low peaks, which can be interpreted as noise, (f) passed through a test to detect samples beyond three interquartile ranges above the upper quartile or below the lower quartile (**g**). Due to the blinks' amplitude prominence, blinks would most likely fall within the estimated outliers. Hence, all non-outliers were zeroed out.

The last step was to locate all positive and negative peaks, by detecting changes from positive to negative values (**h**), as proposed by Kong and Wilson (1998). Each sample was processed in a loop to retrieve the left bound (positive peak) and the right bound (negative peak). The absolute difference between the left and right bounds is considered to reflect the blink's duration (Kong & Wilson, 1998). An example of the blink detection approach can be seen in **Figure 5**.

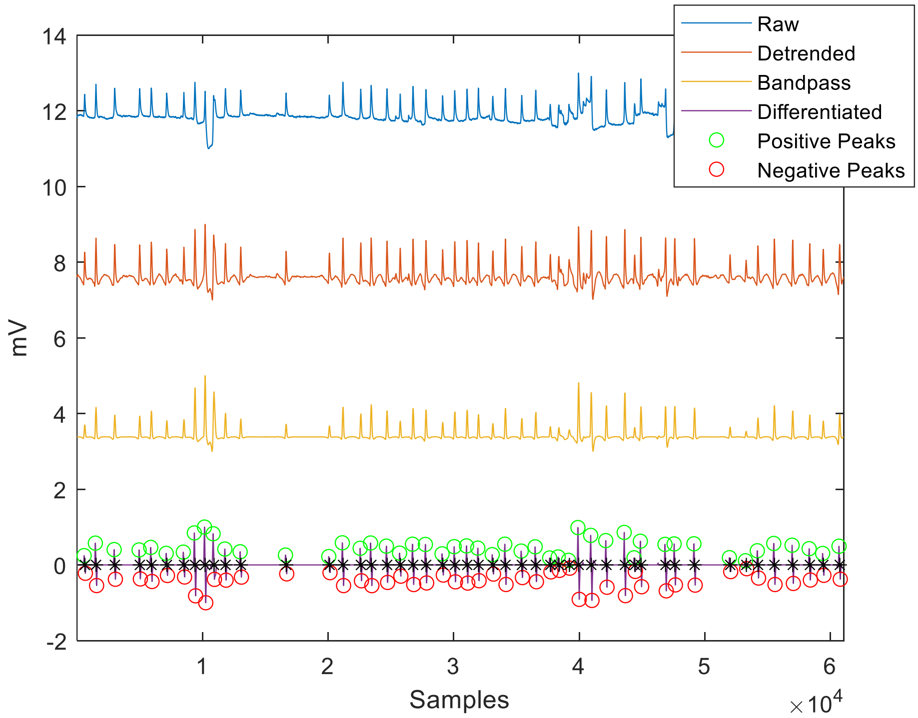


Figure 5: Blinks detected in a sample EOG signal.

The computed features from the post-processed signal are displayed in **Table 1**. The mean inter-blink interval (IBI) was computed as the mean time difference between successive blinks from their preceding. Eye blink rates (EBR) or blinks per minute (BPM) were computed through the count of blinks multiplied by 60 and divided by the length of window. This was done to normalize EBR by the length of the window in one minute. The amplitude (AMP) was acquired from the blink locations (x) and the y value of the sample at x from the normalized filtered signal in the (e)th step. Finally, the duration of the blinks (DUR) was calculated from the absolute difference between the positive and negative peaks of each identified blink (Kong & Wilson, 1998).

Table 1: Features extracted from electrooculography for data analysis.

Feature	Units	Description
Mean Inter-blink Interval (bIBI)	Seconds	The average time difference between consecutive eye blinks (higher values indicate reduced blinking rate)
Min Inter-blink Interval (minbIBI)	Seconds	The minimum time difference between consecutive eye blinks within one window
Max Inter-blink Interval (maxbIBI)	Seconds	The maximum time difference between consecutive eye blinks within one window
STD Inter-blink Interval (stdbIBI)	Seconds	The standard deviation of time differences between consecutive eye blinks within one window
Eye Blink Rate (EBR)	Count/Min	The average number of blinks per minute, normalized by the window's length
Mean Amplitude (AMP)	Normalized mV	The average amplitude of blinks within one window, with min-max normalization of the signal
Mean Duration (DUR)	Seconds	The average duration of blinks within one window

3.4.4 Electrocardiography Preprocessing and Features

Although the pipeline used for peak detection in the EOG data was also able to detect QRS-complex peaks in ECG data, it was decided that a more standardized detection algorithm, specific to ECG, was preferable. Hence, the raw ECG data were processed using an algorithm proposed by Pan and Tompkins (1985), which has been implemented in MATLAB and is publicly available (Sedghamiz, 2014). The resulting data are inter-beat intervals (IBI), which are the time intervals between pairs of successive beats (see **Figure 6**). The process described by Pan and Tompkins applies multiple filters to the signal, then differentiates and squares the output and finally applies a moving window integration. The authors describe this as the learning phase, followed by the detection stage, where an adaptive threshold is implemented as a means of reducing noise. The parameters to identify the QRS morphology include the slope, width and amplitude of the QRS complex (for more details, see Pan & Tompkins, 1985). Inter-beat intervals were further examined for ectopic beats using a threshold of 0.2, i.e., intervals changing by more than 20% of the previous (Malik & Camm, 2005). Ectopic beats are treated as abnormal beats and not necessarily as products of the autonomic nervous system (Choi & Shin, 2018); hence, they were removed. Typically, IBIs after ectopic beat correction are referred to as NNs or normal sinus beats (Shaffer & Ginsberg, 2017).

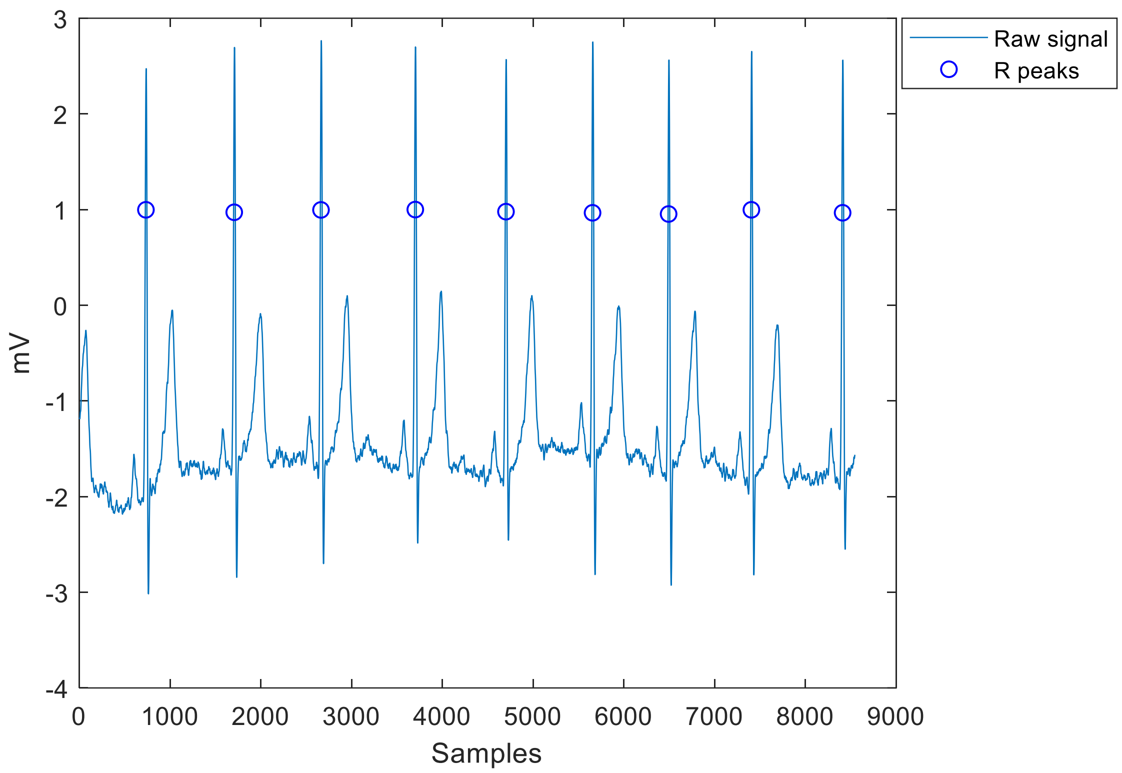


Figure 6: R-wave detection in a sample ECG signal using Pan & Tompkins QRS detection algorithm (1985) and implemented in MATLAB by Sedghamiz (2014).

A common practice in heart rate variability analysis is the extraction of features from the frequency domain (e.g., Shaffer & Ginsberg, 2017). Two terms of interest are short-term and ultra-short-term HRV. The former refers to the analysis of segments longer than 5 minutes, while the latter refers to segments that are less than 5 minutes in duration (Pecchia et al., 2018).

For frequency-domain HRV, there are variable durations of segments recommended. The high-frequency power requires a minimum of 1-minute segments (ultra-short), the low-frequency power requires a minimum of 2-minute segments (ultra-short), the very-low-frequency power requires a minimum of 5-minute segments (short-term), whereas the ultra-low-frequency power requires a minimum of 24-hour segments (long-term) (Shaffer & Ginsberg, 2017). However, the duration of the rounds in the TD-VR game depended on the performance of the player and could be as short as 20 seconds long, thereby making the computation of frequency-domain measures potentially unreliable (Pecchia et al., 2018). However, certain time-domain measures are also inapplicable in this work. For example, SDNNI, SDANN and HRV Triangular Index are best computed in 5-minute segments (Shaffer & Ginsberg, 2017), which exceed the maximum round duration in our study. The final computed features from the post-processed signal are displayed in **Table 2**.

Table 2: Time-domain features extracted from electrocardiography for data analysis.

Feature	Units	Description
Mean Inter-beat Interval (IBI)	Seconds	The average interval between successive heart beats. Heart beats are considered the R-peaks from the QRS complex. Typically, longer intervals denote a state of relaxation.
Min Inter-beat Interval (minIBI)	Seconds	The minimum inter-beat intervals
Max Inter-beat Interval (maxIBI)	Seconds	The maximum inter-beat intervals
Standard deviation of corrected Inter-beat Interval (SDNN)	Seconds	The standard deviation of NNs
Heart rate (HR)	Count/Min	Heart beats per minute
PNN50	%	Percentage of successive inter-beat intervals that differ by more than 50 ms
RMSSD	Seconds	The root mean square of successive inter-beat interval differences

3.4.5 Electroencephalography Preprocessing and Features

The EEG signal was first converted from the *vhd* format to MATLAB-compatible files using the EEGLAB toolbox (Delorme & Makeig, 2004) and the BVA-io plugin [4]. First, the signal was downsampled to 250 Hz (see e.g., Katahira et al., 2018) and digitally re-referenced to the right mastoid. Then, it was bandpass-filtered from 0.1 – 30 Hz using a Hamming-window based windowed sinc FIR filter (Widmann, 2006). Baseline correction was further applied to the filtered signal and Independent Component Analysis (ICA) was used to extract signal components for artifact attenuation. The components were visually inspected for each participant for the rejection process and cross-referenced with ICLabel (Pion-Tonachini et al., 2019), which is a tool that provides an approximation of the component sources (e.g., neural source, heart rhythm, eye blinks, motion, etc.).

The processed signal was split into segments corresponding to the game's rounds, using the triggers file sent from PlayStation 4 and stored in UTC date time on a desktop receiver, where the physiological data were also transmitted. Triggers were converted into milliseconds and then divided by four to procure the adjusted onsets after the signal was downsampled. For the onset of each round, the nearest highest integer was used for millisecond approximation, whereas for the end of each round, the nearest lowest integer was used. This was made to ensure that that the rounding of the onsets would always fall within the corresponding round's data. For each round, the processed signal was submitted to a Welch periodogram with a Hann window of 50% overlap to extract the absolute and relative power from the bands of interest. The window size was set to the ratio of the lower edge of the band of interest, divided by two, and multiplied by the sampling frequency (250 Hz).

The relative power is expressed as the percentage of presence of certain frequencies in the signal. It is computed by the ratio of the absolute power of each band over the total power of the signal. The relative powers were thus extracted for all channels and were preferred over absolute in that the latter may reflect variability due to extraneous factors, such as skull

[4] Available at https://sccn.ucsd.edu/wiki/EEGLAB_Extensions_and_plug-ins.

thickness (Hagemann et al., 2008). Following the findings of previous studies, the bands of interest were narrowed to **theta** (4-8 Hz), **alpha** (8-12 Hz) and **beta** (12-30 Hz), which have been extensively used in studies of flow (e.g., Berta et al., 2013; Katahira et al., 2018; Yelamanchili, 2018). The features in **Table 3** were derived from all EEG channels and then later reduced to a minimal set (more details in **Chapter 6**).

Table 3: Frequency-domain features extracted from electroencephalograms for data analysis.

Feature	Units	Description
Relative Theta	%	Percentage of absolute theta power over the total summed power of theta, alpha, beta. Frequency range: 4-8 Hz.
Relative Alpha	%	Percentage of absolute alpha power over the total summed power of theta, alpha, beta. Frequency range: 8-13 Hz.
Relative Beta	%	Percentage of absolute beta power over the total summed power of theta, alpha, beta. Frequency range: 13-30 Hz.
Relative Theta/Beta	Ratio	Also known as the attention ratio index, theta/beta is used to assess attentional and cognitive control.
Absolute Alpha	μV^2	Used to assess hemispheric asymmetric changes across prefrontal sites, as a marker of motivation and affect. Frequency range: 8-13 Hz.

3.5 Data classification and clustering

3.5.1 Introduction

To identify patterns across different classes of data, research typically employs Machine Learning methods. Though there is an abundance of existing techniques that approach each problem differently, classification algorithms usually attempt to separate observations by maximizing their differences or dissimilarities (e.g., Neelamegam & Ramaraj, 2013; Xiong & Cherkassky, 2005). This is done to reduce classification ambiguity, i.e., the probability that the membership of a sample point is truly representative of a specific class. This training stage is followed by the testing phase, where the classifier is used to predict unseen data. The quality of predictions determines the classifier's generalizability (Gunn, 1998).

This sequence can be used for internal validation, known as cross-validation, that estimates the classifier's performance and calculates the generalization error. It can be done by splitting the data into two parts of unequal size (hold-out method), whereby the e.g., 80% will be used for training and 20% for testing. Alternatively, an equally popular approach is the k-fold cross-validation, where, at its simplest form, the data are randomly partitioned into k sets. The classifier is thus trained k times, and its performance is determined by the mean accuracy reported by all folds. Because the retrieval of samples to populate each fold is random, the mean accuracy of a classifier trained with k-fold cross-validation is not constant (Wong, 2015). Hence, by controlling the randomization seed during the training phase, one can generate reproducible results.

There are other, more intensive cross-validation methods, such as leave-one-out, which only reserve a single observation as validation data while training with the remaining n-1 samples (Kearns & Ron, 1999). These approaches are used in addition to ruling out overfitting problems (Faber & Rajko, 2007), partitioning contingencies (Cawley & Talbot, 2007), but also as better estimators of the classifier's performance (Witten et al., 2016). Importantly, the training stage is vital. If the input data used for training are not efficient exemplars of each class, the separation of the classes may be inadequate (e.g., Cortes et al., 1995; Kavzoglu, 2009). As a result, the predicted outcomes can be highly unreliable.

In this project, we specifically focused on the training stage and explored parameters that could offer enhanced separation between low- and high-flow episodes. These parameters included variable window sizes for feature extraction, as well as injection of personality scores in order to influence the magnitude of each observation. In addition, we explored a combination of supervised learners and clustering techniques to evaluate game design choices based on the players' physiological arousal. In the next section, an overview of the algorithms used in the present project is provided, with a brief description of how they work and how they have been applied in video games research.

3.5.2 Overview of classification algorithms

3.5.2.1 Support Vector Machine

Introduction. Support vector machines (SVM; Cortes & Vapnik, 1995) have been widely used in the classification of mental states during video game playing (e.g., Fortin-Côté et al., 2018; Parsons et al., 2015; Plotnikov et al., 2012) as well as neighboring contexts, such as emotion detection (e.g., Mohammadi et al., 2017; Rakshit et al., 2016). Thus, it is a natural choice for addressing our multi-class classification problem. SVM belong to the family of supervised learning (Caruana & Niculescu-Mizil, 2006), which indicates that the class of each observation is known during the training process. SVM is widely applicable, as it can support non-linear and high-dimensional data (Paiva et al., 2018).

For example, heart rate variability parameters derived from electrocardiography data are known for non-linearity and SVM kernel functions can help separate them with high accuracy (Ghorai & Ghosh, 2017). There are several kernel functions that can be used for SVM models, including Gaussian or Radial Basis Function (RBF), linear, quadratic and cubic. The aim of these functions is to optimally fit a hyperplane surface to the training data, by maximizing the separation boundaries between *N* classes (Di Bono & Zorzi, 2008) and by mapping the data to a potentially high-dimensional space (Lotte et al., 2007). Those boundaries are called support vectors (Meyer, 2019). Support Vector Machine benefits from the regularization parameter, which makes it more tolerant to outliers and errors, allowing for higher generalizability (Lotte et al., 2007).

SVM in video games research. Support Vector Machines have been widely applied in video games research. For example, Plotnikov and colleagues (2012) classified different levels of difficulty in Tetris using the average band power from alpha, gamma, delta, beta and theta bands. The authors found an average 81% accuracy in a binary SVM using the RBF kernel when classifying each participant separately. However, an observable decrement of performance was observed when four different levels were added into a multiclass SVM, which resulted in an average accuracy of 57%. Likewise, an HRV-based feature set was able to discriminate between low- and high-flow episodes with an average accuracy of 72.3% in

Rißler et al. (2018). Berta and colleagues (2013) applied SVM on electroencephalography data in an aircraft battle game. The authors reported accuracy rates of 50.1% and 66.4% for user-independent and user-dependent classification respectively in the assessment of flow, boredom and frustration. SVMs were therefore chosen in this work due to their wide application and robustness in games and other areas of research.

3.5.2.2 K-Nearest Neighbors

Introduction. K-nearest neighbors (kNN) is another widely used non-parametric algorithm that is preferred for its simplicity. It is also common in affective decoding (e.g., Seo et al., 2019 as detailed next). The algorithm holds a data point and calculates its distance from the remaining data points, whereby it finds the nearest neighbors (Hu et al., 2016). Euclidean distance constitutes the most used distance metric, but other metrics have also been implemented, such as cosine, correlation and city block (see Chomboon et al., 2015). A known challenge with kNN is the occasionally arbitrary selection of k, i.e., the number of nearest neighbors, which must be decided a priori (Guo et al., 2003).

The parameter k may determine the admissibility of a specific data point to a class, which can affect the classifier's performance (Lim & Chia, 2015; Zhang et al., 2017). Lall and Sharma (1996) suggested that the square root of the overall number of observations could be used as an optimal k. In addition, the input data have been shown to yield most effective classification rates when they are z-normalized compared to other standardization techniques (Mohamad & Usman, 2013). However, kNN appears to underperform in presence of high-dimensional data (Sedeeq & Romano, 2014). Nevertheless, the advantages of kNN include the ability to update the training samples online but also its flexibility in resolving multiclass problems (Yazdani et al., 2009).

K-NN in video games research. K-nearest neighbors is also a widely popular algorithm in affective and video games research. For example, Seo and colleagues (2019) used EEG to discriminate between boredom and non-boredom. The authors found that activity from the beta and gamma bands were most helpful in identifying boredom and used them as input to a kNN classifier, which achieved an average accuracy of 86.73%. Similarly, in a study by

Jimenez and colleagues (2011), the authors were able to identify when a player was engaged in a game or was relaxed via EEG frontal activity alone.

The performance of k-NN reached 84.26% with k = 5 and 82.10% with k = 10. Basic emotions and complex affective states have also been investigated using k-NN. For instance, joy, anger, sadness, grief, and hatred have been classified using electromyographs (EMG) and respiratory data, with respiration occasionally outperforming EMG (e.g., Meftah et al., 2012). Likewise, Zheng and colleagues (2018) used k-NN to identify states of high-stress and relaxation with an average accuracy rate of 86.7% and 91.26% in Rahman et al. (2015). Finally, Benlamine and colleagues (2017) were able to identify instances of player motivation for mastery and performance using different classifiers. Though k-NN was outperformed by a random forest classifier, it still achieved a promising F1-score of 73%. In this work k-NN was used due to its fast computation and simplicity and its promising performance in the literature.

3.5.2.3 Ensemble of CART Decision Trees

Introduction. Ensemble of CART (Classification and Regression Trees) decision trees are a less popular learning technique in video games research. However, they have been shown to yield high accuracy rates in EEG (Sun et al., 2007), ECG (Mert et al., 2012), and EOG (Banerjee et al., 2014) signal classification. Therefore, in this study, we considered this technique as a candidate for flow classification.

An ensemble is essentially a cluster of multiple decision trees, whose classification boundaries rely on a voting process amongst individual classifiers (Dietterich, 2000). Thus, the advantage of ensemble decision trees is their classification superiority (Dietterich, 2000) and reduced error rate (Hansen & Salamon, 1990) compared to a single decision tree. Ensemble learning implements different algorithms, such as bagging (Breiman, 1996) and boosting (Freund & Schapire, 1996).

Bagging produces subsets (replicates) of the original training set by drawing random samples with replacement in a way similar to bootstrapping, allowing each observation to be selected more than once during the sampling of each subset (Dietterich, 2000). The number of the replicates determines the number of the learning iterations (Quinlan, 1996). The final classifier, the ensemble, outputs the collective decision predicted by the individual classifiers (Bauer & Kohavi, 1999). On the other hand, boosting entails the use of moderately accurate individual classifiers (error rate less than 50%) that are combined to create a stronger learner (Drucker & Cortes, 1996; Freund et al., 1999). Unlike bagging, where the "voting" weights remain constant, boosting applies increased weights to misclassified instances, but lower weights to correctly classified training examples (Dietterich, 2000; Quinlan, 1996). The final classifier is the sum of the applied weights; through an iterative process, the algorithm aims to minimize the weighted error (Dietterich, 2000), by relearning the misclassified examples.

Ensemble CARTs in video games research. Although ensemble learning has been rarely applied in the classification of player experience (e.g., Geisler, 2002), a study has shown that ensemble decision trees may outperform SVMs when using either bagging or boosting methods (Caruana & Niculescu-Mizil, 2006; Vo et al., 2019). Notably, there are also experimental approaches in bagging SVMs, instead of decision trees, which have been similarly shown to outperform a single SVM learner (Kim et al., 2002). Oh and colleagues (2015) investigated cognitive workload in a flight simulator under different levels of challenge. The authors found that both bagging and boosting achieved an average of 88% and 95.7% accuracy rates respectively using different metrics, such as EEG and time-frequency HRV. In a related work, mental stress was best classified using an ensemble of SVMs (77.05% accuracy) compared to an individual SVM or others, e.g., Naïve Bayes or Artificial Neural Networks that remained in the range of 70% or less (Yin et al., 2016). In the context of the flow experience in video game play, our work is perhaps the first to decode flow episodes using an ensemble learner.

3.5.2.4 Hyperparameter selection

The best combination of the classifiers' hyperparameters were determined using Bayesian optimization (Mockus et al., 1978). This method sequentially evaluates an objective function such that it minimizes a target criterion. In the estimation of the optimal hyperparameters, this criterion would be the misclassification rate that is reported from the objective function, i.e., the cross-validation (e.g., Snoek et al., 2012).

What is particular to this method, compared to grid-based or random searches, is that it maintains an acquisition function, which is maximized through a Gaussian process model in order to determine the next evaluation point. The acquisition function identifies whether the next evaluation point is appropriate based on a posterior distribution function (e.g., Gelbart et al., 2014). Hence, Bayesian optimization was preferred over grid or random search, due to its probabilistic sensitivity to previous evaluation tests (Oh et al., 2015). Notably, a recent study has also demonstrated that Bayesian optimization for SVM-based classification outperforms grid search or heuristic choice (Czarnecki et al., 2015). SVMs were optimized through the regularization parameter C and kernel scale. For the BDT ensemble, fine-tuning was performed on the number of weak learners and the minimum leaf size. Finally, for k-NN, the parameters tested were the number of nearest neighbors and the distance metric (Guo et al., 2003; Lora et al., 2003).

3.5.2.5 Balancing class sizes

With respect to the multiclass classifiers, the number of observations per class was inhomogeneous. In TD-VR, flow was designed to be elicited in 16 trials, whereas anxiety and boredom were designed for seven trials each. Imbalanced data may result in a bias toward the majority class (i.e., the class comprising the highest number of observations), resulting in weak classification precision for the minority class(es) (Bhowan et al., 2012). To balance the classes in the multiclass scenario, the Synthetic Minority Oversampling Technique (SMOTE; Chawla et al., 2002) was used. This type of intervention is external, unlike the internal which modifies the classifier's algorithm to incentivize the consideration of the class imbalance issue (Estabrooks et al., 2004; Guo et al., 2008).

This data augmentation algorithm operates on k-nearest neighbors to create additional samples for the minority classes. It computes the difference between the nearest neighbor and the existing sample and multiplies by a random number in the range of zero to one from a uniform distribution. This new value is then appended to the existing feature space as a synthetic data point (Chawla et al., 2002).

Stratified sampling was further employed for partitioning the data into k folds given the imbalance of classes (He & Ma, 2013), which preserves the original class distribution across the folds. Data from each subsequent fold were reserved for validation, whereas the data using the remaining $k - 1$ sets were oversampled using SMOTE. Then, a model was fit using both the synthetic data and the original training examples. Hence, the synthetic observations were not validated during cross-validation but were instead used to complement the training sets. The number of observations equally represented each class after SMOTE, but the classifiers' performance measures reported refer to real data and thus are not artificially inflated. The resulting prediction for each validation fold was matched against the original class the observation belonged to. This procedure was repeated k times.

Notably, the ratio of the majority class (flow) to the minority classes (anxiety or boredom) was 2.3:1. Following the application of SMOTE on the training data, the three classes were represented homogeneously. Indeed, SMOTE defeats data sparsity and can shift the distribution skew of the dominant class (Taft et al., 2009). The overall classification performance metrics computed for the three-way classification of flow, anxiety and boredom were the accuracy, specificity, precision, recall and F1 score. The latter three are sensitive to false positives. No misclassification bias was observed in the minority classes, as expected in this approach (see results in the following chapters).

These metrics were extracted through macro-averaging, as a precaution to the stratified sampling. Macro-averaging treats all classes equally, by simply computing the arithmetic mean of the performance metric for each individual class with respect to the rest, that is, dividing by the total number of classes (Sokolova & Lapalme, 2009). Contrarily, micro-averaging considers the number of true positives, true negatives, false negatives, and false positives in all classes at once; hence, it is more sensitive to and favors larger classes or the majority class (Sokolova & Lapalme, 2009; Tran et al., 2018). The data augmentation and stratified cross-validation, along with the macro-averaging reporting were meant to alleviate any bias toward the flow class. Likewise, metrics other than accuracy that involve false positive rates are more effective against class imbalance (Guo et al., 2008), thereby supporting the classifier's evaluation integrity. As a note, the issue of class imbalance was exclusive to the multiclass scenario, not the binary classifiers that follow in the next chapters.

3.5.3 Backtracing misclassification rates with respect to the game's design

3.5.3.1 Introduction

Given that the classification solutions employed in this project are supervised, the ground truth should be soundly provided. However, due to inter-player differences or other factors, such as time, the ground truth may not be guaranteed; erroneous assumptions may therefore impact the ability of an algorithm to efficiently classify the input data. As argued earlier, different players may respond differently to the same game events. For this reason, a clustering approach was employed to investigate potential misclassification rates.

Clustering methods aim to group observations together based on their distance or similarity of the comprising samples and label those data independently in an unsupervised manner (Xiao & Yu, 2012). A known issue with clustering techniques is that it may be hard to interpret what each cluster represents (Hotho et al., 2002), which has also been pointed out in the domain of player behavior modeling (e.g., Drachen et al., 2012). However, the combination of unsupervised and supervised learners may be a promising approach for evaluating tentative ground truth, as in the case of video game design. In this work, the approach was regarded a necessary step to assess game design choices as per their tendency to yield variable effects on player experience. In the following studies, the condition (flow/anxiety/boredom) and the effect of time (the time point when each condition was introduced) may not consistently meet the assumptions of a specific class.

3.5.3.2 Unsupervised learner: Fuzzy C-means

Fuzzy C-means is a soft clustering technique (Panda et al., 2012). This is different from other clustering methods, such as k-means and k-medoids, which are hard clustering techniques (e.g., Patil et al., 2014). Soft clustering acknowledges that the membership of an observation may not be exclusive to one cluster, as hard clustering methods assume (Khanmohammadi et al., 2017), but may be shared among the desired number of clusters with a degree of membership (Nascimento et al., 2000). Considering that anxiety may not necessitate the absence of flow, but rather a mixture of both states (e.g., Borderie & Michinov, 2016; Roest & Bakkes, 2015), a single class membership may not provide efficient information to the classifier. Given these theoretical observations, these states may be "gradient" and perhaps better addressed with a soft clustering technique. The aim was to understand which classes could fail to provide consistent results and, more specifically, to identify subjective characteristics that may have underlain less reliable outcomes (i.e., increased misclassification rates).

The idea behind FCM (Bezdek, 1981) is that it randomly initializes the center vectors V, whose length is equal to the number of tested clusters C. The iterative process underlying FCM updates the membership degree of each data point x_j as well as the center vectors V; this is done until a termination criterion ε is achieved through m iterations. The matrix U contains the degree of membership of each data point x_j relative to each cluster C_i. A high-degree membership indicates that the data point x_j is close to the center vector V_{C_i}, whereas a low-degree membership indicates that the data point lies far from that center. The degrees of membership U_{ij} for each data point x_j range from 0 to 1. The purpose is to minimize the distance (e.g., Euclidean, Mahalanobis, Cosine) between the data point x_j and V_k, weighted by the membership degree U_{ij}.

Although three class outcomes are of interest, that represented flow, anxiety and boredom, the number of desired clusters was instead evaluated using the Davies-Bouldin index (DBI; Davies & Bouldin, 1979). This was done to prevent the enforcement of our bias on the cluster separation criterion. The computation of DBI entails the ratio of within-cluster compactness over the between-cluster separation. Thus, DBI attempts to maximize the intra-cluster similarities (Sitompul et al., 2019). The lower the value of the Davies-Bouldin index, the better the clustering solution is considered (Davies & Bouldin, 1979). Hence, to select the optimal number of clusters C, a function tests through various cluster numbers from a range of values $C = \{1, \dots, N\}$, computes the Davies-Bouldin index at each iteration $DB_i = \left\{DB_{C_1}^{i}, \dots, DB_{C_N}^{n}\right\}$, and selects C corresponding to the minimum Davies-Bouldin index, min(DB), which is considered the function's criterion.

3.5.3.3 Combining Fuzzy C-means and self-report ratings

The input data to FCM were extracted from the three design considerations separately. Thus, FCM was performed three times on the feature spaces of flow, anxiety and boredom sample data. Using DBI, the optimal number of clusters C was determined for each condition at an inspection range of 1-6. For each cluster C_{ij}, where j is the condition from which the data set was extracted, the degree to which each participant was represented in that cluster was evaluated. The degree was defined as the number of all instances belonging to the p^{th} participant at C_i over the total number of p^{th} observations in j. In other words, the degree of representation is the proportion of participant observations in each cluster i for condition j. This enables to further identify cohesive subsets of patterns within each condition (round type) with different characteristics. The hypothesis is that the emergence of different clusters from the physiological responses may explain poor decoding of the round type when all data are pooled.

To interpret the clustering solution of FCM, and the role of each cluster, the user self-report ratings were used as explanatory measures. The proportions were submitted to a stepwise linear regression as a dependent variable, and FSS-2 and BFI ratings as independent variables. For each significant relationship, the direction of the slope, i.e., the sign of the beta coefficient, was considered (a). Next, depending on the self-report dimension that was found to significantly predict the proportions of observations for each participant, the normalized rating of that dimension for the corresponding participant was obtained (b). The normalization consisted of the rating of the participant for the self-reported dimension divided by the maximum rating that the self-reported dimension can take. Each observation in the original data set X was finally multiplied by the coefficient value (b) with the sign from (a). The resulting coefficient was thus $(1 + b)$ for $a > 0$, and $(-1 - b)$ for $a < 0$. The steps of this approach are shown in **Figure 7**.

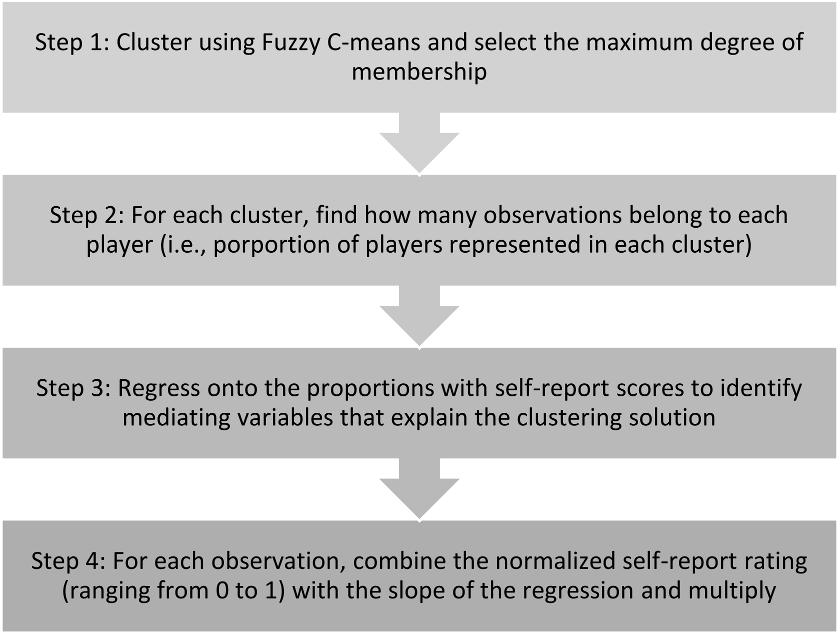


Figure 7: Summary of the steps taken for the deployment of an unsupervised learner to uncover physiological patterns from HRV, EBV and EEG.

The purpose of this approach is to elucidate potential latent factors that may have underpinned experiential variability in the game. These factors are then intermixed with the actual sample data to enhance their separability. This approach can be seen in chapters 5 and 6. However, prior to the examination of the physiological correlates of flow, anxiety and boredom, a pilot study was conducted to evaluate the effectiveness of TD-VR and to confirm the putative relationship of flow with the Big Five personality dimensions of agreeableness, conscientiousness, extraversion, neuroticism and openness. The study is detailed in the next chapter.

4 Study I: Agreeable individuals are more likely to experience flow

4.1 Introduction

The flow state appears to be experienced similarly across individuals and different types of activity (Nakamura & Csikszentmihalyi, 2014). However, the experience may not be ubiquitous to every person – approximately 15% of Americans and 12% of Germans have reported that they have never had such an experience (Csikszentmihalyi, 1997a). Earlier research has indicated that there may be certain particularities that could hinder the expression of this state. For example, temperament and personal characteristics may influence the likelihood of experiencing flow (Teng, 2011).

Others have reported that proneness to flow is related to the density of striatal, dopaminergic receptors (De Manzano et al., 2013; Mosing et al., 2012b), and more specifically the putamen (De Manzano et al., 2013), as well as genetic influences (Mosing et al., 2012a, b). The study of flow in video games research is rarely accompanied by supplementary measures that could be used as surrogate indicators of flow proneness. For example, it may be challenging to assert whether reports of low flow were attributed to the activity characteristics or the individual's proneness to flow. Although the interaction of the activity and the person are expected to determine the ultimate flow experience (Chen, 2007; Nah et al., 2014), individual characteristics require more attention.

Five major personality traits have been studied extensively and have been used to identify core psychological qualities of individuals (Costa et al., 1991; John & Srivastava, 1999). It has been suggested that each personality trait is an abstract construct that comprises multiple attributes of behavioral specificity (Costa & McCrae, 1995). **Openness to experience** has been associated with curiosity, imagination, artistic interest, creativity and willingness to try new activities and ideas (McCrae, 1987; Proctor & McCord, 2009). **Conscientiousness** seems to relate to planning, self-discipline, goal orientation, achievement, responsibility and perseverance (Costa & McCrae, 1995; Golbeck et al., 2011; Kern & Friedman, 2008). Individuals with high levels of **Agreeableness** are cooperative, empathic, social and altruistic (Bakker et al., 2006; Graziano & Tobin, 2009; Graziano et al., 2007; Jensen-Campbell & Graziano, 2001). Extraverted individuals, i.e., individuals characterized by increased levels of **Extraversion**, seem to excel in interpersonal contexts, are talkative, sociable, assertive and achieving (Depue & Collins, 1999; John & Srivastava, 1999; Wilt & Revelle, 2017). Finally, **Neuroticism** is a construct affiliated with many psycho-behavioral patterns, among which are anxiety, emotional instability, irritability, impulse and susceptibility to criticism (Lahey, 2009; Watson et al., 1994; Widiger, 2009).

With regard to the flow experience, the personality traits Conscientiousness (Johnson et al., 2014; Tatalović Vorkapić & Gović, 2016; Ullén et al., 2012, 2016), Extraversion (Tatalović Vorkapić & Gović, 2016; Ullén et al., 2016), Agreeableness (Ullén et al., 2016), and Openness to Experience (Tatalović Vorkapić & Gović, 2016; Ullén et al., 2016; Bassi et al., 2014) have been found to predict and positively relate to flow, whereas Neuroticism seems to present a negative relationship (Heller et al., 2015; Johnson et al., 2014; Tatalović Vorkapić & Gović, 2016; Ullén et al., 2012, 2016). As discussed previously, neuroticism is typically escorted by elevated negative affect, which is counterproductive to flow, a state linked to well-being and positive affect (e.g., Collins et al., 2009). The remaining personality traits relate to goal orientation, creativity and positive affect, all of which have been shown to have a direct relationship with flow (Csikszentmihalyi, 1990, 1997c; Ghani, 1995; Ullén et al., 2016).

In addition to personality, age and years of video game playing have been examined in previous studies as antecedent qualities of the flow experience. These demographic characteristics have been shown to influence the self-reported levels of flow to varying degrees. Age of participants, albeit considered, has limited evidence that it can effectively influence the flow experience. For instance, teenagers aged between 15-19 years were found to report significantly lower levels of flow compared to the 30-39 age group (Sillaots et al., 2020). However, these findings should be carefully interpreted. The range of years considered in the younger group was smaller than the older group, and the number of participants per group highly dissimilar (N = 8 for 15-19 and N = 63 for 30-39). No other differences in flow were found among other age groups (from 20-29 to 60-69 years of age). Another study by Voiskounsky and Wang (2014) found that players younger than 17 years of age reported a significantly higher flow experience compared to older players.

For example, Fong and co-authors (2015) carried out a meta-regression analysis and could not infer linear effects of age as a continuous variable on perceived skills-challenge balance and flow. However, a small effect was observed when age was dichotomized instead – below 30 and above 30 years of age. The authors argued that a non-linear relationship between flow and age may be likely. Likewise, Payne and colleagues (2011) found no correlation between flow and age, and they suggested that flow does not seem to diminish with age. Perhaps, a clear picture is challenging to paint, considering that video games are popular across all age groups (Johnson et al., 2016; McLaughlin et al., 2012).

Contrary to age, the link between years of playing games and flow seems more consistent. Highly experienced players (as measured in years of experience) are more likely to experience flow than novice players (Wu et al., 2013). However, they have also been found to enter flow slower as their experience grows (Wallace, 1999).

In this study, the evaluation of a Tower Defense virtual reality game was made using the five major personality traits as well as self-reports of flow experience. It was hypothesized that flow would peak during the game wherein an average level of difficulty is introduced but drop during the easy and hard difficulty modes (**H_1**). Hence, difficulty of the game was modified in three dedicated conditions, with the aim of influencing flow's dimension of skills-challenge balance. A second hypothesis was that the normal difficulty mode would elicit higher positive affect to the players compared to the easy and hard modes (**H_{2a}**). Similarly, negative affect increases were expected to occur after playing the easy and hard modes, as a result of boredom or anxiety (**H_{2b}**). The third hypothesis was that personality dimensions would significantly explain the overall flow experience as well as its individual nine dimensions (**H_3**).

4.2 Methods

4.2.1 Participants

Although a remarkable increase of female players has been observed in the recent years, research findings and outputs are disproportionately derived from male participants (Wohn et al., 2020). Given the preliminary scope of this study, the participation criteria were limited to a focus group of males aged 18 to 45, as they constitute a major, if not the largest, demographic of video games-related targeted marketing (e.g., Chess et al., 2017; Williams et al., 2008). No familiarity with virtual reality and/or TD games were required. The study comprised 31 male participants ($M = 27.4 \pm 4.9$ years old) who were recruited through internal mailing lists in Sony Interactive Entertainment in London, UK, and students from Bournemouth University in Bournemouth, UK.

Participants reported an average video game playing experience of 15.1 years (SD = 8.03 years). The difficulty conditions comprised easy (N = 10), medium (N = 11) and hard (N = 10). Eleven participants self-identified as novice players (mean years of experience = 7.1 ± 4.5), whereas 20 participants identified as experienced video game players (mean years of experience = 19.55 ± 5.8). The experimental procedures were approved by the local ethical committee at Bournemouth University (ref: 13325).

4.2.2 Materials: Self Reports

The tools used to evaluate the players' experience included a custom player expertise questionnaire (see Appendix, Error! Reference source not found.), the Flow State Scale – 2 [General] (FSS-2; Jackson & Eklund, 2002), the Positive Affect and Negative Affect Schedule (PANAS; Watson et al., 1988), and the Big Five Inventory (BFI; John & Srivastava, 1999).

The PANAS was used to evaluate enjoyment and overall affective shifts, including negative affect. This scale has been previously used either verbatim or modified for games research (e.g., Bessière et al., 2007; Mathiak et al., 2013; Wang et al., 2008). PANAS assesses positive and negative affect via 20 items ranging from 1 ("Very slightly or not at all") to 5 ("Extremely"), depending on the applicability of the statement to the participant's mood at the time. A description of the FSS-2 and BFI questionnaires can be found in the **Questionnaires** section.

Analyses of the self-report scores were carried out using SPSS v23.0 (IBM Corporation, Armonk, NY). Data normality was investigated using the Lilliefors normality test. Significance level was set to α = .05.

4.2.3 Materials: Test Game

The virtual reality Tower Defense game has had three options for difficulty: Easy, Medium (Normal) and Hard. The selection of the menu option was done in front of the participants. However, to prevent potential effects on the player experience, the options were disguised using icons that did not expose their purpose. The participants were randomly assigned to either condition at the beginning of the experiment regardless of personal characteristics such as experience with video games.

Three tower types were used as options to build in the game: Fire (balanced tower, high attack damage), Disease (the lower the health of the attacked opponent, the higher the damage output of this tower) and Nova (multiple enemies were shot at once). Towers with special abilities (i.e., Disease and Nova) were given lesser base attributes, including attack power and attack speed, so that they were more balanced. All towers have had three upgrade levels, which made their attack power more powerful. However, Fire towers, which had no special ability, further received an increase in their attack range at each level. Disease towers have had a naturally long range, but less than a Fire Tower's. Nova towers have had a short range that did not increase per level. There were ten enemies per wave (footmen), except for the boss which was a single unit, and had 5 times the health of a unit from the immediately preceding game round. Bosses appeared every five rounds in the game.

The zones where the towers could be constructed were four, corresponding to the cardinal directions of North, South, East and West, relative to the player's view (**Figure 8**). To build a tower, the player had to gaze at any zone, select it, browse through the available grids in the zone, select their preferred grid and create the tower of their choice. Head tracking for head-driven input has been shown to be highly immersive in VR games (Martel et al., 2015). The zones would remain actively selected even if the player looked away from that zone – a spate of "back" key commands (Circle button on DualShock 4) had to be issued by the player in order to deselect the active area. A glowing hammer was used as a diegetic display to signal that a certain zone was selected. The hammer was selected for its semantic relatedness to the build and upgrade commands. A custom shader was used for the object, which allowed it to be rendered on top of other elements in the scene that would otherwise occlude it. As in other TD games, players were given the ability to accelerate the pace of the game whilst holding down a key (R2 on DualShock 4).

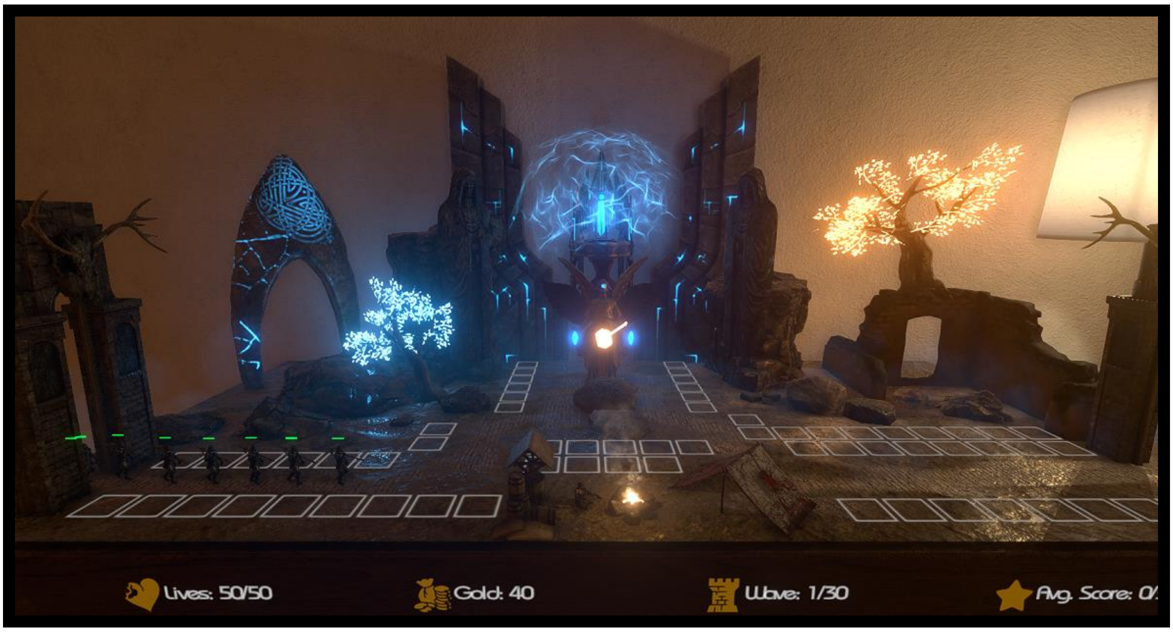


Figure 8: Game table area from Tower Defense VR where the game action takes place. The white grids are usable for tower construction. The enemies on the left have hovering health bars, making it easier to track the damage inflicted by the towers. The gate on the left is the entry point, whereas the gate on the right is the exit point. The enemies walk behind the portrayed statue in the middle and continue towards the exit. The glowing hammer, that follows the player's head rotation, aims to visually aid the head tilt sensitivity required to gaze across different regions for tower construction.

The surrounding environment in the virtual space (**Figure 9**) remained unaffected during gameplay.

Figure 9: Overview of the virtual space surrounding the game table from Tower Defense VR.

For the difficulty, the default health scaling formula that was tested with different users was multiplied by x0.7 for the easy mode, x1 for the medium mode and x1.3 for the hard mode. The received gold from eliminating enemies was also modified based on the difficulty, favoring higher levels of difficulty (proportional to the in-game round number). In addition, gold was passively added to the player's gold pool every eight seconds, the amount of which was equivalent to the worth of one opponent from the most recent round. Upon elimination, bosses granted eight times more gold than a regular footman. The players had to play through 40 rounds, with a 7-second inter-round delay but could be defeated at any point if all their 50 lives were lost. One life point was subtracted once an enemy escaped unless it was a boss. Players who made it through all the rounds were victorious. The outcome of the game was communicated to the player once the game ended. The text of "Defeat" was colored in red, which is a traditional color code used in games to signal undesirable game states (e.g., Zammitto, 2008).

4.2.4 Procedure

The experiment took place in Sony Interactive Entertainment Europe in London, UK and Bournemouth University, Bournemouth, UK. Participants were seated in front of a desk, on top of which the PlayStation VR kit was installed. A thin, cardboard box or a computer screen were used to rest the PlayStation VR Camera, which is used for tracking the VR head-mounted display. The height of the camera was adjusted until the participant felt that their in-game height felt natural. Prior to the experiment, participants were asked to fill in the following documents (arranged in order of administration): consent form, games experience questionnaire and PANAS. Participants were also let known that if discomfort were to be experienced in any way, the experiment would be immediately terminated. The participants were then given written instructions for the game's purpose and controls, regardless of any previous experience in Tower Defense games.

Once the participants reported confidence in recalling the buttons for each action in the game, they were asked to demonstrate the controls on the PlayStation controller (DualShock 4). The game session commenced after the completion of the aforementioned steps in accordance with the designated difficulty level of the randomized assignment (see above, **Materials: Test Game**). The experimenter received visual feedback of what each player was seeing in the game through a TV monitor connected to the PlayStation 4 console. This helped guide the participants during their first steps of their engagement with the game as well as the game's controls. If more than 10 minutes were lost in the training process, the experimenter would restart the game. Volunteers were informed that they would not be interrupted during the game. The participants did not receive any compensation for their participation, and they were thanked for their time and contribution.

When the player either won or lost the game, they were asked to fill out the final set of questionnaires. In order of administration, these were the PANAS, FSS-2 and BFI. As a note, the BFI could have alternatively been administered before the game, but prior to PANAS. The reason why it followed the game session was to spread out the experiment's activities more evenly across time. The PANAS questionnaire was administered in an AB design (or pre- and post-test), which entails double testing in time point A and time point B, with the game session mediating these points (e.g., Mathiak et al., 2011). PANAS was filled out prior to FSS-2, and immediately after the game session.

This decision was made to ensure that PANAS scores would specifically relate to the game experience instead of being polluted with the act of filling out another questionnaire. Multi-item scales, like FSS-2 with four items of varied rephrasing per subscale, may not be well perceived by participants (Robinson, 2018). Contrary to PANAS that asks about one's mood at the present time, such concerns were not present for FSS-2, given that the phrasing of each item relates specifically to the activity under study (i.e., the TD-VR).

4.2.5 Experimental design

For this study, a retrospective approach was adopted for reporting the flow experience, which we identified earlier to be less intrusive to the natural pace of flow (see *Research Approaches* in Chapter 2, **The challenges in measuring flow in video games**). This was done through a between-subjects experimental design, where participants were exposed to a single condition among the three difficulty levels (easy, medium, or hard). This allows us to observe how difficulty may mediate the relationship between personality and the self-reported experience of flow.

It is worth noting that we were not merely interested in exposing a relationship between personality and flow, but how this relationship is challenged by the game's difficulty level. The assignment to the difficulty conditions was randomized. To assess whether subjects assigned to each experimental condition were not different in a systematic way, a multivariate ANOVA was used, testing for difficulty group differences on the study's dependent variables, positive and negative affect at baseline - i.e., prior to the game session (see e.g., Rogatko, 2009; Suresh, 2011).

4.3 Results

In order to verify successful randomized assignment, a one-way MANOVA was used with the difficulty group as a between-subjects factor, and Positive affect pre-game (PANAS) and Negative affect pre-game (PANAS) as dependent variables. The results did not reveal any significant differences in either variable among the groups, suggesting that all groups can be treated equally.

The player's progress (round reached divided by 40) as well as the average score of each player were similarly tested against the difficulty group. These two variables, player progress and player average score, are distinct in that the former determines how far in the game the player managed to play, whereas the latter determines how well they did on average. A main effect of difficulty group was observed for both the progress [$F(2, 28) = 5.32$, $p = .011$] and the average score [$F(2, 28) = 6.43$, $p < .01$] in a one-way MANOVA. Subsequent pairwise comparisons revealed that the progress made was significantly different across conditions, except for easy versus medium ($p > .05$).

Players in the easy mode progressed an average of 11.3 (±3.54) rounds more than players in the hard mode ($p < .01$). Similarly, players in the medium difficulty group progressed an average of 7.7 (±3.46) rounds more than players in the hard mode ($p = .034$). Though the normalized player's progress and the average score accumulated in the game were naturally highly correlated ($r = .965$, $p < .001$), the effect of difficulty on the average score was slightly different in the pairwise comparisons. Players in the easy mode scored significantly higher than players in both medium ($p < .05$) and hard modes ($p = .001$). However, no significant differences were observed for medium versus hard modes. This was also confirmed with a chi-square test, where an association between win-loss condition and game difficulty was observed, $\chi^2(2) = 12.87$, $p < .01$.

4.3.1 Flow scores (FSS-2)

To test H_1, the flow scores from FSS-2 were analyzed using a one-way MANOVA. The difficulty condition in which the participants were randomly assigned was set as the independent variable and the flow dimensions from FSS-2 were added as dependent variables. Although the results showed a trend toward the expected outcomes (as per the inverted-U reported in Peifer et al., 2014), they did not yield a statistically significant main effect of game difficulty on any of flow's individual dimensions or their sum, i.e., the overall flow score (**Figure 10**). Hence, H_1 was not successfully rejected. These results were also true when controlling for age or years of experience.

Further, an independent-samples t-test was performed with the game's outcome as the grouping variable. Significant differences were found between flow's dimension of time perception distortion, $t(29) = 2.89$, $p < .01$. Winning players reported higher time perception distortion than players who lost. When the novice player group ($N = 11$) was removed from the analysis, a second flow dimension reached significance. Experienced players who won reported a higher balance between their skills and the game's demands, $t(18) = 2.28$, $p < .05$.

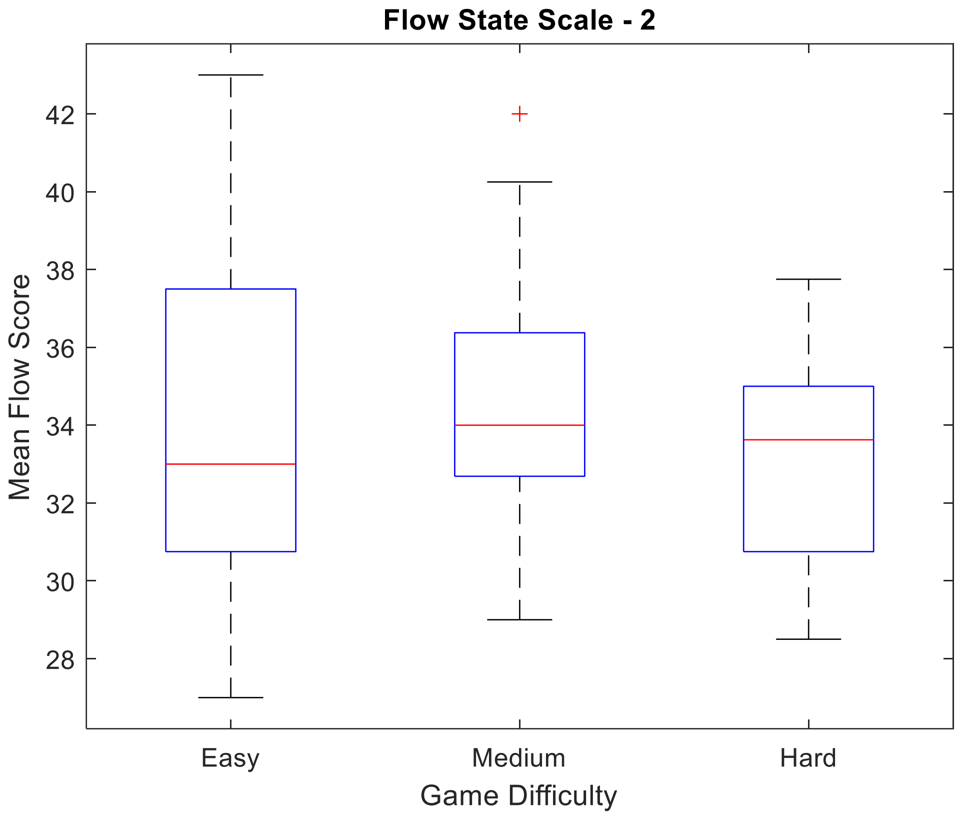


Figure 10: The overall flow score comprising the sum of all flow dimensions from FSS-2. There is a visible trend where reported flow score is maximized during the medium difficulty mode, which was designed to be optimal.

4.3.2 Positive and Negative Affect (PANAS)

To examine whether the game was enjoyable as a function of the game's difficulty (H_2), a one-way ANCOVA test was conducted using the pre-game positive affect (PANAS scores) as a covariate, the post-game positive affect as a dependent variable and the flow difficulty condition as between-subjects factor. All variables passed normality tests (Lilliefors). Levene's test was not significant (p = .148), indicating that the assumption of homogeneity of variance was met.

The test revealed a significant main effect of difficulty condition on the adjusted post-game positive affect scores, $F(2, 27) = 4.006$, $p < .05$, $\eta_p^2 = .229$. Follow-up pairwise comparisons without adjustment showed that the differences found occurred between the Easy and Hard difficulty conditions [$t(27) = 5.955$, $p = .025$] and Medium and Hard [$t(27) = 6.254$, $p = .017$], indicating that the easy and medium difficulty modes contributed to higher positive affect in comparison to the pre-game equivalent score (**Figure 11**). The same test was carried out for the negative affect scores of the PANAS scale, but they were not significant ($p > .05$).

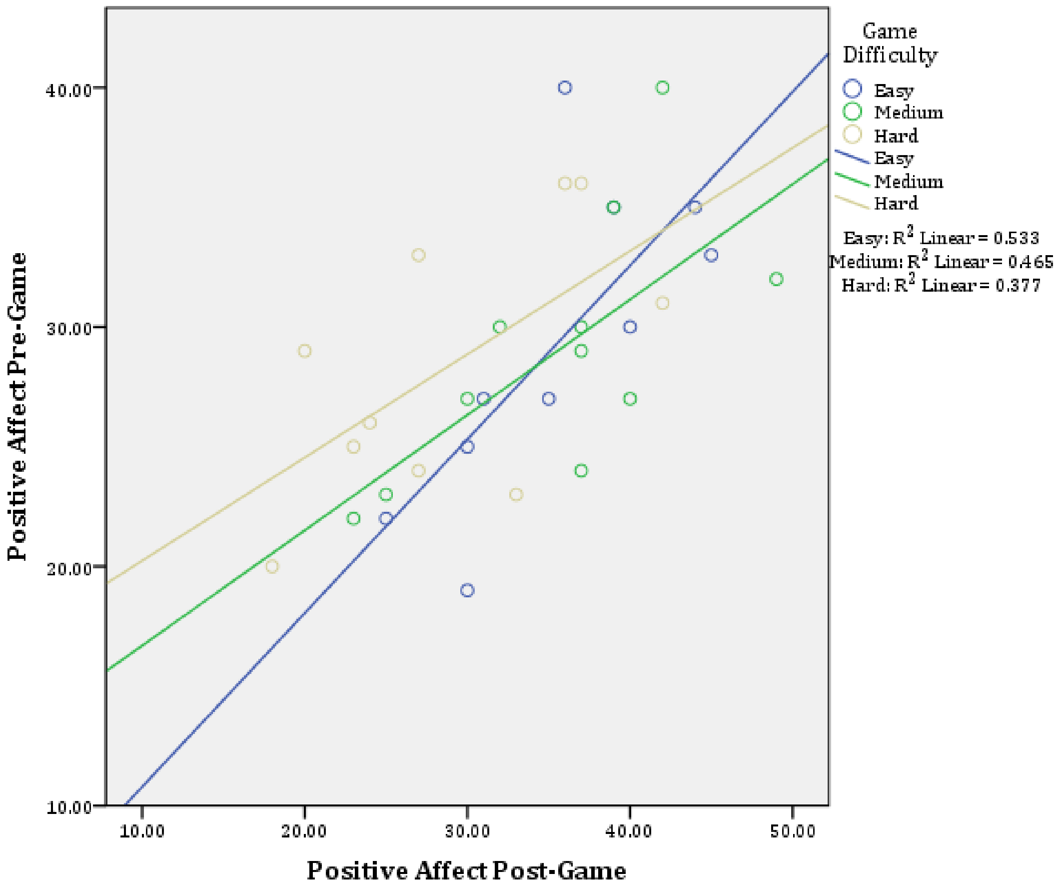


Figure 11: Scatterplot of positive affect shifts across the three difficulty game modes.

Further, the game's outcome (Victory/Defeat) was added as a between-subjects factor in a one-way ANCOVA with PANAS pre-game and post-game scores added as a covariate and a dependent variable respectively. Significant differences were observed when controlling for pre-game affect between the groups, $F(1, 28) = 5.57$, $p = .026$, $\eta_p^2 = .166$; players who won in the game reported higher positive affect ($M = 35.71$, $SD = 1.44$) than players who lost ($M = 30.8$, $SD = 1.49$) (**Figure 12**).

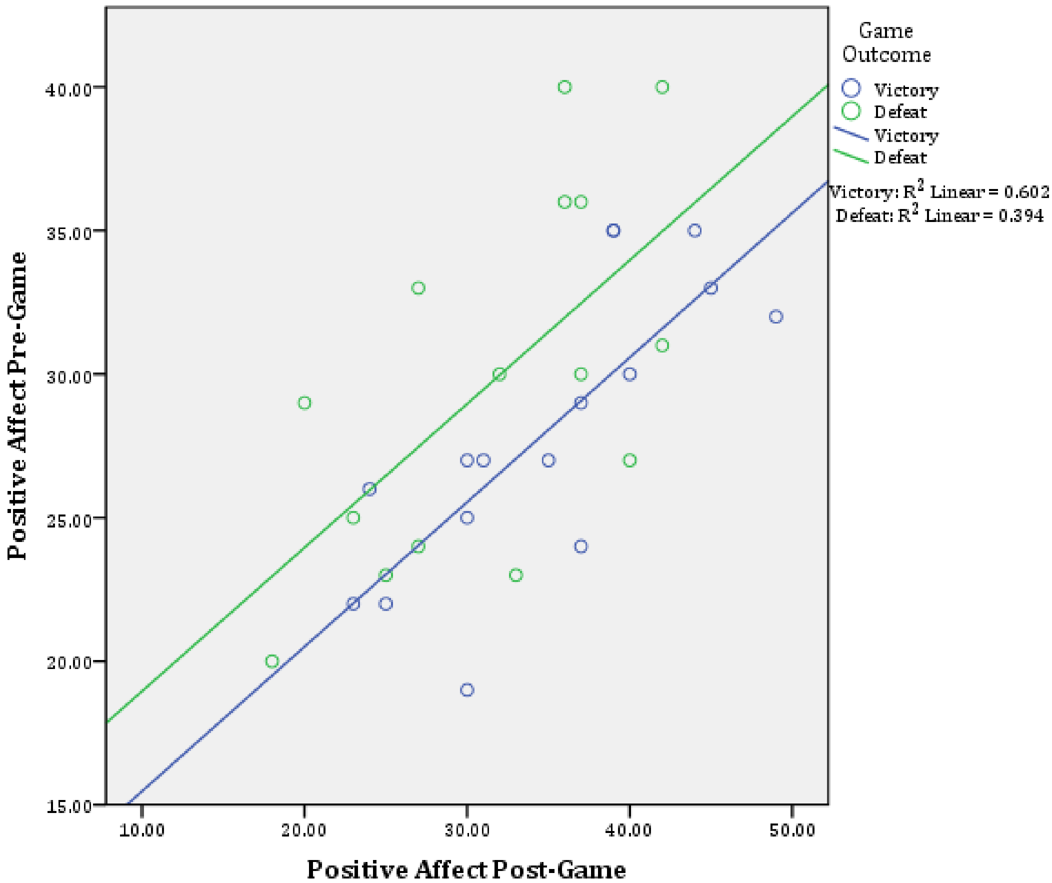


Figure 12: Scatterplot of positive affective shifts that occurred between the two different game outcomes.

4.3.3 Personality scores (BFI)

To test the relationship between personality traits and the reported flow experience (H3), a set of stepwise multiple linear regressions were run. The dependent variables were flow's individual dimensions. All analyses were examined for multicollinearity; the lowest tolerance score found from all the analyses was .745 and the highest Variance Inflation Factor (VIF) score was 1.343, suggesting that there were no multicollinearity issues (Hair et al., 2011). Except for feedback clarity, sense of control and concentration, the remaining flow dimensions were predicted by the BFI scores (**Table 4**). A final, multiple hierarchical regression was conducted on the overall flow score. As expected, the model revealed Agreeableness as the most significant predictor of total flow ($p < .05$), explaining 20.1% of the flow scores' variance. With respect to positive affect and negative affect, a stepwise multiple linear regression was run, using the personality traits as independent variables and the positive or negative affect pre-game scores as the dependent variables. The results showed that positive affect was predicted by agreeableness ($\beta = 4.582$, $p < .05$, $R^2 = .156$) and negative affect by neuroticism ($\beta = 2.36$, $p < .05$, $R^2 = .191$).

To address the influence of personality scores in the flow experience across the difficulty conditions, a one-way MANCOVA was conducted with the model built from interaction terms of individual personality traits (covariates) and the game difficulty mode (independent variable) on flow dimensions (dependent variables). Interaction terms were considered in this case because the difficulty mode may have had different effects for different levels of personality traits (e.g., Beck & Bliwise, 2014). The results showed that when controlling for Openness, Extraversion or Neuroticism, a significant interaction emerged for the flow dimension of Automaticity, $F(3, 12) = 5.986$, $p = .01$, $F(3, 12) = 9.605$, $p < .005$, and $F(3, 12) = 3.756$, $p < .05$ for each personality trait respectively.

When controlling for Agreeableness, a significant interaction occurred for the flow dimensions of Skills-Challenge Balance, $F(3, 12) = 5.174$, $p < .05$ and Clarity of Goals, $F(3, 12) = 4.848$, $p < .05$. No other interaction was significant for the remaining flow dimensions. The adjustment of the mean flow dimensions from the collective set of all personality traits revealed a significant main effect of game difficulty mode on Clarity of Goals, $F(2, 30) = 4.418$, $p < .05$. Follow-up pairwise comparisons revealed that significant differences were found between the easy and medium difficulty modes – goal clarity was highest in the medium difficulty mode.

Table 4: Summary of multiple linear regressions using personality traits (columns) as independent variables and individual flow dimensions (rows) as dependent variables. Rows containing more than one entry for personality traits indicate that a significant model was selected, containing all the displayed traits.

<table>
<tr><th></th><th>Agreeableness</th><th>Extraversion</th><th>Neuroticism</th><th>Openness</th></tr>
<tr><td rowspan="2">Automaticity</td><td>β = .854, p < .01</td><td>β = .37, p < .05</td><td></td><td></td></tr>
<tr><td colspan="2">Extraversion and agreeableness explained 34.4% of the total variance in automaticity; p < .01</td><td></td><td></td></tr>
<tr><td rowspan="2">Autotelic experience</td><td></td><td></td><td>β = .333, p < .05</td><td></td></tr>
<tr><td></td><td></td><td>Neuroticism explained 15.6% of the total variance in autotelic experience.</td><td></td></tr>
<tr><td rowspan="2">Clarity of goals</td><td>β = .618, p = .012</td><td>β = .333, p < .05</td><td></td><td></td></tr>
<tr><td colspan="2">Extraversion and agreeableness explained 28% of the total variance in clarity of goals; p = .012</td><td></td><td></td></tr>
<tr><td rowspan="2">Loss of self-reflection</td><td>β = .811, p = .011</td><td></td><td></td><td></td></tr>
<tr><td>Agreeableness explained 21.1% of the total variance in loss of self-reflection.</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Skills-Challenge balance</td><td>β = .741, p < .01</td><td></td><td></td><td></td></tr>
<tr><td>Agreeableness explained 27.8% of the total variance in skills-challenge balance.</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Time perception distortion</td><td></td><td></td><td></td><td>β = .817, p < .01</td></tr>
<tr><td></td><td></td><td></td><td>Openness explained 26.8% of the total variance in time transformation.</td></tr>
</table>

4.4 Discussion

In this study, personality as well as positive and negative affect were examined in relation to flow. The results corroborate previous findings suggesting a relationship between personality and flow, which may help researchers to better understand individual differences in experiencing flow. Agreeableness appeared consistently as a predictor of flow as well as its individual dimensions (see Ullén et al., 2016). According to Johnson and colleagues (2012), agreeable individuals are more likely to feel enjoyment during video game play and to be resistant to annoyance. Agreeableness has also been associated with happiness and well-being (Johnson et al., 2012; McCrae & Costa, 1991), both being relevant concepts to the flow experience. Extraversion and openness were also significant predictors of flow's dimensions.

Openness and Extraversion. Individuals high in openness tend to form connections and identify with their characters (Dunn & Guadagno, 2012; Soutter & Hitchens, 2016). Though the players did not actively control any avatar in the game, Ermi and Mäyrä (2005) suggested that character identification is a facet of emotional investment in the game and part of enjoying the fantasy unveiled within. Thus, openness may have related to the fantasy content featured in the TD game used in this study as well as the tendency of being strongly motivated (McCrae & Costa, 1997). Similarly, extraversion has also been shown to have ties to character identification (Soutter & Hitchens, 2016) and by extension the imaginative aspect of the game (Ermi and Mäyrä, 2005). Extraverts engage in sensation-seeking which contributes to enjoyment during video game play (McCrae & Costa, 1985; Fang & Zhao, 2009). Additionally, extraverts derive pleasure from achievement and mastery of challenges (Watson & Clark, 1997), which constitute intrinsic motivators of digital game play (Wang et al., 2008). Importantly, the combination of agreeableness and extraversion as predictors of flow dimensions confirms previous work in the area (e.g., Ross & Keiser, 2014).

Conscientiousness. Earlier research has also identified the role of conscientiousness in flow propensity (e.g., Ullén et al., 2012; 2016), which was not confirmed in this study. However, Johnson and colleagues (2012) identified an inverse relationship between conscientiousness and flow. The authors suggested that conscientious individuals may be less likely to "let go" and may not allow themselves to be fully immersed in the game, as a result of being highly goal-driven. Conscientious individuals display high levels of self-discipline and are more likely to engage in activities of personal growth than leisure (McElroy et al., 2007; Moore & McElroy, 2012). For example, conscientiousness has been shown to relate to flow in the context of job performance (e.g., Demerouti, 2006). However, contrary to Johnson et al. (2012), Teng (2008) has shown the relevance of conscientiousness in the context of video game playing.

Insights about the contradictory findings can be seen in Peever and colleagues (2012)'s work, who demonstrated that conscientiousness was significantly related to preference toward specific genres, i.e., sport, racing, simulation and fighting games. Perhaps, Tower Defense games do not conform with the mechanics found in those genres that would make them attractive to conscientious individuals. However, conscientiousness has been linked to thorough planning (Golbeck et al., 2011), which is reminiscent of the strategy game genre due to the high involvement of planning involved in the gameplay.

Indeed, strategy games may be preferred by conscientious individuals (e.g., Gabbiadini & Greitemeyer, 2017). Tower Defense games also incorporate elements that require strategic planning (Brich et al., 2015) and are considered a subgenre of strategy games (Avery et al., 2011). Thus, they could likewise be appealing to players with increased levels of conscientiousness. In the game used in this study, selecting the timing and an appropriate tower type for construction, the zonal grid to build a tower on as well as the timing for its upgrade played a critical role in the progression of the game. Paradoxically, however, conscientiousness did not appear to predict flow when playing a Tower Defense game. Perhaps, more research is needed to explicate the absence of this relationship.

The relationship between autotelicity and neuroticism. Another unexpected finding was an observed positive relationship between autotelicity and neuroticism. Previous research has consistently shown that neuroticism is antithetical to flow (Johnson et al., 2014; Tatalović Vorkapić & Gović, 2016; Ullén et al., 2012, 2016). This is because neuroticism has been traditionally conceived as a personality trait that is accompanied by elevated levels of negative affect (e.g., Giluk, 2009; McNiel & Fleeson, 2006; Smits et al., 1995). Indeed, this was confirmed in the present study, where a positive relationship was found between negative affect and neuroticism. Individuals with high neuroticism present a strong tendency towards addiction (e.g., Cao & Su, 2007; Hardie & Tee, 2007; Li et al., 2006; Mehroof & Griffiths, 2010).

This is also true for the flow experience, which has been shown to circumstantially devolve into addiction (e.g., Chou & Ting, 2003; Csikszentmihalyi, 1990), perhaps as a result of personal characteristics (Blaszczynski, 2008). Hence, it is likely that autotelicity, as a dimension of flow, could be used as a surrogate measure for addictive tendencies in individuals with high neuroticism. A related observation may be drawn from Conway and Rubin (1991)'s work, who found that individuals experiencing high levels of anxiety are more likely to engage in media as a means of escaping reality. These accounts may offer an interesting angle in the domain of video games addiction and the flow experience. However, due to lacking research in the manifestation of a positive relationship between flow and neuroticism, these explanations can only be speculative.

Flow as a result of game difficulty. Critically, the main hypothesis that flow dimensions vary as a function of game difficulty was not met. It was only after adjusting for personality that a flow dimension reached significance. Though the allocation of participants to each difficulty group was randomized, the predefined difficulty settings may have posed a limitation to this study. The overall flow score reported by the participants revealed visible trends, yet this was not confirmed in the statistical results. There are many possible explanations for the absence of evidence. First, the sample size was admittedly small, with roughly 10 participants allotted to each difficulty mode. Second, Tower Defense games increase in difficulty at each round, and each wave of enemies becomes more resistant than the previous. This automatically inherits a model of incremental difficulty, which conforms with an adaptive difficulty mechanism. However, the scaling of the enemies' health over time may have been rather drastic and may have not differentiated the three difficulty modes adequately. Klarkowski and colleagues (2015) were similarly unable to find substantial differences between two conditions of differing difficulty levels. The authors suggested that the flow experience may still emerge regardless of a skills-challenge imbalance, but also acknowledged potential design flaws.

Moreover, in this study, flow ratings were accumulated after the game, and the participants may have relied on the experience of the most recent rounds (see Sabet et al., 2019), which were admittedly harder than their preceding rounds regardless of the difficulty mode. Finally, the player was able to modify the pace of the game by simply holding down a key, which was perhaps an unnatural mechanic for VR. It may have been possible that game acceleration took place during instances of boredom or anxiety, as a means of skipping predictable or unwanted outcomes. Likewise, game acceleration could have impinged on the flow experience, by altering the perception of time that is already transformed during flow. The potentially confounding effects of the game acceleration mechanic were not originally accounted for and no supplementary measures were taken to record the frequency of its use.

In line with previous research, and as expected, the results suggested that winning contributes significantly to positive affect (Rieger et al., 2014). Over 60% of the positive affect shifts were explained by victory. Positive affect was best predicted by the personality trait of agreeableness, which is in accordance with earlier studies (e.g., Johnson et al., 2012). Notably, a similar relationship was not found for negative affect, suggesting that an in-game defeat did not significantly contribute to increased negative affect. This may indicate that players were not particularly interested in winning or losing, but that a victory was a pleasant outcome, whereas a defeat caused indifference rather than frustration. Victorious and defeated players reported on average comparable overall flow scores – 34.7 and 33.03 respectively out of a maximum score of 45.

Overall, the study revealed an interesting set of findings, which are mainly in accordance with other studies investigating the relationship of similar constructs, such as personality. However, flow should be more finely distinguished across the difficulty conditions if it is to be detected by means of physiological changes. Participants were only able to play one level of predetermined difficulty, which had had uncertain impact on player experience. This may be treated as a limitation in the game's design. Difficulty scaling, imposed by design, is perhaps insufficient without extensive testing in large player bases, such as beta or alpha testing as traditionally done in the commercial sector. Hence, improvements were made that are described in the next study.

5 Study II: Decoding flow from heart-rate and eye-blink variability using personality scores

5.1 Introduction

In the previous study, we showed that personality was a relevant concept to the flow experience in video games. However, there was a potentially deleterious effect originating from the way difficulty was determined (see sections Chapter 4, **Materials: Test Game** and Chapter 4, **Flow scores (FSS-2)**), which hindered our interpretability of the outcomes. In this study, we applied a dynamic difficulty adjustment formula in order to allow for the game to adapt to the player's performance. However, to investigate the effects of the adaptation, a self-report may not suffice, considering it is collected after the game and thus a flat score may not be truly representative of specific time periods during game play. Physiological measures may bridge this gap and provide valuable insights from patterns that can emerge as a function of the game design and difficulty adaptation. This is motivated by earlier research, where physiological measures have been employed as a means of assessing flow episodes during game play (e.g., Bian et al., 2016). Essentially, the purpose is to match the self-reported experience of flow with bodily changes that occur during play.

5.1.1 Electrocardiography

The most common physiological measures to have been employed in the study of flow are cardiac features (Knierim et al., 2018). The peripheral nervous system includes the autonomic nervous system (ANS), which further bifurcates into the sympathetic and parasympathetic branches. The functions of ANS are mainly regulatory and involved in metabolic and homeostatic processes (Loewy & Spyer, 1990). The sympathetic subdivision is widely known to be associated with the fight-or-flight response (McCorry, 2007), which is a stress response (Hamill & Shapiro, 2004) and otherwise known as the ergotropic response (Benson et al., 1974). On the other hand, the parasympathetic branch balances sympathetic activity and produces relaxation, i.e., the trophotropic response (Benson et al., 1974; Moss, 2004; Suter et al., 2009).

Heart rate variability (HRV) is a conglomerate of statistical measures that aim to identify the differences between successive heartbeats as a function of time or frequency (McCraty & Shaffer, 2015). Based on the distinction between the sympathetic and parasympathetic branches of the autonomic nervous system, the former being excitatory and the latter inhibitory (Appelhans & Luecken, 2006), a few hypotheses can be formed in relation to the Four Channel Flow Model (Massimini & Carli, 1986). Anxiety and frustration, as high-arousal states, should engage higher sympathetic activity (Giakoumis et al., 2010). On the other hand, flow's relation to a balanced mental workload is suggestive of parasympathetic modulation (e.g., Peifer et al., 2014; Tian et al., 2017). Indeed, Ullén and colleagues (2010) stressed that the regulation of the two branches may be an important marker of the flow experience. Finally, boredom should relate to decreased sympathetic activity, as a low-arousal state (Tozman et al., 2015). Given that the sympathetic and parasympathetic branches may exhibit antagonistic effects (Appelhans & Luecken, 2006), boredom should also relate to increased parasympathetic activity. The relationship between flow, boredom and anxiety, with respect to sympathetic arousal, appears to be quadratic, where flow hovers at the peak of an inverted-U shape; low arousal indicates boredom and high arousal indicates anxiety (Peifer, 2012; Peifer et al., 2014).

Heart rate appears to increase proportionally to the game's difficulty (Chanel et al., 2008; De Sampaio Barros et al., 2018) and may reflect mental workload (Jorna, 1993; Meshkati, 1988). Indeed, increases in heart rate during the flow experience have been consistently reported (Bian et al., 2016; De Manzano et al., 2010; Gaggioli et al., 2013). On the frequency domain, HRV is typically bisected into the low-frequency and the high-frequency bands (McCraty & Shaffer, 2015). Harmat and colleagues (2015) used a computer game in their study and asked participants to play three different levels of difficulty, easy, optimal and hard. The authors found that the optimal difficulty was characterized by the highest flow scores on a self-report scale and lower low-frequency HRV. In presence of mental load, HRV is expected to decrease (Dong, 2016). However, in the context of flow, this has not been consistently confirmed (cf.

Bian et al., 2016; Keller et al., 2011a). The low-frequency component of HRV remains controversial with respect to the underlying physiological mechanisms that it represents – e.g., baroreflex activity, sympathetic/parasympathetic interaction or parasympathetic tone (e.g., Goldstein et al., 2011; Houle & Billman, 1999; Reyes del Paso, 2013). Nevertheless, previous findings typically converge in the relatedness of cardiac measures as potential indicators of the flow experience, mental load (Keller et al., 2011a), stress in VR games (Ishaque et al., 2020) and automaticity (Peifer, 2012).

5.1.2 Electromyography

Another measure often used in flow research is electromyography. Electromyography is a method able to differentiate states of varying levels of arousal and valence (Cacioppo et al., 1986). In an influential study by Ekman and colleagues (1983), the authors were able to distinguish different emotions using electromyography. Though emotional expressions have long been interpreted as messages occurring mainly in social contexts (e.g., Ekman, 1970; Kraut & Johnston, 1979), several studies have shown that they may also form during video game playing. For example, contraction of the muscles zygomaticus major and orbicularis oculi has been linked to the experience of flow (e.g., De Manzano et al., 2010; Nacke & Lindley, 2010; Ullén et al., 2010), indicating positive emotions (Ravaja et al., 2006, 2008). On the other hand, the corrugator supercilii are highly relevant to negative affect (e.g., Larsen et al., 2003) and have likewise been associated with negative events in video game playing (e.g., losing to an opponent; Hazlett, 2006).

It should be noted that, in the study of Hazlett (2006), there was an observable overlap of the corrugator muscle group with positive events (e.g., winning) besides negative ones. Van den Hoogen and colleagues (2012) reported a similar finding where negative events (in-game deaths) potentiated activity in the zygomaticus major. These opposite trends, consistent with Ravaja et al. (2006), Salminen and Ravaja (2007), and Kivikangas and Ravaja (2013) findings, may suggest that a negative event does not necessitate a negative experience (Ravaja et al., 2006). Thus, they solidify the possibility that flow's suspension may not occur under negative affect as Jennett and colleagues (2008) suggested. In addition, it might be helpful to consider that flow may not have overt ties to a particular direction in the valence spectrum (e.g., Csikszentmihalyi, 1990, 1996). This may also explain the absence of a relationship between flow, zygomaticus major and orbicularis oculi reported in a study conducted by Kivikangas (2006).

In summary, electromyography is a promising measure, but robust evidence for flow is currently limited. Furthermore, we previously argued that the relatedness of flow to valance may not be exclusive to either positive or negative, and thus any inferred positive or negative states from facial muscle activation may not necessarily reflect the presence or absence of flow.

5.1.3 Electrodermal Activity

Electrodermal activity (EDA), not specifically pertaining to the facial musculature, has also been used as a marker of the flow experience. Unlike electromyography, EDA has been notoriously inadequate at detecting valence, but reliable for detecting arousal (e.g., Horslen & Carpenter, 2011; Kron et al., 2015; Öhman et al., 1993). However, recent studies employing techniques from the field of machine learning have claimed to have detected valence with a high accuracy (e.g., Canento et al., 2011; Feng et al., 2018; Greco et al., 2016). Electrodermal activity has been found to increase during challenging tasks (Frijda, 1986). Though Drachen and colleagues (2010) failed to replicate this finding, suggesting that there may have been opposing definitions of challenge in the self-report measure that they used, they found EDA to have a significant, positive relationship with frustration.

However, the same positive relationship has been found for flow (Nacke & Lindley, 2010; Peifer, 2012). Again, we are confronted with the dichotomy between flow and frustration. Although it was stressed earlier that their similarities may have previously been neglected, thereby rationalizing the inconsistency of findings reported in the literature, an important observation can be drawn from the study of Kron and colleagues (2015). The authors used a multidimensional set of features for EDA, in order to discriminate valence indices, which is known to provide comparably better classification results than a small set of features (e.g., Janecek et al., 2008). Hence, there might be more EDA parameters needed to disentangle flow and frustration as two separate high-arousal states.

5.1.4 Electrooculography

A final method of considerable implications for flow is electrooculography (EOG), which measures electric potentials evoked by oculomotor activity (e.g., Gratton, 1998). Although this approach has surfaced only recently in the study of flow, valuable insights may be drawn from its application in neighboring contexts. Typical parameters of interest from ocular activity include eye blinks, saccades, fixations and pupil dilation (e.g., Pfleging et al., 2016). It was previously thought that eye blinks were only helpful in lubricating the cornea (Stern et al., 1994), but studies have established an additional relationship with increased cognitive processing (e.g., Rosenfield et al., 2015; Veltman & Gaillard, 1996).

Eye blinks are thought to reflect endogenous processes, such as mental workload (Chen et al., 2011) and concentration (Wu et al., 2011), but they are highly variable among different individuals (Martins & Carvalho, 2015; Nourbakhsh et al., 2013). Thus, baseline measurements are recommended for interpretable outcomes (Egloff & Schmukle, 2002). A general rule is that mental workload results from the interplay of mental processes that occur whilst one is faced with a difficult task (Mulder, 1979). These processes may be manifested through significant physiological changes (Mulder, 1979). Similarly, low gazing dispersion (Reimer, 2009), duration of fixations (De Greef et al., 2009), duration of blinks (see Kramer, 1990), greater pupil dilation (De Greef et al., 2009; Palinko et al., 2010) and fewer saccades (Harbluk et al., 2002) have all been correlated with higher levels of cognitive load.

According to Recarte and colleagues (2008), mental workload is not a meaningful concept unless both the user and the task are taken into consideration. Mental workload has been shown to produce higher inter-blink rate and lower duration of the eye blinks as a result of high visual demand (Veltman & Gaillard, 1998). However, in Recarte et al.'s study, visual load and mental load were found to have opposite effects – eye blink rate increased during mental load but decreased during visual load (Recarte et al., 2008). This is consistent with previous studies, where decreased blink rates were observed under high visual demand (e.g., Charles & Nixon, 2017; Chen et al., 2011; Van Orden et al., 2001). The current evidence is therefore mixed, especially when considering studies reporting opposite findings, i.e., blinking decline during high mental load (e.g., Bednarik et al., 2018; Mallick et al., 2016; Wanyan et al., 2018;

Wu et al., 2011, 2014; Zheng et al., 2012). This issue between visual load and mental load was recently acknowledged, suggesting that previous research may have misinterpreted a visually demanding task as being mentally demanding (Marquart et al., 2015).

A mechanistic view of humans as systems with limited capacity acknowledges the limitations imposed by working memory capacity (Mulder, 1979; Sweller et al., 2011). The Cognitive Load Theory suggests that when this capacity is overflown, it leads to what is understood as cognitive overload and a difficulty in or a possible suspension of information processing (Sweller et al., 2011). Under these conditions, differential blinking patterns may occur, as well as performance deficiency (Bernadine et al., 2013). Chen and Epps (2014) suggested that blinks are better indicators of perceptual than cognitive load, and blink rates may increase as a means of alleviating mental tension. However, in the specific case of the flow experience, cognitive load is expected to be moderately high yet attuned to the individual's skills (Pilke, 2004). Chang and colleagues (2017) found that flow is related to *germane cognitive load*, which is a concept derived from the Cognitive Load Theory and describes the effort required for the construction of cognitive schemata that relate to learning (Sweller et al., 1998). This finding is reminiscent of the automaticity component of the flow experience (e.g., Csikszentmihalyi, 1990).

Rau and colleagues (2017) found that participants who reported a high flow experience during video game play, demonstrated significant changes in blinking rates throughout the game at different times, with a markedly significant decrease relative to a baseline. Notably, the game that they used, Warcraft III, involves a visually demanding activity, which is argued to be generating lower blink rates (e.g., Marquart et al., 2015; Recarte et al., 2008). Nevertheless, the resulting connection between visual load and self-reported flow scores is interesting. In a similar vein, but not contextualized to the flow experience, Wu and colleagues (2014) employed a video game and monitored ocular activity during playing (high-concentration task). In comparison to listening to music, which was deemed a low-concentration task, the authors found significant differences between blinking rates, with observable decreases during the high-concentration task. However, a small concern with this study is that the trial times were relatively low (150 sec), which may not be adequate in capturing significant changes in blinking behavior.

Overall, the peripheral nervous system seems to be highly responsive to variations in task difficulty, manifested through fluctuations in arousal levels. Hence, task difficulty alterations may facilitate our understanding of flow's onset. Although each individual measure could, to some extent, differentiate flow from other mental states, a multimodal measurement is perhaps more desirable (e.g., Bian et al., 2016).

5.1.5 Study objectives

In this study, both electrocardiography (ECG) and electrooculography (EOG) were used to investigate the flow experience by modifying the difficulty of TD-VR. These measures were chosen as indices of arousal and mental load that earlier studies have demonstrated (e.g., Bednarik et al., 2018; Martins & Carvalho, 2015; McIntire et al., 2014; Recarte et al., 2008). In addition, EOG appears to be a particularly interesting measure for the flow experience (e.g., Rau et al., 2017) with only a pair of electrodes needed to extract the physiological responses of interest. Research has so far applied EOG in virtual reality mainly for enabling interaction with a game as an alternative physical input (e.g., Kurma & Sharma, 2016). However, there is little evidence on the use of EOG as a surrogate measure for flow in virtual reality (e.g., Ishiguro et al., 2019).

The purpose of this study is to investigate the physiological patterns from both ECG and EOG modalities during virtual reality game play. This will also enable us to discern whether these temporal patterns contain sufficient information to decode flow, anxiety and boredom. We expect that when the difficulty adapts to the player's performance, the flow experience is more likely to trigger (e.g., Denisova & Cairns, 2015a). It is further hypothesized that anxiety will trigger the highest arousal levels, followed by flow and finally by boredom (e.g., Keller et al., 2011a; Mikulas & Vodanovich, 1993; Peifer et al., 2014).

Notably, what is meant by arousal is the physiological arousal under the effects of mental effort (e.g., Critchley et al., 2000). In video game studies, sympathetic arousal has been shown to increase (see Subahni et al., 2012). It is considered that elevated arousal levels can be indexed through heart rate variability metrics, such as higher heart rate (Kennedy & Scholey, 2000), and thus shorter inter-beat intervals (Gentil et al., 2009), as well as decreased pNN50 and RMSSD (e.g., Hammoud et al., 2019; see also **Electrocardiography Preprocessing and Features**). For EOG, increased arousal is expected to be manifested through decreases in spontaneous blink rates and blink duration in presence of visual load (Brouwer et al., 2014a). The opposite, an increase in eye blink frequency, may allude to higher distractibility (Strenge et al., 1999).

Those arousal differences provide a useful input for the classification of gameplay experiences, and thus offer the prospect of monitoring periods of experiential discontinuities in real-time settings. In addition, they may uncover latent patterns for validation of the original game design choices, which are documented in the next sections.

5.2 Methods

5.2.1 Participants

Twenty-nine volunteers took part in the study (mean age = 31.5, *SD* = 7.0) who were employees in Sony Interactive Entertainment in London, United Kingdom. The sample comprised 19 males and 10 females with an average video games experience of 19.5 years (*SD* = 8.7). As such, no gender or age group screening took place for this study. Those restrictions were lifted such that the samples from the present and the next study (**Study III - Chapter 6**), both employing biometric measurement, would generalize better to the diversity of the gamers' population. The study was approved by the Bournemouth University Research Ethics Committee (ref: 17333).

5.2.2 Data Acquisition

Electrocardiography (ECG) and electrooculography (EOG) recordings were obtained with BIOPAC BSL MP45 (BIOPAC Systems Inc., California, USA), using the proprietary software Student Lab 4.0, at a sampling rate of 1 kHz, which has been deemed an appropriate sampling rate for HRV time-domain measures (Hejjel & Roth, 2004).

A set of events for the start and end of the recordings from Student Lab 4.0 were added to the message list on the receiving desktop computer (for details, see Chapter 3, **Data acquisition**). The recordings were always started (BIOPAC_RECORDING_STARTED) and ended (BIOPAC_RECORDING_ENDED) using the key combination Ctrl + Space. The custom application on the desktop computer that listened to incoming messages from PlayStation used the native input detection to capture the aforementioned key combination when a window starting with the name "Biopac" was active. This was done to track the Coordinated Universal Time (UTC) timestamp the recording was started on Student Lab 4.0, thereby achieving synchronization with the game events during the data analysis.

For ECG, the placement of the electrodes followed Einthoven's triangle principle in a LEAD I configuration (e.g., Gargiulo et al., 2018). The ground electrode was attached on top of the right ankle, the positive electrode on the left wrist and the negative on the right wrist. For EOG, the positive electrode was attached on top of the left eye and the negative electrode below the left eye, both approximately at 1 cm + half the size of the disposable electrodes. Vertical EOG was preferred, as eye blinks, a feature of interest in the study, produce observable spikes (e.g., Heo et al., 2017; Ma et al., 2015) more prominently on the vertical axis (Oliveira et al., 2018; Pander et al., 2008).

Two Ag/AgCl disposable cloth electrodes were used (EL504, BIOPAC Systems Inc., California, USA) for EOG and three Ag/AgCl disposable pre-gelled electrodes (EL503, BIOPAC Systems Inc., California, USA) were used for ECG. The data were imported into MATLAB R2019b (The MathWorks Inc., Natick, USA) for trial segmentation and analysis. In addition to the physiological data, two self-reports were used, the Big Five Inventory (BFI; John & Srivastava, 1999) for measuring personality traits and the Long Flow State Scale - General (FSS-2; Jackson & Eklund, 2002) for measuring the flow experience.

5.2.3 Materials: Test Game

5.2.3.1 Scoring System

TD-VR received major updates from its earlier version. Unlike archetypal games from the Tower Defense genre and the first version of TD-VR, the lives system was eliminated in the following experiments. This was done to prevent variable data durations across participants that could have resulted from individual performance advantages and ensured that there were data from all the volunteers in similar sections of the game. Instead, at the end of each round, a pop-up panel was used to display the score for each round (**Figure 13**). This score was calculated based on the ratio of enemies eliminated over the enemies originally instantiated. The panel also served to alert the player about the type of enemies that would follow in the next round. At the end of the game, an overall score averaged from all the rounds was shown to the player with a maximum score of five. Hence, even though a player could neither win nor lose in the game, their sense of progression was not entirely deprived.

Figure 13: Scoring system at the end of each round in TD-VR. This was used as an alternative to a life pool system that commonly appears in Tower Defense games, to ensure that all players played the full game, despite individual performance advantages. The pop-up displays the performance of the player for that round as well as prepares them for the type of enemy units in the next round.

5.2.3.2 Environment and Gameplay Modifications

The environment and milieu (the visual genre of the game; Apperley, 2006) maintained its fantasy theme, but the environment was significantly enhanced in terms of visual detail. Visual details included animated fire sources, fog scattered across the game table, an animated windmill, a small lake and several structures aiming to depict a small village (**Figure 14**). The locations at which the player could build towers were also given a different visual appearance. The previously used white squares were replaced with hovering stone platforms, which allowed them to visually blend with the environment.

The original towers were also visually replaced by gems, whose color matched their tower type, as a means of associating color with tower functionality (**Figure 14**). Red gems were used for the Fire towers (incurred a burning effect, which caused the attacked enemy to receive a percentage of their maximum health as damage every second for three seconds – known as "Damage over Time (DOT)" effect). For clarity, given the naming consistency amongst the studies, Fire towers did not have a special ability in the previous study. Blue gems were used for the Arcane towers (briefly paused the attacked enemy, by locking them in place – commonly referred to in video games as "stun" effect) and purple gems for Nova towers (caused enemies surrounding the main target to also receive a percentage of the inflicted damage – known as "area of effect (AOE)" effect).

Due to an observed misuse of the Disease towers in the first study, those towers were replaced with the Arcane towers. The issue was that many participants constructed them in the beginning of the path, which did not allow them to function to their full potential. That was because Disease towers maximized their damage when the enemy's life was low. Hence, constructing them at the start of the enemies' route hindered their capability, because the enemies were freshly spawned. All towers in the present study were designed to work in a complementary way to one another – in game design terms this is typically referred to as *synergy* (e.g., Rocha et al., 2008). Players could also sell existing towers if they felt that their current layout did not perform well, and they would be compensated for 50% of the tower's overall cost in return (**Figure 15**).

Figure 14: Visual enhancement of the modified game table from Tower Defense VR (top) and close-up of the game table action in progress (bottom). The three different tower types are displayed.

Figure 15: Tower selection in TD-VR. This is an example of a level-1 Fire Tower whose special ability is to trigger a burning effect on the enemy that persists for three seconds. The red ring on the ground is used as an indicator of the tower's attack range. The black depicted gems activate based on the towers' level. The value next to the gold pot on the bottom indicates the cost to upgrade the tower.

5.2.3.3 Rounds and Dynamic Difficulty Adjustment

There were three round-types introduced, coded as flow, anxiety and boredom. The player was shown aliases instead: Extermination, which corresponded to flow, and Assault, which corresponded to either anxiety or boredom. The game comprised 30 rounds (trials) grouped into three consecutive blocks with an overall average duration of 28.77 min (SD = 0.35 min). Each block included rounds designed to induce flow, anxiety and boredom (ordered as mentioned). The experimental variables were the health of the enemy units as well as the currency reward magnitude bestowed whenever the player eliminated an enemy (see e.g., Sutoyo et al., 2015). Reward manipulation served to hinder or to facilitate game progression and by extension player performance, thus adding to the perceived difficulty of the game.

Flow-designed rounds featured an adaptive difficulty scaling mechanism that was inspired from earlier work in TD games (Sutoyo et al., 2015). The ratio of enemies escaping over enemies originally spawned was indexed through a predefined table of ratio ranges (e.g., {0.4–0.6}). The lower the ratio, the higher the difficulty of the next round and the lower the reward bonus. Conversely, the higher the ratio, the lower the difficulty of the succeeding round and the higher the reward bonus. On the other hand, anxiety and boredom round-types featured bosses, which were more powerful (**Figure 16**). Unbeknownst to the player, those enemies were programmatically modified to be invincible near the end of their life. When the bosses appeared, a random value was generated between 1 to 5%; when the health of the boss reached that percentage, further damage would no longer be inflicted.

Figure 16: Unit-types introduced in the second version of the TD-VR game. From left to right, (a) the first unit is a fast boss with high health and high reward bonus in the anxiety rounds. In the second rhombus (b), there are ten units in the flow-designed rounds, whose health and reward bonus increase proportionately to the player's performance in the previous rounds. In the last rhombus (c), a slow-moving boss with very high health and low reward bonus was introduced for the boredom rounds.

The bosses first appeared very early in the game (rounds 4 and 5), but they were made vulnerable and easily eliminated. This was done to prevent players from apprehending the experimental manipulation at later stages of the game. Anxiety rounds featured a fast-moving boss with high health, and a high reward promised, whereas boredom rounds featured a slow-moving boss with very high health, but low reward in return. This was done to increase motivation and goal-directed behavior for the anxiety rounds, but reduce it during the boredom rounds (e.g., Johnson et al., 2018; Mitchell, 1982). We therefore anticipated that anxiety would kick in when players were unable to obtain the promised reward as a result of high challenge, and boredom to result from the absence of an interesting incentive as well as the sense that their goals were blocked (Bench & Lench, 2013). Anxiety and boredom rounds were repeated three times in a row. The boredom rounds followed anxiety, because repetitive play of the same difficulty level is more likely to induce boredom (Chanel et al., 2011). Fleeing bosses did not precipitate difficulty adjustment for subsequent rounds – only flow rounds did. This is because the bosses would have erroneously signaled low player performance, which was an intended outcome.

In addition, bosses were made immune to the towers' effects (for example, they could not be slowed down by the Arcane tower) and were only susceptible to the towers' raw damage. This was done to minimize the risk of revealing the experimental manipulation. Near the end of the designated path, the enemies' speed was doubled to accelerate the end of each round. The point at which the acceleration occurred was outside the range of the highest attack-range tower. The acceleration was introduced to preserve some level of player autonomy from the previous study, in which the players used a key to manually accelerate the game. The new mechanic did not impact the gameplay, as it occurred automatically during a period of inactivity or when the player could not have prevented the enemies' escape.

5.2.4 Procedure

Participants were first asked to sign a consent form, informing them about the purpose of the study and their right to withdraw from the experiment at any point. The games experience questionnaire and BFI were first filled out, followed by a training stage to help familiarization with the controls of and the game. Once they reported confidence with handling the game, the electrodes were attached, and a resting baseline was recorded for one minute. Participants then played the full game and finally filled out the FSS-2 questionnaire. Including the setup and training, the experiment lasted approximately 60 minutes per participant. As a note, participants were not made aware that the game featured indomitable rounds and content that adapted to their performance before the experiment. This was done during the debriefing (i.e., the conclusion of the individual experimental session after all self-reports were filled out).

5.2.5 Experimental design

Contrary to the previous study (see Chapter 4, **Experimental design**), the design amendments of TD-VR in this present study called for a different experimental design. All participants were exposed to all the experimental conditions of flow, anxiety, and boredom. In this case, a within-subjects design was more appropriate. There were two reasons behind this decision. First, the recruitment of participants proved challenging in the previous study. This becomes an issue because between-subjects designs require a high number of participants that is proportional to the number of experimental conditions (Brysbaert, 2019).

Second, a within-subjects design was preferred for observing the evolution of the volunteers' physiological responses across the experimental conditions. In general, players may waver through multiple affective states whilst playing (e.g., Shin et al., 2012). Hence, we considered that the scenario of testing participants through different conditions in one gameplay session promoted higher external validity. Additionally, we needed to obtain a more granular measure of how personality interacts with the gameplay experience across different stages in the game to better understand how games can be improved.

The technical constraint of TD games, whereby difficulty increases at every round, precluded the randomization of the experimental conditions. Thus, any effects caused by the order of the conditions were not controlled. The game design of the boredom rounds relied on repetition. Consequently, shuffling the three game conditions would have risked obfuscating the reliability and literature support behind the states they were intended to induce (see **Rounds and Dynamic Difficulty Adjustment)**. Similarly, the dynamic difficulty adjustment mechanic would have been impractical if the anxiety and boredom conditions had occurred at random.

A known concern with within-subjects designs are the carry-over effects, which can be alleviated by introducing inter-condition breaks (Greenwald, 1976). In the current study, volunteers were given the autonomy to confirm when to progress to the next game round by pressing a button. After they confirmed, a fixed 5-second window followed before the next round commenced. These considerations were made to suppress carryover effects not only between-, but also within-conditions.

5.2.6 Statistical Analysis

The statistical analyses were carried out using SPSS v23.0 (IBM Corporation, Armonk, NY). Each game round was split into three windows. The analyses comprised the mean values extracted from the second window, which guaranteed that the player engaged in the different content offered by each condition. First, the input data were extracted from the second window of each game round, in order to capture the most salient differences of game play content offered by the different round-types (flow/anxiety/boredom). The first and third windows contained periods of player inactivity, where the player waited until the enemy units arrived at the outer tower's range of effect or until the enemy units exited, respectively.

Though these peripheral windows may have contained information relevant to the state of interest for each round, they were also likely confounded with player inactivity, essentially discounting a major part of the gameplay. The factors used in a 3x3 within-subjects design were the round-types, i.e., Condition, which refer to Flow, Anxiety and Boredom rounds. The bosses who first appeared in the 4th and 5th rounds (see **Rounds and Dynamic Difficulty Adjustment**) did not interrupt the game's pace, as they did later in the game, and were thus treated as flow trials. The second within-subjects factor, Block, was also added as a factor with three levels, i.e., Block 1, Block 2 and Block 3, each reflecting the repeated sequence of Flow-Anxiety-Boredom rounds for each block in the game. In other words, block can be viewed as three separate time points in the game.

Data normality was investigated using the Lilliefors normality test. Significance level was set to $\alpha = .05$ (two-tailed). For violations of sphericity (Mauchly's test), reported results are after the Greenhouse-Geisser adjustment for degrees of freedom. All post-hoc pairwise comparisons were carried out with Bonferroni adjustment. The data were first transformed into percentage changes from the baseline measurement, taken prior to the game session. This was calculated using the difference between the observation and the baseline value of the corresponding participant and then divided by the baseline value.

The assumption of the normality of residuals for repeated-measures ANOVA was consistently violated and the data were severely skewed, which is a common scenario in HRV analysis (e.g., Katz-Leurer & Shochina, 2005). A routinely used solution for HRV analysis is to transform the data to natural logarithms (e.g., Caldwell & Steffen, 2018; Katz-Leurer & Shochina, 2005; Saboul et al., 2016). However, the non-normal distributions persevered in our data (see also, Kekecs et al., 2016). Hence, the data were instead rank-transformed (rank 1 corresponding to the smallest value) to purge or remediate the skewness (e.g., Conover & Iman, 1981; Thurber et al., 2010). After rank transformation, the residuals were again tested for and passed normality of distribution using the Lilliefors test.

5.2.7 Data Segmentation

Each round was segmented into $n = \{1, 2, \ldots, 9\}$ epochs to examine whether epoching can minimize the probability of error (e.g., Tax et al., 2000). Generally, longer windows tend to boost classification performance, provided that the mean and variance vary slowly, at the expense of high computational times, as they contain a higher volume of information for determining the classes' separation decision (Da Cruz et al., 2015; Smith et al., 2010). On the other hand, shorter windows provide richer training sets, which can also improve classification performance (Darvishi et al., 2013).

Three approaches were considered for trial-data segmentation and subsequently the selection of an optimal windowing method: (*a*) fixed-*length* windows (the length is variable), (*b*) partition each trial into three segments (the number of windows is constant), and further segment the *second* window, thereby trimming the beginning and the tail of each trial, and (*c*) partitions dictated by *n* windows (the number of windows is variable).

(a) The first possible approach entails partitioning each game round (trial) into non-overlapping windows of a fixed length (fixed-length windows). The problem in this case is that the game's rounds have had variable durations, which depended on the player's success at eliminating the enemies. This causes the final data matrix to be evidently skewed in favor of flow and boredom rounds. On the other hand, anxiety rounds featured a fast-moving boss, and hence have had a lower average duration than the remaining types of rounds, making class definition highly imbalanced, in addition to the already imbalanced classes of the game's design (16 rounds dedicated to flow, 7 to anxiety and 7 to boredom). Fixed-length windowing was thus deemed impractical, as the resulting number of observations per class was dependent on the duration of each game round.

(b) The second approach is similar to the one employed for our statistical analysis (see previous section). The beginning of each trial, as well as the end of each trial contained periods of player inactivity. Hence, they may not be as informative as the periods of pure player activity which occurred in the second segment of each game round. This windowing approach will thus split each game trial into three segments and maintain only the second segment. Thus, we may use the second segment to further partition it and to enhance the training, training and validation sets.

However, considering that the duration of certain trials was as low as 19 seconds, this would entail a window length of 19/3 = 6.67 seconds. The implemented Pan & Tompkins (1985) algorithm for HRV computation (see **Electrocardiography Preprocessing and Features**) requires at least 2 × Sampling rate (1000) to calculate thresholds that are necessary to discriminate between noisy or correct peaks (Sedghamiz, 2014). What this means for the present dataset is that each window must not be less than 2000 ms, otherwise the IBI data, necessary for HRV calculation, will not be possible to retrieve. Therefore, we would be unable to partition the second window lasting for 6.67 to more than 3 windows, as 4+ windows would produce a window length of less than 2000 ms, making HRV computations unlikely.

(c) The third approach is to partition each game round (trial) into n non-overlapping windows. Thus, instead of extracting e.g., n × [2000-ms] windows from each round (fixed-length windowing), we extract [1…n] windows that are on average of equal length. Differences in window lengths within each trial will be due to rounding, with an offset of 1-2 samples added to the final segment. In this sense, the approach is similar to the fixed-length window approach described above, except that each round is guaranteed to produce the same number of windows per game round, regardless of its duration. Hence, the number of observations each game round will generate for training and validation sets will be the same. The class imbalance from the game's design will still be present, but this can be alleviated with resampling that we will describe below.

Based on the aforementioned considerations, we opted for the third approach, by segmenting each trial into n equally-spaced and non-overlapping windows. However, the issue of short-duration trials still lingers for HRV calculations. Hence, we limited the search range of optimal window counts to 1…9. This search range was also applied to the data segmentation of EOG to maintain consistency and comparability between HRV and EBV.

5.2.8 Feature Extraction

For eye blink variability, the first parameter extracted from each epoch was the mean inter-blink interval (bIBI; seconds), i.e., the number of seconds between two consecutive blinks. The remaining features were the minimum inter-blink interval (minbIBI; seconds), maximum inter-blink interval (maxbIBI; seconds), standard deviation of inter-blink interval (stdbIBI; seconds), eye blink rate (EBR; count per minute), mean blink duration (DUR) and mean amplitude of the blinks (AMP) (see **Table 1**). HRV metrics comprised seven time-domain measurements, calculated for each epoch: mean IBI (ms), minimum IBI (ms), maximum IBI (ms), SDNN (ms), mean HR (count/min), pNN50 (%) and RMSSD (ms) (see **Table 2**).

Features comprised {29 participants × 30 trials × n windows} samples. Each observation (row of features) was first subtracted from the resting baseline of the corresponding participant (Chanel et al., 2011; Vicente et al., 2016) as well as the pre-trial baseline in order to suppress physiological leftovers from the immediately preceding trial. The pre-trial baseline was a period during which the game expected the user's input to launch the next game round. The input data were standardized using z-score transformation, in order to control for personality scaling and inter-feature numerical dominance and provided the input to three competitive classifiers separately – Support Vector Machine (SVM), Ensemble of CART Decision Trees with Bagging (BDT) and k-Nearest Neighbors (kNN) (see Chapter 3, **Overview of classification algorithms**).

Ground truth was tested by **a**) splitting the sample into two parts based on the 50^{th} percentile (median), using the self-reported overall flow score from FSS-2 (the sum of all flow dimensions), and **b**) the labels governed by the game's design. Thus, both binary (low-flow/high-flow) and multiclass classification (flow/anxiety/boredom) methods were employed. The multiclass training data were oversampled during each fold cross-validation using SMOTE (see Chapter 3, **Balancing class**). Tests were carried out for each physiological measure separately and in conjunction, i.e., HRV (7 features), EBV (7 features), HRV and EBV (14 features). Generally, feature fusion is expected to provide better classification decision performance under categorical labels (Mangai et al., 2010). In this study, we employed serial feature level fusion, by concatenating the feature sets (Fu et al., 2008) of the selected biomarkers.

5.3 Results

5.3.1 Statistical results

5.3.1.1 Comparison with Study I

To estimate the success of the new design compared to its predecessor (Chapter **4**), including the adaptive difficulty mechanism, the flow ratings were compared between Studies I and II (two-way t-test). The test revealed significant differences in the self-reported flow dimensions of Skills-Challenge balance, $t(58) = 2.828$, $p < .01$, Automaticity, $t(58) = 2.168$, $p < .05$, Autotelicity, $t(58) = 3.228$, $p < .01$, and the overall flow score, $t(58) = 2.210$, $p < .05$. However, these dimensions were rated significantly higher in the first study compared to the present study.

The same analysis was also applied to demographic variables. The average age of the participants in this study (31.5 ± 7) was significantly higher than in the first study (27.4 ± 4.9; $t(58) = -2.64$, $p < .05$). The reported years of experience in playing video games was similarly significantly higher in the present sample, $t(57) = -2.01$, $p < .05$, with an average difference of four years. Finally, for the Big Five Inventory questionnaire, only Conscientiousness was found to be significantly higher in the present study, $t(57) = -3.36$, $p < .01$, with an average rating of 3.91 ± .55 versus 3.4 ± .61 in the previous study.

5.3.1.2 Cardiac measures – Heart rate variability (HRV)

A 3x3 repeated-measures analysis of variance with Condition (Flow, Anxiety, Boredom) and Block (Block 1, Block 2, Block 3) as within-subjects factors was conducted on each dependent variable (mean HR, mean IBI, pNN50 and RMSSD). The model was specified to output the main effects of Block and Condition, as well as the interaction terms of Block × Condition. The variables pNN50 and RMSSD did not yield a significant F-statistic ($p > .05$).

Inter-beat intervals (IBI) and Heart Rate (HR). For IBI, the main effect of Condition was significant, $F(1.59, 44.48) = 5.732$, $\eta_p^2 = .17$, $p = .01$. Pairwise comparisons with Bonferroni adjustment showed that the condition of Boredom was characterized by a significantly reduced inter-beat interval compared to Flow ($p = .014$) and Anxiety ($p = .037$). As expected, HR was also found to be significant for the main effect of Condition, $F(1.61, 45.1) = 5.617$, $\eta_p^2 = .198$, $p < .01$ (**Figure 17**). Heart rate was found to be statistically significantly higher, relative to the baseline, during Boredom than during Flow ($p = .029$) and Anxiety ($p = .031$).

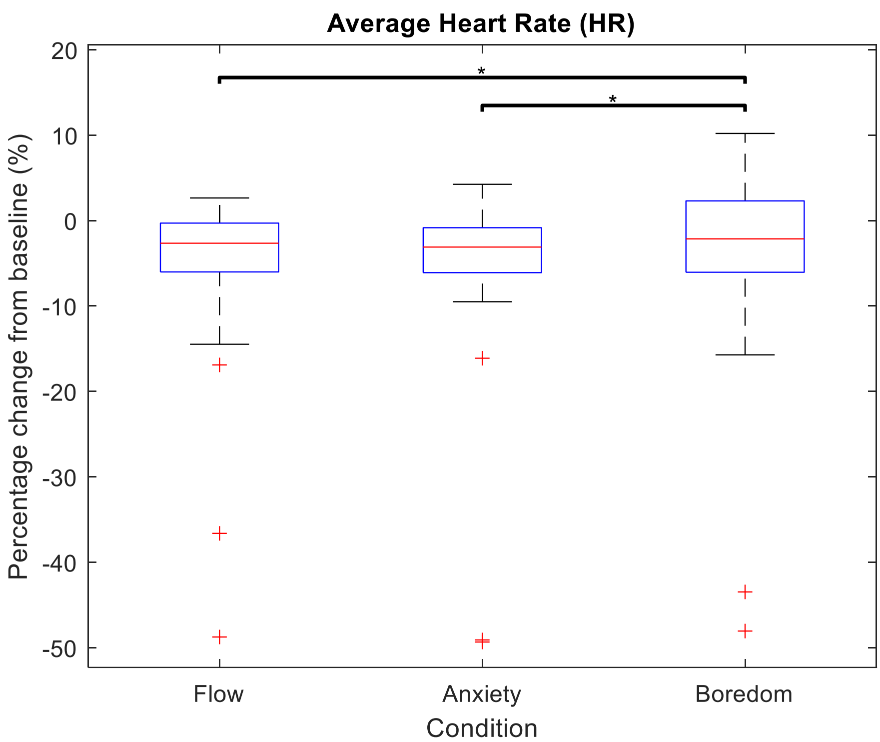


Figure 17: Mean heart rate across the three conditions of the TD-VR game. The omnibus repeated-measures ANOVA indicated a significant main effect of Condition. The post-hoc

comparisons showed that the difference in the average heart rate occurred between the pairs of Flow and Boredom and Anxiety and Boredom at $p < .05$.

5.3.1.3 Ocular measures – Eye blink variability (EBV)

The same test was carried out for the EBV variables (eye blink rate, blink duration and blink amplitude). The model was specified to output the main effects of Block and Condition, as well as the interaction terms of Block × Condition.

Eye blink rate (EBR). A main effect of Condition was found for eye blink rate, $F(2, 56) = 11.771$, $\eta_p^2 = .296$, $p < .001$. The post-hoc comparisons revealed that EBR significantly increased during Boredom as opposed to Flow ($p < .05$) and Anxiety ($p < .001$). However, flow and anxiety did not yield a significant difference, $p = .126$ (**Figure 18**).

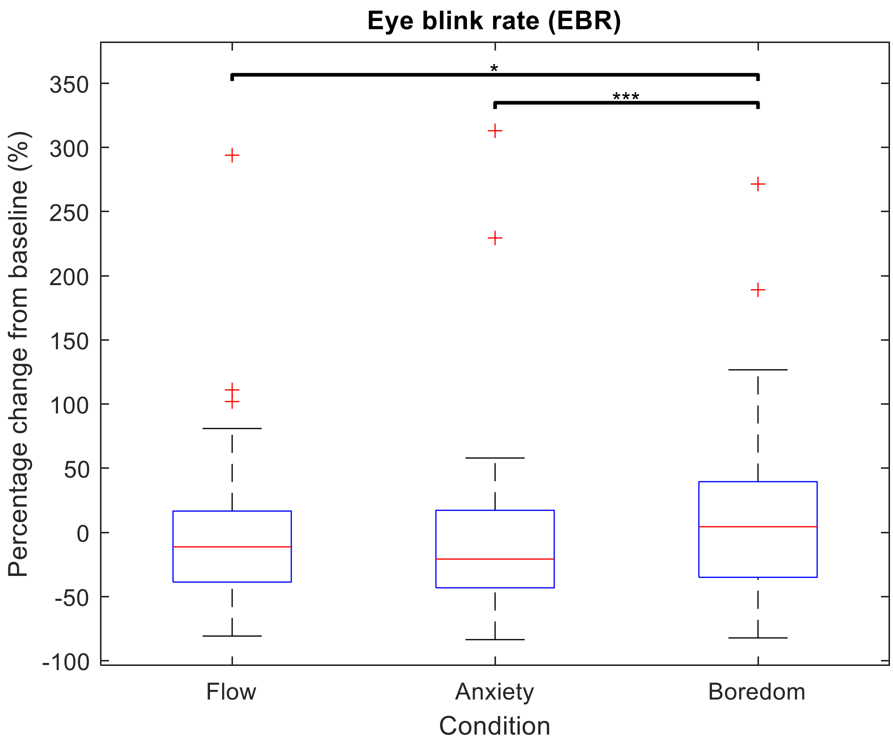


Figure 18: Mean blink rate across the three conditions of the TD-VR game. The omnibus repeated-measures ANOVA indicated a significant main effect of Condition. The post-hoc comparisons showed that the average blink rate was significantly higher during Boredom than during Flow and Anxiety at $p < .05$ and $p < .001$ respectively.

Blinks duration (DUR). The main effect of Condition was significant for the blinks' duration, $F(2, 56) = 6.413$, $\eta_p^2 = .186$, $p < .01$, whereby Boredom was characterized by significantly higher blink durations than Flow ($p < .05$) and Anxiety ($p < .05$). Relative to the baseline, the shortest blink durations were found during anxiety, followed by flow and finally boredom (**Figure 19**).

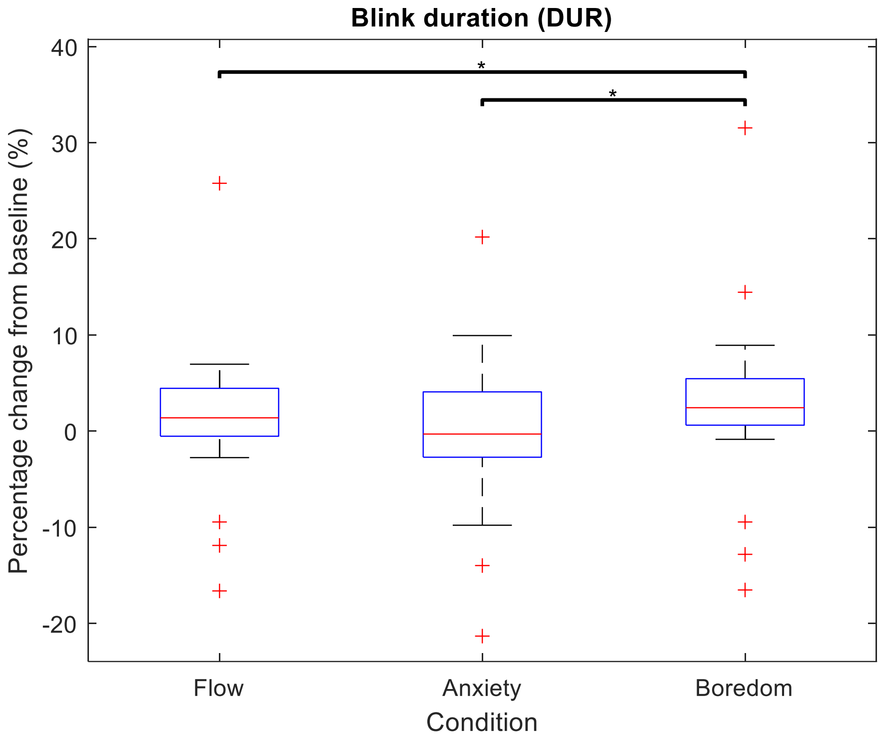


Figure 19: Mean blink duration across the three conditions of the TD-VR game. The omnibus repeated-measures ANOVA indicated a significant main effect of Condition. The post-hoc comparisons showed that the average blink duration was significantly higher during Boredom than during Flow and Anxiety at $p < .05$.

Blinks amplitude (AMP). The main effect of Condition for the blink amplitude was significant $F(2, 56) = 4.453$, $\eta_p^2 = .137$, $p = .016$. The post-hoc tests indicated that Anxiety has had a significantly higher blink AMP than Boredom, $p = .021$ (**Figure 20**). The remaining pairs as well as the main effect of Block were not significant.

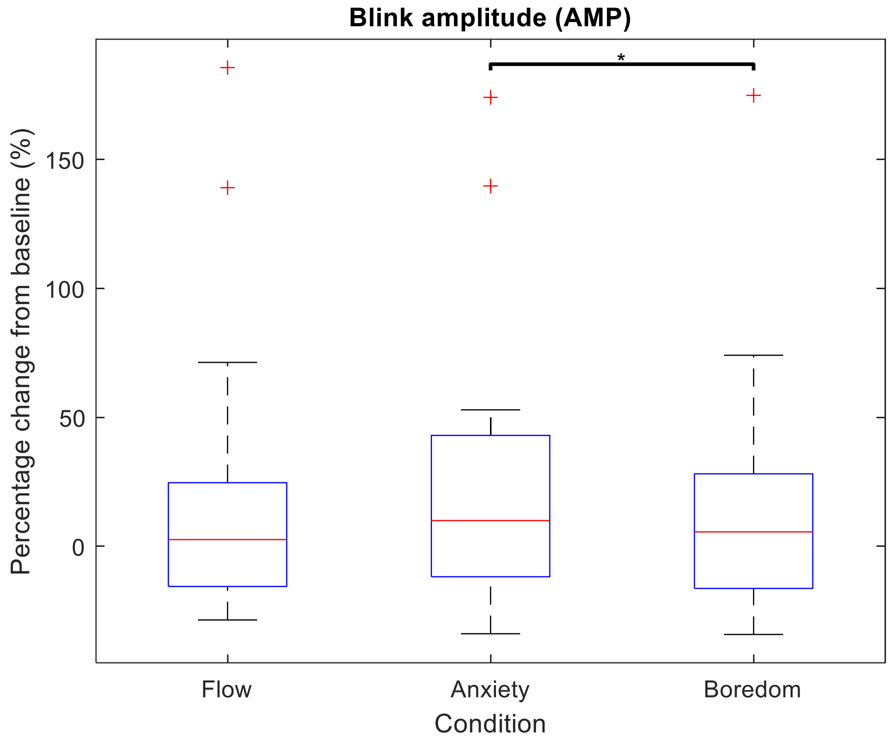


Figure 20: Mean blink amplitude across the three conditions of the TD-VR game. The omnibus repeated-measures ANOVA indicated a significant main effect of Condition. The post-hoc comparisons showed that the average blink amplitude was significantly higher during Anxiety than during Boredom at $p < .05$.

5.3.1.4 Discussion of statistical findings

The statistical findings of this study suggested that the physiological responses were consistent over time. This was evident from the absence of a significant main effect of Block and a Block × Condition interaction. Thus, the physiological responses were resistant to the effect of time and hence, the design conditions may be treated similarly across time.

Self-reports. Participants' ratings on skills-challenge balance were more inflated in the previous version of the game. Additionally, the previous version of TD-VR effected a more enjoyable experience overall, as evidenced by the dimension of autotelicity. Indeed, a perceived balance of challenge can produce higher satisfaction (Andrade et al., 2006). This is somewhat expected of a finding because the flow rounds were interspersed with the anxiety and boredom rounds that appeared to have been very challenging. As mentioned in the methodological section earlier, these rounds were impossible to win, albeit this was not communicated to the players until they were debriefed after the experiment.

In a similar vein, the self-reported flow dimension of automaticity was significantly higher in the first version of TD-VR. Dietrich (2003) posited that automaticity may relate to the level of expertise the individual has in a particular activity – individuals with higher expertise attend activities in a more automatic way than novices. A common measure of expertise in the video games literature is the number of years an individual has been playing video games (e.g., Latham et al., 2013). However, the trend in the comparison of our two studies is opposite – the participants of the present study have had a significantly higher number of years invested in video games, but their ratings on automaticity were lower.

First, this discrepancy of expertise may be due to the average age differences between the two samples; the participants of the second study have had a higher average age. Therefore, it is rather natural that this is similarly reflected in their proportionately higher number of years playing video games. A second possible explanation for the reduced ratings of automaticity in the present study is that the participants were bewildered by the anxiety and boredom rounds. According to Carson and Collins (2016), stressful situations may shift into conscious awareness which regresses into a state akin to newly learned skills, which dilutes the concept of automaticity. Due to the high challenge posed by those types of rounds in TD-VR, participants may have reviewed multiple strategies for defeating the bosses. Hence, the FSS-2 items pertaining to doing things automatically, without thinking, were consequently rated lower [5].

This section epitomizes the inadequacy of retrospective video game experience assessment (see Chapter 2, **The challenges in measuring flow in video games**). The holistic experience, reflected through the retrospective self-ratings, does not seem to provide enough information to evaluate moment-to-moment experiential fluctuations during the game. Our addition of a

[5] As a note, the exact questionnaire items from the Flow State Scale – 2 (FSS-2) cannot be referenced here for clarity, due to licensing restrictions.

dynamic difficulty adjustment mechanic in the second version of TD-VR cannot be reliably assessed from these findings alone. For this reason, they ought to be complemented with secondary measures. In the next section, we review our findings on the recorded physiological responses, which may constitute a richer source of information.

Physiology – HRV. Our results evinced consistent physiological differences between boredom and the remaining conditions. We noticed that heart rate was significantly higher during boredom, relative to the baseline, than during flow and anxiety. As heart rate is a well-known index of autonomic arousal (Jorna, 1992; Wang et al., 2018a), it is implied that boredom was characterized by higher arousal than the other two conditions. This finding is somewhat surprising, as boredom has been widely viewed as a low-arousal state (e.g., Mikulas & Vodanovich, 1993) even in animal studies (e.g., Burn, 2017). It is thus tempting to reevaluate boredom as a high-arousal state, which has also been proposed by Bench and Lench (2013). However, this may not be properly verified through the present study, as the block design was not randomized. The boredom rounds always followed the anxiety rounds at a repetitive sequence of six rounds in total (i.e., three rounds dedicated to the anxiety condition and another three to the boredom condition). Hence, an alternative explanation is that the boredom rounds may have contained carryover effects from the anxiety rounds. Yet, these conditions manifested significant physiological differences.

Design-wise, the differences between anxiety and boredom trials were the reward magnitude, the health and speed of the enemy boss. However, they were also conceptually similar – a single enemy unit that was seemingly impossible to eliminate; both types of trials were named "Extermination" and regarded as low-flow conditions (see **Rounds and Dynamic Difficulty Adjustment**). Given that the players were unable to defeat the bosses during those six rounds, this monotonous scenario may have provided scaffolding for increasing levels of frustration. In light of this assumption, Van Hooft and Van Hooff (2018) found that boredom may morph into frustration when the perceived autonomy is low.

This is highly relevant to the present study, as the players have indeed had low autonomy during the boredom rounds. Low autonomy was enabled by a difficulty to eliminate the bosses as well as the absence of fiscal rewards during the said rounds, which limited the opportunity for interaction with the game elements. Consistent with our findings, Van Hooft and Van Hooff (2018) associated high arousal of negative valence with boredom, stressing the moderating role of low autonomy. Thus, this observation does not invalidate the presence of boredom in our study, but rather suggests an amalgam of boredom and frustration.

Of note is also a study by Ralph and colleagues (2017) who found that task monotony may reduce vigilance. However, if the monotonous task is interrupted by a cognitively challenging task, vigilance may be restored. These observations can be similarly transposed to our context; given the nonrandomized trials, flow rounds may have acted as the cognitively challenging tasks that interrupted the monotonous sequence of the anxiety-boredom complex. However, anxiety rounds were also significantly different from the boredom rounds, attesting that anxiety rounds may also be viewed as cognitively challenging. This was likewise corroborated by the absence of differences among the flow and anxiety rounds. The reader should note that Keller et al. (2011a) were similarly unable to identify differences in the physiological responses of flow and anxiety.

Taken together, it becomes increasingly clear that flow, anxiety and boredom are intertwined in a triadic circumplex, wherein each state has a polymorphic activation: the activation conditions of each state may depend on subtle changes of perceived control. This circumplex can perhaps take the visual form of a Venn diagram. Yet, it makes it challenging to evaluate the appropriateness of the design for the boredom trials. However, a different perspective may be derived from the eye blink variability (EBV) measures in the next section.

Physiology – EBV. As mentioned earlier, the literature evidence on EBV remain ambiguous with respect to mental load. In this study, flow and anxiety were characterized by a decreased blink rate (EBR), relative to baseline, compared to boredom. This finding is corroboratory of earlier research demonstrating decreased blink rates during mentally taxing tasks (e.g., Bednarik et al., 2018; Mallick et al., 2016; Ohira, 1996; Wanyan et al., 2018; Wu et al., 2011, 2014; Zheng et al., 2012) and during the flow experience relative to the baseline (Rau et al., 2017). Reductions in blinking frequency have also been associated with visually or auditorily demanding tasks (Charles & Nixon, 2017; Chen et al., 2011; Recarte et al., 2008; Stern & Skelly, 1984; Van Orden et al., 2001).

This is putatively relative to information intake, whereby blink inhibition facilitates attention to stimuli (Fukuda et al., 2005). Our findings for the flow condition can thus be partially justified, since the number of visible enemies was ten times higher than the boredom condition and thus more visually demanding. However, the same explanation is not equally convincing for the anxiety rounds. Both anxiety and boredom rounds featured a single enemy unit but were characterized by significantly different blink rates. In a similar vein, anxiety and flow rounds did not present significant differences, even though they featured a disproportionate number of enemy units. Hence, the potential effect of visual demand on the observed blink rate gap may be untenable for this pair.

The condition of boredom was intended to be difficult, corroborating the link between increased blink rate and task difficulty (Nakayama et al., 2002; Ramachandran et al., 2017; Tanaka & Yamaoka, 1993). Earlier studies have asseverated a link between high blink rates and high arousal (e.g., Weiner & Concepcion, 1975). This perspective is befitting in that higher arousal was also inferred from the increased heart rate observed in the boredom rounds. Furthermore, we previously adumbrated the possibility that boredom rounds may have engendered a mixed state, i.e., boredom and frustration. The increased blinking rate observed during the boredom rounds complement this speculation, as high blink rate has also been associated with frustration (e.g., Peters, 2010; Ramachandran et al., 2017).

Another notable finding is related to the blink durations, which were shown to be shorter during the flow and anxiety rounds, compared to boredom. Their differences can be considered an indicator of increased visual workload (Benedetto et al., 2011; Goldstein et al., 1985; Veltman & Gaillard, 1996). Yet, as argued earlier about the blink rates, these differences cannot be attributed solely to visual overload. Long blink durations have also been associated with fatigue, vigilance decrement and drowsiness (e.g., Caffier et al., 2003; Ingre et al., 2006; McIntire et al., 2014), all of which have been viewed as symptoms of boredom (e.g., Aidman et al., 2015; Eastwood et al., 2012; Parikh & Micheli-Tzanakou, 2004). This gives further support to our theory that the boredom rounds may have induced a mixed state of frustration and boredom. Moreover, blink duration decreases have been associated with higher mental load (Marquart et al., 2015), which asserts higher command of cognitive processing during the flow and anxiety conditions.

Contrary to the previous pattern observed in the data, wherein boredom was significantly different from flow and anxiety in all measures, the only variable to have diverged from this pattern was the blink amplitude. This EBV measure was only significantly different between the boredom and anxiety pair and presented higher blink amplitudes during anxiety than during boredom. Blink amplitude seems to become larger as a result of increased cognitive load (Fukuda & Matsunaga, 1983; Ohira, 1995). This observation fits our design, whereby anxiety was meant to be mentally overloading. The apparent challenge in eliminating the bosses in the anxiety rounds may have encouraged players to consider different strategies in order to acquire the coveted reward.

Overall, the physiological responses allow for richer conclusions compared to the self-reported evaluations of the flow experience. The results of this study provided interesting insights on how the target mental states were not always consistent with the game design's intended states. For example, the differences in physiological arousal amongst the three conditions were not in line with our hypotheses, but the distribution of cognitive load across the conditions seems to have been verified. This incompatibility with our original hypothesis might pose classification uncertainty in supervised learners because the classes used as ground truth are given a priori and may thus be unreliable. However, an encouraging (null) finding was the absence of a significant effect of time, which implies that the ground truth preserves some internal consistency over the course of the game. These issues are explored and addressed in the next section.

5.3.2 Classification results

This section is divided into four subsections, (a) the window selection, (b) the original game design labels used as the ground truth, (c) the self-reports of flow used as ground truth and (d) the results obtained from an unsupervised learner as a diagnostic tool, used to inform our design. As a note, the classification performance results in the following sections are exploratory. The exploratory phase involves cross-validated classifiers but with fixed regularization parameters[6] for comparability. This was done to minimize the long optimization times required for the hyperparameter regularization operations. However, the final section (**Diagnosing the findings using an unsupervised learner and regression models**) consists of the optimized classifiers, based on insights garnered from the exploratory phase.

5.3.2.1 Data partitions and window length

This subsection is intertwined with the multiclass approach of the original game design labels as ground truth (exploratory phase). As previously mentioned, we partitioned each game round into 1…9 windows, extracted the features, and constructed the training and validation sets. A classifier was trained at each step. We opted for a Support Vector Machine with a radial basis kernel function as a testbed to perform the search for an optimal window count, because it has been used in classification problems of comparable complexity (e.g., Chanel et al., 2008). The regularization parameters of the model were kept constant at each iteration. The process was repeated for HRV, EBV and the combined HRV and EBV. The results are displayed in **Figure 21**.

The accuracy rates obtained from each fold and all physiological measure (HRV, EBV and HRV and EBV) were aggregated and added to a paired-samples t-test to evaluate differences among consecutive pairs of $window_{n-1}$ and $window_n$. The results indicated that significant differences were found for the following pairs of windows, 2 – 1 ($p < .001$), 5 – 4, 8 – 7, and 9 – 8 ($p < .05$). Due to the inconclusive results, the accuracy rates were collapsed into three regions, by averaging the accuracy rates across three consecutive windows, namely region (a) 1 – 3, (b) 4 – 6, and (c) 7 – 9. A second paired-samples t-test indicated that windowing in region (b) yielded a statistically significantly higher average classification accuracy than in region (a), *p*

[6] SVM: Parameter *C* = 1, *Kernel* = RBF (optimal kernel scale determined via a heuristic procedure, see {'kernelScale', 'auto'} name-value pair parameter in Matlab's *fitcsvm* and *templateSVM* for *fitcecoc* functions); k-NN: *K* = 3, *Distance* = Euclidean; BDT: *Number of learner trees* = 30, *Minimum leaf size* = 5.

< .001, whereas region (c) was statistically significantly higher than regions (a), $p < .001$, and (b), $p < .001$.

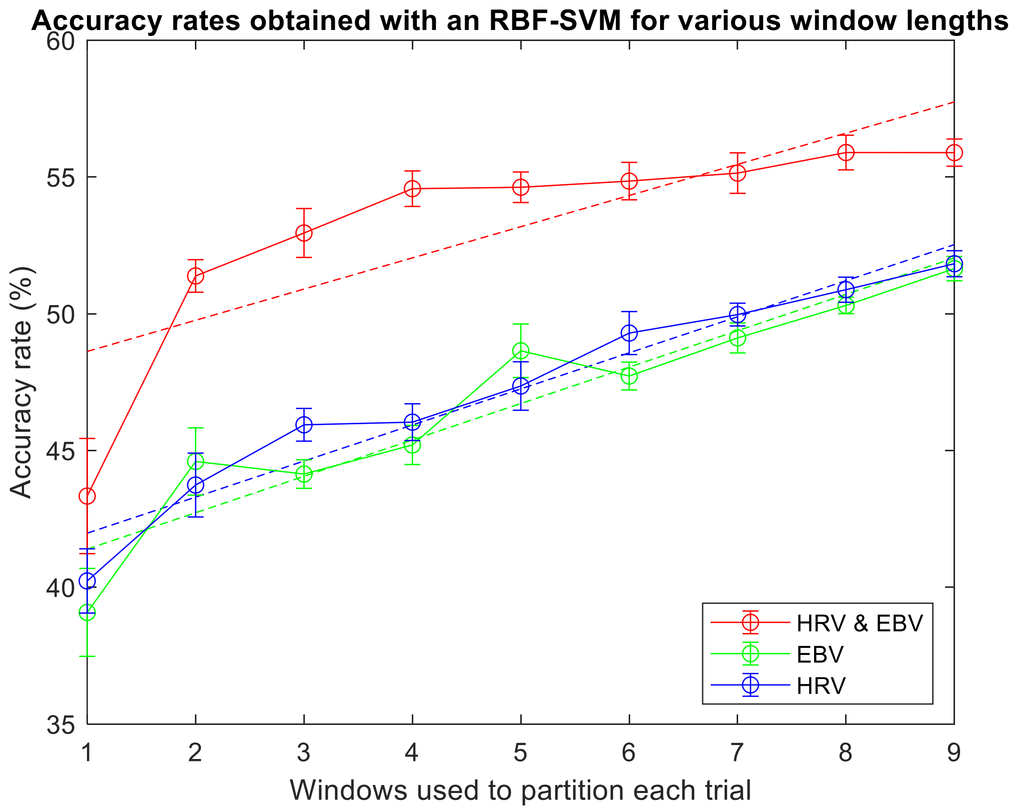


Figure 21: Average classification accuracy rates obtained when testing with a variable number of windows for HRV, EBV and [HRV EBV] data using 10-fold cross-validation. The dashed lines, color-coded to correspond to each physiological measure, show a linearly increasing trend (in the least-squares sense), as the number of windows increases. Most notably, the conjunction of HRV and EBV data provides lower classification ambiguity (error) compared to the individual physiological measures. The error bars represent the standard error.

Although the window length appears to influence the classification of the physiological data, the average performance at each window is somewhat poor. The issue is especially evident when the window count is set to one (i.e., averaging across entire trials). Thus, for practical purposes, we considered seven windows (the lower boundary of region (c)) a reasonable number to partition each trial.

5.3.2.2 Ground truth derived by design in a 3-class classification setting

At a next step, we proceeded with partitioning the data from the two biomarkers (ECG and EOG) with a fixed seven partitions per trial. The data were then used as input to three competitive classifiers, in order to gain a better estimate of the classes' separability. We tested the fusion of as well as singular biomarkers, i.e., HRV and EBV, HRV, EBV. The results of this analysis are displayed in **Table 5**.

Table 5: Classification performance (%) metrics for heart rate variability (HRV), eye blink variability (EBV) and the combination of HRV and EBV, using three different multi-class classifiers and the original game design classes. The results are from a 10-fold cross-validation. The anxiety and boredom classes were oversampled to create balanced datasets for each class, but the validation was performed on the non-oversampled data.

Performance	Support Vector Machine			K-Nearest Neighbors			Bagged Ensemble of CART Decision Trees		
	HRV	EBV	HRV& EBV	HRV	EBV	HRV& EBV	HRV	EBV	HRV& EBV
Accuracy	49.97	49.11	55.14	45.34	43.86	49.44	51.18	50.89	56.47
Specificity	72.26	72.16	75.03	72.17	71.76	74.36	73.81	74.26	76.65
Precision	45.49	44.84	51.53	43.90	43.12	48.55	47.61	48.33	53.30
Recall	45.43	45.08	51.52	45.48	44.79	50.98	48.44	49.95	54.45
F1-score	45.46	44.94	51.51	43.72	42.63	48.40	47.91	48.72	53.74

Following the first impressions from the previous section, the classification results once again demonstrated a significant but poor separability of the flow, anxiety, and boredom trials, regardless of the classifier used. Ergo, they challenge the design approach employed in the game. As expected, the Bagged Ensemble of CARTs generated the lowest out-of-bag errors and presented an advantage in discriminating amongst our classes compared to the other two classifiers. Nevertheless, neither classifier was able to map the physiological responses onto the predicting classes with sufficient accuracy for practical purposes. Hence, we opted for using a different class definition, using the participants' self-reported flow, detailed in the next section.

5.3.2.3 Self-reports of flow as ground truth in a binary classification setting

As a follow-up analysis, the flow scores were split into two halves, corresponding to low- and high-reported flow experience, based on the sum of all flow dimensions per participant. This approach generated binary flow-experience indices for each participant. Given the seven partitions per trial, each observation belonging to a participant was now given the class label of low-flow or high-flow, depending on the participant's rating of their flow experience. The same series of classifiers (SVM, k-NN and BDT) were trained, but this time using a binary class definition.

Table 6: Classification performance (%) metrics for heart rate variability (HRV), eye blink variability (EBV) and the combination of HRV and EBV, using three different binary classifiers. The results are from a 10-fold cross-validation averaged from seven time-windows in which the original data were partitioned.

Performance	Support Vector Machine			K-Nearest Neighbors			Bagged Ensemble of CART Decision Trees		
	HRV	EBV	HRV& EBV	HRV	EBV	HRV& EBV	HRV	EBV	HRV& EBV
Accuracy	78.33	79.90	90.38	74.83	76.45	77.24	81.67	83.73	89.16
AUC	84.80	86.90	96.27	81.13	82.54	82.77	90.27	91.30	95.98
Precision	75.71	81.09	91.07	74.49	76.24	77.28	81.98	84.35	89.89
Recall	81.12	76.12	88.78	72.79	74.42	74.86	79.52	81.39	87.38

As can be seen in **Table 6**, the classification performance using the binary-coded scores from flow's self-reports demonstrates a substantially improved class separability. This indicates the presence of patterns in the physiological data that may not align well with our original hypothesis on the game's design. Naturally, these results are not directly comparable with the multiclass learners shown earlier, since the number of classes differ (uniform distribution chance levels at 50% and 33% respectively). A more balanced comparison will be addressed in the next section.

5.3.2.4 Ground truth derived by self-reports in a 3-class classification setting

The previous section demonstrated the robustness of the nonlinear mapping among physiological variables and the overall flow experience as reported by the subjects. The success of the binary classification based on self-reports suggested the possibility of further enhancing the resolution of the approach by considering a third class. Thus, the same procedure was carried out using a 3-class split of the FSS-2 scores, as was done previously with the binary classifiers. The flow scores were split into three regions thresholded by the percentiles at $\{x <= 33.3\%\}$, $\{33.3\% < y <= 66.6\%\}$, $\{z > 66.6\%\}$.

Naturally, the 3-class split of the self-reports does not necessarily correspond to the game-design labels. This is because the flow scores are one-dimensional and cannot be transposed to the anxiety and boredom states in a straightforward manner. That is, a very low flow score does not warrant or exclude the existence of either boredom or anxiety. Nevertheless, we proceeded as such to demonstrate how the same data with a different class definition at the input – not based on ad-hoc design hypothesis, but on subjects' perception – can enhance the classification performance rates.

Table 7: Classification performance (%) metrics for heart rate variability (HRV), eye blink variability (EBV) and the combination of HRV and EBV, from three different multi-class classifiers using a 3-class split of the scores derived from the flow questionnaire FSS-2. The results are from a 10-fold cross-validation.

Performance	**Support Vector Machine**			**K-Nearest Neighbors**			**Bagged Ensemble of CART Decision Trees**		
	HRV	**EBV**	**HRV&EBV**	**HRV**	**EBV**	**HRV&EBV**	**HRV**	**EBV**	**HRV&EBV**
Accuracy	67.72	73.45	84.33	67.60	69.95	80.87	72.05	76.40	84.14
Specificity	83.67	86.76	92.21	83.72	85.01	90.52	85.83	88.36	92.10
Precision	67.62	73.12	84.02	67.47	69.67	80.55	72.17	76.20	83.84
Recall	67.52	73.55	84.39	67.64	70.13	80.83	71.85	76.62	84.16
F1-score	67.57	73.19	84.14	67.46	69.77	80.60	71.96	76.15	83.96

From the results presented in **Table 7**, it can be observed that, when using three classes, the performance of the classifier naturally degrades than when using two classes (cf. **Table 6**). Yet, the number of classes does not seem to severely affect misclassification rates that would justify the poor performance of our original design classes (cf. **Table 5**).

Although the self-reports provided a promising basis for distinguishing physiological responsiveness to the TD-VR game, they are not highly insightful. The problem is that they only describe a one-time snapshot, i.e., a retrospective perception of subjective experience. The challenge in this case is interpreting the factors that contributed to these ratings as well as the adequacy of the experimental manipulation at any point in the game. As such, a classification learner that is agnostic to our original hypothesis may help elucidate these findings. This is detailed in the next section.

5.3.2.5 Diagnosing the findings using an unsupervised learner and regression models

The reason behind the inadequate mapping of the physiological responses to the game's design remained unclear (**Table 5**). This result contrasts with the reasonably accurate mapping when classes are derived from subject's reported experience (**Table 6**, **Table 7**). To gain a deeper understanding of the misclassification rates in **Table 5**, an approach was devised to identify and understand the heterogeneity within the game-design classes. The approach is based on a robust unsupervised learner, Fuzzy C-means (FCM; see details in Chapter 3, **Unsupervised learner: Fuzzy C-means**), to segment observations into clusters with potentially different characteristics, and in regression models to interpret those clusters, which will finally inform the classification model (the approach is described in the section **Combining Fuzzy C-means and self-report ratings**). Clustering was performed on flow-, anxiety- and boredom-only data separately on both HRV and EBV uncorrected for baseline, and *z*-transformed to reflect the normalization process employed in the supervised learners in the earlier sections.

Cluster interpretation. According to the Davies-Bouldin's optimal *C* selection (Davies & Bouldin, 1979), data were deemed to be optimally separable into three clusters for all design considerations. Thus, the algorithm detected 9 clusters, that are termed 1-3 (for flow round-type), 4-6 (anxiety round-type) and 7-9 (boredom round-type). The proportion of data per participant belonging to each cluster were used as a dependent variable in a stepwise linear regression. The independent variables were the flow scale (FFS-2) and personality scale (BFI) scores of each participant. Results showed that, for each design consideration in the game (flow, anxiety, and boredom), a consistent significant trend arose with the personality trait ratings of BFI (**Table 8**).

Table 8: Results from separate stepwise linear regressions for data of each design condition (columns), and the proportion of participant observations in each Fuzzy C-means cluster (rows). The results demonstrated significant associations with personality as well as flow-related dimensions. Note that double cells in designated rows indicate that more than one linear model were able to explain the data with a significant R^2 change (α = .05). Greek lower-case betas are the slopes of each regression model.

<table>
<tr><th></th><th>Flow</th><th>Anxiety</th><th>Boredom</th></tr>
<tr><td rowspan="2">Cluster 1, 4, 7</td><td rowspan="2">Agreeableness, $\beta = 25.49$, $R^2 = .142$, $p < .05$</td><td>Flow's Sense of Control, $\beta = 8.48$, $R^2 = .144$, $p < .05$</td><td>Conscientiousness, $\beta = -13.66$, $R^2 = .166$, $p < .05$</td></tr>
<tr><td>Extraversion and Flow's Sense of Control, $\beta_1 = -11.51$, $\beta_2 = 8.2$, $R^2 = .414$, $p < .01$, R^2 change = .27, $p < .01$</td><td>Conscientiousness and Agreeableness, $\beta_1 = -14.79$, $\beta_2 = 10.49$, $R^2 = .29$, $p < .05$, R^2 change = .123, $p < .05$</td></tr>
<tr><td rowspan="2">Cluster 2, 5, 8</td><td rowspan="2">Extraversion, $\beta = -31.43$, $R^2 = .364$, $p < .01$</td><td>Agreeableness, $\beta = 12.11$, $R^2 = .156$, $p < .05$</td><td rowspan="2">Extraversion, $\beta = -13.04$, $R^2 = .397$, $p < .001$</td></tr>
<tr><td>Agreeableness and Conscientiousness, $\beta_1 = 13.24$, $\beta_2 = -13.16$, $R^2 = .30$, $p < .05$, R^2 change = .144, $p < .05$</td></tr>
<tr><td rowspan="2">Cluster 3, 6, 9</td><td rowspan="2">Conscientiousness, $\beta = 34.16$, $R^2 = .278$, $p < .01$</td><td>Flow's Automaticity, $\beta = -8.11$, $R^2 = .138$, $p < .05$</td><td>Flow's Automaticity, $\beta = -9.27$, $R^2 = .173$, $p < .05$</td></tr>
<tr><td>Flow's Automaticity and Conscientiousness, $\beta_1 = -5.29$, $\beta_2 = 13.16$, $R^2 = .326$, $p < .01$, R^2 change = .187, $p < .05$</td><td>Flow's Automaticity and Conscientiousness, $\beta_1 = -6.61$, $\beta_2 = 12.37$, $R^2 = .331$, $p < .01$, R^2 change = .158, $p < .05$</td></tr>
</table>

As seen in the table, the personality dimensions of agreeableness, extraversion and conscientiousness, and two flow's dimensions (sense of control and automation), significantly mediated the physiological responses to the original game design (R^2, $p < 0.05$).

Cluster-informed classification. The next step in the approach was to use subsets of data (clusters) that best represented the design conditions (flow, anxiety and boredom). For instance, one combination could include cluster 2 (flow), clusters 5 and 6 (anxiety) and cluster 9 (boredom). Through all combinations of clusters, the optimal 10-fold cross-validation performance was obtained when retaining Cluster 1 for the flow rounds, Cluster 4 for anxiety and Cluster 8 for the boredom rounds. It should be stressed that the original design classes were maintained as the class definition for the training and testing procedures, and the data of each cluster were oversampled, as in the case of the multiclass approach (**Ground truth derived by design in a 3-class classification setting**). The summarized results from this selected subset of data are displayed in **Table 9**. A Friedman's test on the accuracy rates of each fold did not indicate significant differences among the three classifiers' (SVM, KNN, BDT) performance (p = .368).

Table 9: Classification performance (%) of 10-fold, cross-validated, and optimized Support Vector Machine (SVM), K-nearest Neighbors (k-NN) and Bagged Ensemble of CART (BDT) using a subset of the original data. The data were informed by a clustering technique (Fuzzy C-means), consisting of cluster 1 for flow, cluster 4 for anxiety and cluster 8 for boredom. Overall, BDT and SVM were slightly more performant than k-NN.

Performance	Support Vector Machine	K-nearest Neighbors	Bagged Ensemble of CART
Accuracy	84.32	82.34	83.21
Specificity	92.65	91.56	92.08
Precision	81.75	80.33	81.05
Recall	81.67	80.22	80.90
F1-score	81.68	80.20	80.87

The table above shows that the classification performance improved considerably from the classification pooling all data (**Table 5**). Thus, through these subsets of representative data samples, it is possible to effectively predict the game round-type from the physiological responses.

Personality-informed classification. The interpretation of clusters via the regression analyses (**Table 8**) and the success of the classifiers in coherent subsets of the original dataset (**Table 9**), suggested that the classification performance in **Table 5** could be refined. This analysis comprised the entire, original data matrix with all observations. However, each observation was multiplied by each self-report dimension involved in the cluster it belonged (**Table 8**), respecting the sign of the beta slopes (**Combining Fuzzy C-means and self-report ratings**).

For instance, assume the j^{th} observation of the i^{th} participant was assigned to boredom's cluster 7. Since both conscientiousness and flow's automaticity explain the variance of the participants' representation in cluster 7 (**Table 8**), this observation was multiplied by the normalized conscientiousness and automaticity scores of the participant it belonged to (in this example, the multiplier was negative for conscientiousness and positive for automaticity). In short, the approach scales observations by the slope direction and self-report scores as a unitary coefficient; it is thus expected to produce more precise inter- and intra-class separability. Results with the full and transformed dataset are shown in **Table 10**.

Consistent with the hypothesis of the approach, the classifiers using the personality-based weighted approach resulted in a considerably higher predictive capability than the unscaled observations originally used (please refer to **Table 5** for comparison). It should be noted that the personality scores were naturally not used during the game design, and thus the out-of-sample classification shown here is genuine.

To summarize, a specifically designed, personality-informed classifier explained the mapping between physiological responses and game rounds with reasonable success. In the next section, the implications of these findings are discussed.

Table 10: Classification performance of 10-fold, cross-validated, and optimized Support Vector Machine (SVM), K-nearest Neighbors (k-NN) and Bagged Ensemble of CART (BDT) using the normalized personality scores of each participant for three classes, i.e., flow, anxiety, and boredom.

Performance	Support Vector Machine	K-nearest Neighbors	Bagged Ensemble of CART
Accuracy	82.32	83.02	83.81
Specificity	90.74	91.54	91.87
Precision	78.43	78.67	79.79
Recall	77.31	79.86	80.30
F1-score	77.82	78.89	79.97

5.4 Discussion

This study explored the use of heart rate variability and eye blink variability in relation to the flow experience. The results indicated that both heart rate and eye blink variability differ among low and high flow experiences and that these physiological measures may be used as a proxy for mental effort that is common during flow (e.g., Mansfield et al., 2012). Furthermore, it is suggested that there may be differential autonomic activity patterns that are subject to interpersonal traits (e.g., Brouwer et al., 2014b; Stemmler & Wacker, 2010). There are three main considerations worth mentioning about the findings.

5.4.1 Identifying representative cases via self-reports and clustering

The classification of the data using our original game design labels demonstrated a significant but poor performance. However, we were able to identify observations that were more representative of the original design considerations, by using an unsupervised learner along with the scores obtained from two different self-reports (FSS-2 and BFI). The results of this approach testified that different game conditions do not evoke consistent physiological responses to individuals who are exposed to the same tasks.

After selecting a subset of the original data guided by personality scores, a considerable increase was observed in the classifiers' predictive performance. An alternative approach could entail a custom kernel function that would morph the hyperplane based on heuristic rules governed by the regression slopes and their slope magnitude, which would prevent tampering with the actual sample points. However, this was beyond the scope of the present study. It appears that the TD-VR game was not as ill-designed as the original classification performance purported it to be, and that individual differences are necessary for consideration. This leads to the second consideration described next.

5.4.2 Explicating the role of personality

This study substantiated the moderating role of personality and how the latter can sporadically question our assumptions regarding the design of a game. Most notably, our design considerations seem to have been hindered or facilitated by individual personality traits, including extraversion, agreeableness, and conscientiousness. Interestingly, both agreeableness and conscientiousness were illustrated as being conducive to flow in our previous study (see Chapter 4, **Personality scores (BFI)**). However, we had failed to find a similar relationship with conscientiousness, which was instead revealed in the present study.

It should be noted that the game in the present study underwent significant design improvements from the previous, which may have helped expose the role of conscientiousness in the gameplay experience. Previous research has mainly focused on the relationship between personality and the flow experience (e.g., Ullén et al., 2016). However, our understanding is limited when it comes to the game mechanics that best accommodate specific personality types. Extraverts, for instance, tend to appreciate the social aspects of gaming (Park et al., 2011), but other personality types have not been similarly scrutinized in existing research (see Nagle et al., 2016; Nagle et al., 2018; Van Lankveld et al., 2010). This is an important topic for future consideration, as it may help us understand how to successfully reach broader audiences.

5.4.2.1 Flow design

By reviewing our findings, one can notice that the game rounds designed to elicit flow have had a positive relationship with agreeableness and conscientiousness. To recapitulate the mechanics that were unique to these rounds, we adopted a dynamic difficulty adjustment system that increased the health of the enemies and decreased the reward when eliminating them, if the player performed well. In contrast, the health of the enemies was decreased, and the reward magnitude was increased, when the player did not perform well.

Conscientiousness and agreeableness have both been positively related to flow (Johnson et al., 2014; Tatalović Vorkapić & Gović, 2016; Ullén et al., 2012, 2016). These traits have been discussed in the previous study (see Chapter 4,

Discussion). Yet, conscientiousness was not found to relate to flow in our previous study, but this link surfaced in the present one. Conscientious individuals tend to exert more effort in tasks (Yeo & Neal, 2004), especially when the rewarding schedule is based on their performance (Nagle et al., 2018).

In our study, the performance of each individual led to a decrement or increment of reward magnitude, based on the dynamic difficulty scaling, which is corroboratory of the conclusion by Nagle and colleagues (2018). By extension, the additional effort that characterizes conscientious individuals, as per Yeo and Neal (2004)'s notes, may have facilitated the onset of the flow experience (Waterman, 2005). Notably, users in flow are by definition more involved in a task than those who are not (Webster et al., 1993). In conjunction with the aforementioned observations, there appears to be a cycle of flow maintenance in conscientious individuals, mediated by the game's reward mechanics.

On the other hand, even though extraversion has been shown to have a similar, positive correlation with flow (Tatalović Vorkapić & Gović, 2016; Ullén et al., 2016), our flow rounds presented a negative relationship instead. A possible explanation for this finding is that the game was single-player and did not have a social layer to satisfy the motives of extraverts in video game play (e.g., Fang & Zhu, 2011; Park et al., 2011). Hence, extraversion may be thought of as an inhibitory factor to experiencing flow in our design. This explanation is in line with Jurnet and colleagues (2005), who argued that introverts are more likely to become immersed and less distracted in a virtual environment – the virtual reality platform used in this study facilitates that.

Further support can be drawn from a recent study, where it was shown that extraverts are more likely to experience flow in social activities than solitary (Liu & Csikszentmihalyi, 2020). This has also been confirmed in the context of video game play, where individuals with high levels of extraversion preferred social activities, whereas those with low levels preferred solo play (Graham & Gosling, 2013). It would be interesting to examine whether a collaborative TD-VR would amplify the expression of extraversion during the flow rounds. Arguably, this conclusion may discount other properties that could have engendered the observed negative link between extraverted individuals and the flow/anxiety conditions of the TD-VR game. For example, extraverted individuals may exhibit strong preference towards avatars, leaderboards and point-based systems (Denden et al., 2018). However, the current study lacks data to specify the properties that can explain the observed association.

It should be noted that the negative finding on extraversion may be interpreted as contradictory with the positive link found in the previous study (**Personality scores (BFI)**). However, the relationships identified there were based purely on self-reports, i.e., FSS-2 being the dependent and BFI the independent variables. Contrarily, this study addresses the relatedness of personality dimensions to each experimental condition based on the physiological data of each player and their membership to clusters decided by an unsupervised learner. Thus, there were no direct comparisons made between the FSS-2 and BFI scores.

5.4.2.2 Anxiety design

Similar to flow, anxiety rounds revealed three distinct patterns. A fast-moving unit which would grant a generous reward to the player if eliminated was the main mechanic introduced as part of the design. However, this reward was impossible to attain, because of the experimental manipulation. Amongst the physiological patterns, we identified a negative association between the physiological responses during the anxiety rounds and extraversion and flow's automaticity. In addition, we found a positive relationship with flow's sense of control and agreeableness, whereas conscientiousness appeared to have had both a positive and a negative relationship. It should be noted that the finding on conscientiousness is not contradictory. The stepwise regressions demonstrated that a negative relationship with conscientiousness was found in Cluster 5, and a positive relationship in Cluster 6. Hence, they are pertaining to different partitions of the data.

The positive link with conscientiousness may suggest that a portion of participants vied for the high reward promised and perhaps engaged in different strategies to acquire it. Indeed, planning, diligence and goal-directed behavior have long been shown to characterize high scorers on conscientiousness (e.g., Roberts et al., 2009; Witt et al., 2002). Because of the planning involved, automaticity is reasonable to have presented a negative link in the same model, as it relates to spontaneity (Csikszentmihalyi, 1990). In contrast, a different model (Cluster 5) evinced low conscientiousness during anxiety rounds in presence of high agreeableness, suggesting an interaction of the two traits. According to Johnson and Ostendorf (1993), the inverse-directional combination of these two traits alludes to easy-going individuals, who are typically less impatient and less susceptible to time pressure (Froggatt & Cotton, 1987). Perhaps, the combination of these traits was the least compatible with our intention to induce anxiety.

Another significant pattern that emerged from the resulting clusters was the combination of high sense of control and low extraversion, accounting for a large portion of the first cluster's data (41.4%). This finding denotes that there is a distinct physiological activation when these traits interact, but this physiological signature becomes more apparent within the anxiety design. Indeed, the combination of these traits increased the likelihood that the corresponding participants were grouped under the same cluster. Introverts tend to feel more autonomous in non-social contexts (Feist, 1999). In addition, they are more susceptible to anxiety, but exhibit higher levels of cognitive control (Hurd, 2005; Matthews & Dorn, 1995). According to this definition and the finding that introverts are characterized by high autonomic arousal (Eysenck, 1981), it may not be surprising that the aforementioned traits were identified by

FCM as a distinct pattern in the anxiety design. As mentioned earlier, this cluster provided the best decision boundaries for the anxiety rounds in comparison to the other two clusters.

5.4.2.3 Boredom design

As per the rounds designed to elicit boredom, a negative relationship with extraversion and flow's automaticity was found, but a positive one with agreeableness. As with the anxiety design, boredom presented the same relationship with the flow's dimension of automaticity and conscientiousness. In addition, both designs evinced bi-directional conscientiousness in two different clusters. As a reminder, in the discussion of the statistical findings (**Discussion**), it was pointed out that boredom may take the form of anxiety when the perceived sense of autonomy is subverted (see also, Van den Hooft & Van den Hooff, 2018). The results here extend this theory and promote conscientiousness to a possible mediating factor. Perhaps the combination of automaticity and conscientiousness was not exclusively related to the reward mechanics, as argued earlier in the anxiety design. Instead, it could relate to the overall strategic planning that may have ensued from the perceived high challenge of those rounds.

In a study by Sulea and colleagues (2015), conscientiousness was also shown to be negatively related to boredom. The authors explained this finding by suggesting that high competence, that characterizes individuals with high conscientiousness, may suppress the onset of boredom. Indeed, conscientiousness has been associated with decreased proneness to boredom (Seib & Vodanovich, 1998; Von Gemmingen et al., 2003). In addition, as noted earlier in the flow design section, conscientious individuals will feel more motivated to perform well in a task where their performance is rewarded (Nagle et al., 2018). In the design of boredom in TD-VR, participants were informed that the reward incentive for the bosses in these rounds was very low. Perhaps, this design further amplified the negative link between conscientiousness and the rounds aimed to elicit boredom.

5.4.3 Finding the right audience

The results of this study allude to the TD-VR game being more suitable to specific personality types, which is an insightful remark for game studios who can tailor their marketing strategies to specific target audiences. Several studies have shown that when the promotional content of an advertisement has certain characteristics, the advertisement will be received more positively by individuals with matching personality (e.g., Hirsch et al., 2012).

For example, conscientious individuals may prefer content that implies efficiency and the pursuit of goals, individuals with high levels of extraversion may prefer messages pertaining to social rewards, whereas openness to experience may better tap into messages portraying creativity and intellectual stimulation (Hirsch et al., 2012). Identifying the optimal demographic for a video game may leverage similar marketing tactics.

Alternatively, further work is recommended to determine how different game mechanics best attract various personality archetypes (as in e.g., Nagle et al., 2016; Nagle et al., 2018). Such work would illuminate the path to improving problematic areas in the game, i.e., game levels or mechanics that could reach broader audiences. Nevertheless, it should be acknowledged that a game may never be able to fully accommodate every player's preferences (Tognetti et al., 2010).

To conclude, the approach detailed in this chapter provides a first step to accomplishing this, primarily by casting light onto individual differences that may be incompatible with the design of a game. While the exploration was limited to the Big Five personality traits and flow, other psychological constructs from self-report assessment measures may likewise provide rich insights about a game's ability to engender the desired mental states.

6 Study III: EEG theta/beta ratio from six scalp sensors can reliably predict the flow experience

6.1 Introduction

6.1.1 Neural correlates of the flow experience

The multifaceted structure of flow suggests that there is a large functional system in the brain able to support flow's experiential components. The nervous system comprises two main branches, the peripheral nervous system (PNS) and the central nervous system (CNS) (Brodal, 2004). The CNS is involved in the modulation of mental processes, which occur through the complex communication among neurons as well as ensembles of neurons (e.g., Brodal, 2004). These processes include, but are not limited to, perception, attention, memory, emotion recognition and experience, motor execution and decision-making.

6.1.1.1 Functional connectivity

Flow has been associated with the dopaminergic system (De Manzano et al., 2013; Gyurkovics et al., 2016; Marr, 2001; Weber et al., 2009), the sensorimotor network (Klasen et al., 2012; Sanchez-Vives & Slater, 2005), the combination of anterior cingulate cortex and temporal pole (Yun et al., 2017), and reduced activity in the prefrontal cortex (Bavelier et al., 2012; Dietrich, 2004; Goldberg et al., 2006; Gusnard et al., 2001; Ulrich et al., 2014). The prefrontal cortex (PFC) remains a debated brain structure for its role in flow. Dietrich (2004) hypothesized that flow reflects reduced frontal brain activity, and basal ganglia are deployed to allow a state of automatic behavior and executive functioning – a state known as transient hypofrontality (Dietrich, 2004).

The state of hypofrontality has not been consistently confirmed (Harmat et al., 2015; Harris et al., 2017b; Leroy & Cheron, 2020; Yoshida et al., 2014). In a recent study by Causse et al. (2017), the authors found that activity in the PFC is proportional to the task's demands. According to Dietrich (2004), information processing is achieved through the implicit and the explicit systems, which are mainly distinguished by their projections to consciousness. Dietrich argued that each system is not exclusive to task performance and the system to be employed may depend on the nature of the task (Dietrich, 2004). Notably, explicit control on the task may not be as essential during flow as the implicit (e.g., Bavelier et al., 2012; Dietrich, 2004; Taylor, 2002).

Indeed, task automation is a distinctive feature of flow (e.g., Csikszentmihalyi, 1975; Quinn, 2005), and it progressively loses its dependency on the PFC as expertise increases (Léger et al., 2014). Klasen et al. (2012) recruited expert players and a conjunction analysis of flow factors revealed a neocerebellar-somatosensory network. Contrarily, Yoshida and colleagues (2014) found increased activity in the left and right ventrolateral PFC during flow. However, Yoshida et al. did not specify whether the participants in their study were skilled players. This is an important factor, in that flow may orchestrate neural networks differently for novice and expert players (e.g., Hafeez et al., 2021; Hilla et al., 2020; Kirschner & Williams, 2014), with novice players requiring more explicit control (e.g., Dietrich, 2004).

Similarly, evidence on the loss of self-reflective thoughts during flow has revealed reduction in the activity of the medial PFC (Ulrich et al., 2014), whereas top-down attention implicated in flow (Harris et al., 2017a) has been associated with the lateral PFC and frontal eye fields (Buschman & Miller, 2007). These contradictory findings should be interpreted with care; flow is perhaps related to a localized hypofrontality, and not as universal as Dietrich (2004) originally speculated (Harris et al., 2017b). It can be suggested that frontal brain reduction may be a function of time. For example, Yun et al. (2017) found that flow peaked after 25 minutes of game playing. Hence, the player may require time to transition from explicit to implicit control and to allocate attentional resources that will make him/her resistant to distractions (e.g., Bavelier et al., 2012; Jackson & Csikszentmihalyi, 1999; Nuñez Castellar et al., 2016; Jennett, 2010). As Dietrich (2004) argued, the explicit system may still be necessary to some extent, regardless of the individual's expertise, and the task's requirements will solicit the contribution of either the explicit or the implicit system.

Klasen and colleagues (2012) have additionally identified the role of parietal regions (i.e., superior parietal cortex, precuneus and intraparietal sulcus) in flow. However, an occipital and frontoparietal network has also been found to underlie spatial awareness in the first-person perspective (Vogeley & Fink, 2003; Vogeley et al., 2004), which is the same perspective Klasen et al. used in their study. Yoshida and colleagues (2014) cautioned this case, in that concurrent neural networks, specific to the task, may have been misidentified for flow dimensions. Different game genres can have specific cognitive demands (e.g., Latham et al., 2013), and, as such, certain video games (e.g., strategy games) may require more explicit control than others (e.g., Spence & Feng, 2010).

These games might stimulate higher prefrontal activation, without necessarily implying that it is a substrate of flow. Importantly, the frontoparietal network's contribution to the experience of flow remains questionable (Bavelier et al., 2012; Léger et al., 2014; Nah et al., 2017). The responsibility of the frontoparietal network to allocate attentional resources and the reduced activation thereof in video game playing (Bavelier et al., 2012) suggests a functional unrelatedness to flow. This is evident from flow's dimension "merging of action and awareness", which signifies that attention is already focused (Csikszentmihalyi, 1975).

An interesting proposal put forward by Weber and colleagues (2009) is the synchronization theory of flow. According to the authors, the synchronization process refers to groups of neurons firing in a synchronous rate to produce discrete state transitions, given a skills-challenge balance during the task, thereby conceptualizing flow's make-up as a complex neural infrastructure. The theory further posits that the synchronization is energetically economical, which makes flow perceived by the individual as a pleasant, rather than a mentally taxing, experience. Their core idea is that flow is orchestrated by the synchronization of attention and reward networks. The attentional networks consist of an alerting network of fronto-parietal origin, whereas a secondary attentional network, the orienting network, comprises the superior and inferior parietal lobe, the frontal eye fields and the superior colliculus. On the other hand, the reward networks include the dopaminergic system, the orbitofrontal cortex, the ventromedial and dorsolateral prefrontal cortex, the thalamus and the striatum. The theory concludes with the proposal that focused attention is inherently rewarding, thereby justifying flow as the result of the synchronization between attentional and reward mechanisms (Weber et al., 2009).

Reward networks have been traditionally linked to dopamine release (Iversen & Iversen, 2007), which occurs during reward anticipation as well as the delivery of rewards (Drecher et al., 2009). Studies investigating the effects of rewards on the activation of dopaminergic pathways typically employ extrinsic rewards (Baldassarre, 2011; Murayama et al., 2010). Though this approach is still relevant to video game play, as a facet of extrinsic motivation (see Chapter 2, **Reward systems**), it is not clear whether a similar neural relationship between the dopaminergic system and intrinsic motivation exists (Di Domenico & Ryan, 2017). The self-conflicting semantics of autotelicity that renders it both an antecedent and an outcome of the flow experience (Michailidis, Balaguer-Ballester, & He, 2018) can be associated with the two dominant views on dopamine and reward processing.

Berridge (2007) pointed out that dopamine release is mostly related to "wanting" (incentive salience view) rather than "liking" (hedonic impact view). Under this distinction, autotelicity as an outcome matches the hedonic perspective, which accompanies the self-rewarding experience of flow (e.g., Berridge, 2007; Keller & Bless, 2008). In contrast, autotelicity as an antecedent would presumably be an instance of incentive salience, in that the individual is motivated to perform the activity before the actual experience of flow (e.g., Delle Fave et al., 2011b). This is also supported by a previous study, wherein the number of volitional engagements with the task was used as an assessment criterion for intrinsic motivation (see Murayama et al., 2010), i.e., the number of times the participants *wanted* to perform the same task. In the specific case of flow, individual proneness to flow has been shown to relate to nigrostriatal dopamine (e.g., De Manzano et al., 2013) but also the mesolimbic pathway (e.g., Klasen et al., 2012), which relates to behavioral persistence (Salamone & Correa, 2016). Even though a few theories have been put forward with the aim to identify the neural underpinnings of intrinsic motivation (e.g., Baldassarre, 2011; Di Domenico & Ryan, 2017), there is still little evidence to draw firm conclusions.

Based on these accounts, it can be inferred that flow's complexity of a construct is likewise reflected in the ambiguity stemming from neural studies of flow. Hence, the inconsistency of existing findings makes it challenging to interpret the contribution of each flow dimension to the ultimate flow experience. This section highlighted the potential presence of confounding factors that may have made the deduction of flow's neural underpinnings more challenging. These factors included the particularities of the task under which flow is expected to be elicited and the dichotomy of extrinsic and intrinsic incentives. Of note is also that the aforementioned findings have been derived from magnetic resonance imaging, which is an unnatural setting for video game playing. In the next section, findings from studies using electroencephalography are reviewed, a method with the benefit of enhanced temporal resolution.

6.1.1.2 Electroencephalography

Flow has also been estimated with electroencephalography (EEG), which is the process of recording electrical signals from the scalp (Jatupaiboon et al., 2013; Teplan, 2002). Under the international 10-20 EEG guidelines (Jasper, 1958), the placement of electrodes on the scalp is standardized using the nasion and inion as the reference points along the sagittal axis, and the pre-auricular points along the coronal axis. 10-20 thus refers to the percentual distances between consecutive electrodes over the total length of the corresponding axes, thereby accounting for individual differences in head sizes. The output of EEG is the voltage difference between two electrodes at a given time point, and thus voltage values are relative (Cohen, 2014). This approach allows for high temporal resolution, but its spatial resolution is relatively weak (Srinivasan, 1999). The advantage of time precision is an important consideration to brain-computer interface (BCI) applications, which should be responsive in a timely manner. Although a slower detection of physiological changes in a real-world video-game application may be tolerated, making the high temporal precision of EEG somewhat superfluous, empirical research has demonstrated the potential of EEG as a tool to improve player experience.

The idea behind the use of EEG in games research is akin to the general purpose of BCI applications, i.e., to translate neural signals into information that is interpreted and handled by a system layer that mediates the software and the user (Minguillon et al., 2017). The player's interaction with a video game is commonly achieved via input to a control device, e.g., a mouse, a keyboard, a gamepad, or a touchscreen (e.g., Cummings, 2007). Likewise, brain activity can be deconstructed as a second-level input. Brain-computer interfaces research has demonstrated the successful utilization of brain signals as an alternative input method to control and/or adapt a game based on the user's collective neural information at a given time (e.g., Liao et al., 2012).

The EEG waveform has been conventionally divided into five main bands, i.e., alpha, theta, beta, delta and gamma. These are determined by the frequency range in which the electrical oscillations occur. In practice, the range of each band is somewhat standardized, though differences in their definitions are not uncommon (Berta et al., 2013). Such differences may be found not only in the upper-band, but also in the lower-band limit, or both.

Nevertheless, common ranges for each band are as follows: alpha (8 – 13 Hz), theta (4 – 8 Hz), beta (13 – 30 Hz), delta (0.5 – 4 Hz), gamma (30 – 100 Hz; Li & Lu, 2009; Schutte et al., 2017; Teplan, 2002). Through spectral analyses, activity from these bands can have different implications on the underlying cognitive functions, especially when considering an apparent topographical bias. For instance, alpha appears more prominently in posterior/occipital sites (Knyazev, 2007; Teplan, 2002), whereas beta is found mainly in the frontal lobe (Taywade & Raut, 2014). Likewise, the alpha rhythm typically implies reduced cortical activation (Allen et al., 2001, 2004; Davidson, 1988), whereas the beta rhythm has the opposite implication (Rangaswamy et al., 2002).

Flow and neighboring constructs, such as that of engagement, have been approached with EEG. Nuñez Castellar and colleagues (2016) used a modified 10-20 EEG placement system and recorded activity from the locations Fz (midline frontal), FCz (midline frontocentral) and Cz (midline central) while the participants played a video game. Among their three conditions for flow, boredom and frustration, they used auditory distractors during the game session. The authors measured reaction times and accuracy of the reactions when an auditory distractor was presented to the participants.

Flow condition was characterized by longer reaction times (also, Ko & Ji, 2018) and lower accuracy, and increased alpha band activity in the electrode FCz, compared to the boredom and frustration conditions. The authors postulated that increased alpha in FCz indicated the interplay between executive control and attention, which alludes to flow’s criterion of "merging of action and awareness" (automaticity). Finally, they suggested that this process ensues from the contribution of the anterior cingulate cortex, an area that has been shown to be of relevance to flow (e.g., Klasen et al., 2012; Ulrich et al., 2016; Yun et al., 2017).

Beta activity from the occipital sites was also found to correlate with the flow experience in Tay (2016). The author used a visual tracking task and introduced four levels of difficulty, each level adding one additional object (dot) to track (four objects at the fourth level). Participants were asked at the end of each trial to specify which dots they were originally asked to track. The ratio of incorrect trials over the total number of trials was used as a behavioral measure of error rate. The results showed that beta amplitudes increased proportionally to the task difficulty for the high-flow group but decreased for the low-flow group. Given that the nature of the task was visual, it is not surprising that the apparent resource allocation in the visual cortex scaled with the task's difficulty. Interestingly, the error rate of the high-flow group similarly increased according to the beta amplitude, suggesting that the participants felt challenged. In a similar vein, Yelamanchili (2018) found a reduction of occipital alpha during Tetris game playing in a difficulty level aimed to induce flow.

Other studies have approached a broader construct of flow, namely engagement, and identified this state by a set of indices. McMahan et al. (2015a) proposed that engagement can be estimated by the [Beta / (Alpha + Theta)] (also, Pope et al., 1995) bands averaged from all channels. In their follow-up work (2015b), they proposed that an arousal index, computed from the beta and alpha frequency bands at the F3 and F4 sites of the bilateral prefrontal cortex, [BetaF3 + BetaF4 / AlphaF3 + AlphaF4], should accompany the engagement index, in which case the user is thought to be disengaged when arousal is low, and to be engaged when the arousal index is high. Of note is also the engagement index proposed by Biercewicz et al. (2020) who found that (Theta / Alpha) best correlated with the self-reported levels of engagement.

What is particularly interesting is that the engagement index by McMahan and colleagues (2015b) suggests that higher engagement can be inferred from theta activity relative to beta. Activity in the theta band has been shown to emerge during video game play (e.g., He et al., 2008; Nacke et al., 2011; Sheikholeslami et al., 2007), and evidence has shown a link to the flow experience as a result of cognitive control (Katahira et al., 2018; Metin et al., 2017). Murphy and Higgins (2019) measured EEG wave amplitudes in five frequency bands during three different conditions, relaxation, concentration and stress. While they could not infer a stable pattern for concentration, they found increased theta and low alpha to underlie the stressful condition. Consistent with these findings, Bodala et al. (2016) identified increased theta activity during challenge integration, particularly in the frontal midline, which can be considered a marker of vigilance (also, Yamada, 1998). In Metin et al. (2017), theta increases were found across a frontocentral network exhibiting positive correlations with enjoyment and intrinsic motivation self-ratings. Similarly, Derbali and Frasson (2010) reported a positive correlation between frontal theta and motivation. Meditative states have likewise been found to be accompanied by increases in theta power (Baijal & Srinivasan, 2010; Lagopoulos et al., 2009) while several studies have drawn parallels between meditation and the flow experience (e.g., Delle Fave et al., 2011a; Krygier et al., 2103).

The theta wave has given rise to another useful index in the literature, the theta/beta ratio, also referred to as the "attention ratio" (Derbali et al., 2011; Putman et al., 2010). The slow wave of the theta band over the fast wave of the beta band is thought to reflect subcortical processes regulated by endogenous cortical activity (Schutte et al., 2017). This interpretation has further been associated with the regulation of affective processes and attentional control (Putman et al., 2010) and a disputed proxy for arousal levels (Clarke et al., 2019). Angelidis and colleagues (2018) argued that the theta/beta ratio reflects cognitive control over affective information. In other research works, low values of this ratio have been found to reflect distractibility (Kobayashi et al., 2020; Ming et al., 2009). The theta/beta ratio has also been investigated in video games as a marker of stress (Saputra & Iqbal, 2017). These implications make the theta/beta ratio an interesting measure for the study of flow, anxiety and boredom.

Though this section did not exhaust the literature in the electrophysiological findings of flow, it still is a nascent area that remains largely unexplored. Notably, the alpha, theta and beta bands are well investigated in the literature of the flow experience. Contrarily, research examining the delta and gamma waves in relation to flow is exceptionally scarce, with only a few studies linking these bands to immersive experiences during video game play (e.g., McMahan et al., 2015c; Metin et al., 2017; Nacke et al., 2011). Thus, although there is an evident gap in the literature of flow with respect to these bands, the main deterrent as per their omission is that their relatedness to flow is not well understood. However, technical challenges may also present. For example, the gamma band typically coincides with the spectral range within which muscle artifacts fall (approximately 20-300 Hz; Muthukumaraswamy, 2013), making it challenging to isolate in experiments involving head motion as part of the user interaction. The frontal-theta power increases linked to flow and motivation are reminiscent of a neighboring metric that is common to the conventional EEG research but less popular in the flow research. Alpha asymmetry is reviewed in the next section.

6.1.1.3 Alpha asymmetry

Frontal alpha asymmetry (FAA) is another EEG metric that may be relevant to the quantification of flow episodes. This metric was used by Davidson and Fox (1982), who observed greater left-hemispheric activation, relative to right, as a neural response to happy video clips, specifically in the frontal regions of the cortex. Since then, FAA has garnered considerable interest in the field of affective neuroscience, and it has been associated with emotional experience (Davidson et al., 1990; Fox, 1991), affective regulation (Reznik & Allen, 2018), approach and withdrawal motivation (Briesemeister et al., 2013), but also motivational tendencies as facets of personality from resting state EEG recordings (e.g., Sutton & Davidson, 1997).

The FAA index is expressed as the difference between the alpha activity of the right hemispheric sites and the left contralateral sites in a logarithmic format, i.e., *log*(alpha_right)-*log*(alpha_left). The *relative* activation of each hemisphere implied by the FAA index is based on the notion that increases in alpha power are associated with decreased cortical activation (Allen et al., 2004; Davidson, 1988; Salminen et al., 2009). In addition, a consistent association between lateralized frontal brain activity and affective valence has been shown, but also been disputed in several works. In earlier works, inter-hemispheric activity differences in the alpha band were associated with the motivational branches of approach and withdrawal (Davidson, 1984). Approach motivation has been characterized by dominant left activity, relative to the right, through positive FAA index values, whereas avoidance behavior by dominant right activity relative to the left (for a review, see Briesemeister et al., 2013). In later works, greater left activation was thought to represent positive affective experience, whereas greater right activation was thought to represent negative affective experience (e.g., Heller, 1990).

Notably, due to the illustrated relationships, literature conflated the terms of emotion and motivation (Harmon-Jones, 2003), perhaps explaining why the relationship between the FAA index and the valence spectrum of emotion remains inconclusive (e.g., Spielberg et al., 2008). For example, a recent study by Tandle and colleagues (2016) found empirical evidence that sadness elicited higher right-frontal activity, whereas happiness elicited higher left-frontal activity, after exposing participants to different types of musical tracks. The authors specifically argued that right-frontal sites may relate to anxiety. These findings are germane to the view that frontal asymmetry relates to valence.

The roots of this confusion can be traced to earlier theories of motivation. For example, in the work of Dickinson and Dearing (1979), motivation was theorized to comprise two systems, the attractive/appetitive and aversive/defense systems. According to the authors, affective valence was determined by the system that would be activated in a particular moment, leading to positive emotions (appetitive) or negative emotions (defense). On the other hand, affective arousal was suggested to reflect the intensity of the motivational mobilization of the corresponding system. Likewise, Gray (1994)'s theory postulated that two systems underpinned motivation, the behavioral approach system (BAS) and the behavioral inhibition system (BIS), which remain the prevalent theoretical considerations behind the FAA index (also, Elliot, 2006). According to Gray, the approach system is engaged under conditioned rewards or instances of non-expected punishment, whereas the avoidance system is employed under conditioned punishment, absence of rewards or innately aversive stimuli (such as fearful stimuli).

These theories indicate a close link between emotion and motivation, rendering affective valence and the approach/withdrawal systems naturally aligned constructs. Indeed, Pessoa (2009) acknowledged this familiar link and suggested that it stems from the constructs' dependence on the relationship between the person and the environment. Another example, that demonstrates how intertwined these concepts are, is McClelland (1985)'s view that motivation is an affective force that orients individuals toward positive or negative stimuli. More recently, Harmon-Jones (2003) provided insights that differentiate these two constructs. The author argued that anger, an emotion that is widely regarded as negative in valence, can often result in an approach behavior such as attacking. The author further posited that this example severs the putative connection between FAA and affective valence, and instead suggested that FAA reflects motivational direction, which the author clearly distinguished from valence.

However, in the specific case of flow, no particular stance was adopted throughout the present work with respect to its affective valence characteristics. Albeit a large body of literature suggests that flow is a state of positive affect (Jennett et al., 2008; Poels et al., 2012), empirical research in video games holds that both positive and negative emotions can coexist without necessarily compromising the sustenance of the flow state (Csikszentmihalyi, 1990; 1996; Kaye et al., 2018; Walker, 2010). Csikszentmihalyi (1990, 1996) pointed out that flow may not be a positive experience while it occurs, but it may be appreciated only after the end of the activity, as an after-effect. This leads to the interpretation that, during flow, affective valence may not be an important indicator of the quality of the flow experience. Indeed, Mercier and colleagues (2019) found that *arousal*, not valence, was the affective dimension that made players appreciate their game experience.

On the other hand, motivation has a clearer relationship with flow. Earlier, both intrinsic and extrinsic motivation were argued to be necessary constituents of the flow experience in video games (see Chapter 2, **Reward systems**). To recapitulate the main points of that discussion, flow as an intrinsic experience encourages replay intention over several time instances (Van den Hout et al., 2018). However, video games are also rife with extrinsic rewards which empower one's motivation to prolong engagement with the activity within a single time instance. Hence, extrinsic motivation should be easier to examine, given that the extrinsic rewards can be manipulated during an ongoing game session.

As a reminder, approach/withdrawal motivation theories have traditionally employed extrinsic rewards as a means of demonstrating how approach and withdrawal are activated, based on their benefit and desirability for the organism (Davidson, 1984; Dickinson & Dearing, 1979; Gray, 1994). The behavioral biproducts of approach and withdrawal may be expressed, for example, with postural changes, whereby leaning forward indicated approach motivation, whereas leaning backwards may imply withdrawal motivation (Price & Harmon-Jones, 2011). D'Mello and colleagues (2007) found that, during video game playing, participants leaned forward and closer to the computer screen, which the authors interpreted as a flow episode. In the same study, postural changes away from the computer were interpreted as boredom.

This may also be extrapolated to the state of anxiety. For instance, Davidson and colleagues (2000) found increased right-frontal activity, relative to the left, when people with social phobia anticipated a public speech; along with heart rate variability, those changes accounted for more than 48% of their self-reported levels of anxiety. Though Elliot (2006) suggested that negative affective states or negative stimuli may be mostly associated with withdrawal or avoidance motivation, Harmon-Jones and Allen (1998) found that anger can violate that assumption. An empirical example of this in the context of video games can be seen in the study by Salminen and colleagues (2009).

The authors argued that challenging events may steer toward withdrawal behavior, manifested through frontal alpha asymmetry and right-frontal activation specifically. Yet, the authors also found that the participants reported positive valence even after the game session, a finding corroborating that withdrawal behavior during the game may not necessarily reflect affective valence, as Harmon-Jones (2003) argued. According to Davidson (1984), affective valence, albeit not lateralized, is manifested as such because it is associated with approach/withdrawal behavioral biproducts.

Based on these accounts, affective valence appears to have an implicit relationship with FAA, but research converges to the fact that approach and withdrawal motivation can be expressed through this index. Since it was argued earlier that flow may be affect-independent while it occurs, i.e., not part of flow's core dimensions (Michailidis, Balaguer-Ballester, & He, 2018), integrating FAA merely on the basis of approach and withdrawal, without making direct implications on its affective valence underpinnings, fits well with this conception of flow. Unfortunately, very few studies have utilized the FAA index metric in flow research, making it a tempting quantification tool to explore. Benlamine and colleagues (2017) are among the few works to have used alpha asymmetry as a proxy measure. The authors did not provide any statistical findings, but instead integrated the FAA index directly within a classification approach for predicting flow episodes.

To summarize, EEG research offers several proxy tools to estimate underlying cognitive mechanisms with a high temporal accuracy. Although the spectral subdivision of EEG activity has been criticized for being somewhat outdated, it is also acknowledged to constitute a golden standard for more than 80 years (Newson & Thiagarajan, 2019). Spectral dynamics in the alpha, theta and beta bands have been shown to support the investigation of immersive experiences during video game playing and cognitively challenging tasks. While some works have identified a relationship between activity in these bands and cognitive-affective functions commonly engaged during flow, others have made explicit connections to the flow experience. Affiliated indices, such as alpha asymmetry, engagement index, and the theta/beta ratio are important markers to consider, that provide a promising basis for their application in virtual reality settings.

6.1.2 Study objectives

The present study aims to extend our findings on flow, using EEG, to examine its feasibility in the commercial frontline. Given that medical-grade EEG typically outnumbers the sensors used in consumer-grade equivalent options, this study incentivizes the selection of an optimal subset thereof for flow-aware technologies. The approach employed in the previous chapter (and detailed in **Combining Fuzzy C-means and self-report ratings**), that combines clustering of physiological responses to a VR game and regression models onto self-reports of inter-personal traits, showed significant class discriminability amongst the game's experimental conditions.

The primary goal is to identify a set of channel/frequency band pairs to optimally disambiguate the three design conditions of flow, anxiety, and boredom. Thus, we examine the electrophysiological responses to the TD-VR game using a relatively high-resolution EEG and later reduce it to a subset of channels and frequency bands that are most relevant to the flow experience.

Regarding the frequency bands of interest, literature does not seem to provide clear evidence on their topographical relevance for flow that could effectively guide our hypotheses. Nevertheless, theta, alpha and beta have all been shown to be of relevance in studies examining flow or neighboring mental states. Based on existing findings, our hypotheses for each frequency band are as follows:

- Theta may be more pronounced in the flow rounds, as a result of cognitive control and emotional arousal. The trials aimed to induce anxiety may also modulate theta power.
- Beta rhythm, due to its relation to excitability, arousal and active cognitive processing, is expected to be particularly relevant in the flow and anxiety conditions.
- Alpha activity is expected to be observed primarily during the boredom condition as a result of reduced player activity and task repetition. Following Dietrich's (2003) hypofrontality theory, alpha activity may also be observed in frontal sites during flow episodes.

Finally, a relationship between alpha asymmetry and the different game design considerations (flow, anxiety and boredom) will be examined. Approach and withdrawal fit well with the manipulation of the rewards in the Tower Defense VR game (see also, Gray, 1994). Based on previous literature, it is expected that the flow rounds will elicit higher alpha activity in the right-frontal sites relative to the left, as a means of an approach behavior. In contrast, boredom rounds are expected to elicit higher left-hemispheric alpha activity relative to the right, as a result of low reward and repetition, thereby indicating withdrawal. However, this may not be necessarily true for the anxiety rounds, as the high fiscal incentive, during those rounds, may engender approach motivation. Noteworthily, our previous study demonstrated a physiological overlap between anxiety and flow rounds, which may be similarly shown in the present study, given that no modification was made to the game in between the two studies.

6.2 Methods

6.2.1 Participants

Thirty-two volunteers took part in the study (mean age = 22.13, *SD* = 5.4) the majority of whom were students at Bournemouth University, Bournemouth, United Kingdom, and employees in Sony Interactive Entertainment in London, UK. The sample comprised 26 males and 6 females with an average video games experience of 12 years (*SD* = 4.77) and an average starting age of playing games at 8.6 (*SD* = 3.42) years old. The study was approved by the Bournemouth University Research Ethics Committee (ref: 17333).

6.2.2 Materials: Test Game

The same version of the game from **Chapter 5** was maintained for this study (see **Materials: Test Game**). In debriefing, no participants reported to have realized the experimental manipulation, including the dynamic difficulty adjustment and the impossibility of eliminating the bosses allotted to the anxiety and boredom conditions. Because the same version of the game was used, the study maintained its repeated-measures design (see also Chapter 5, **Experimental** design).

6.2.3 Data Acquisition

Electroencephalograms were recorded using the BrainAmp DC 32-channel amplifier (BrainProducts GmbH, Gilching, Germany) and the proprietary software BrainVision Recorder (v. 1.21.0102) at a sampling rate of 1 kHz. The BrainVision Recorder did not support control over the recordings through the keyboard, as was possible with the Student Lab 4.0 (see Chapter 5, **Data Acquisition**). Hence, the custom application on the receiving desktop that listened to incoming messages (see Chapter 3, **Data acquisition**) utilized the native input detection to capture mouse clicks when a window starting with the name "Record" was active – corresponding to the start of the BrainVision Recorder window name. To amend superfluous registration of click messages, Alt + Tab key combination was used to switch application windows. The purpose of detecting these events was to register the Coordinated

Universal Time (UTC) timestamp the recording was started for synchronization with the game events.

The EEG electrodes (Ag/AgCl) were attached on the actiCAP Electrode Cap (EASYCAP GbmH, Herrsching, Germany) and re-referenced offline to the right mastoid M2 and grounded to AFz. The thirty-two channels that were used for the recordings were fixed to the scalp with a conductive gel to minimize impedance. The channels were fairly distributed across the scalp, including Fp1, Fp2, AF3, AF4, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, C5, C3, Cz, C4, C6, TP7, CP3, CP4, TP8, Pz, PO7, PO3, PO4, PO8, O1 and O2 (**Figure 22**).

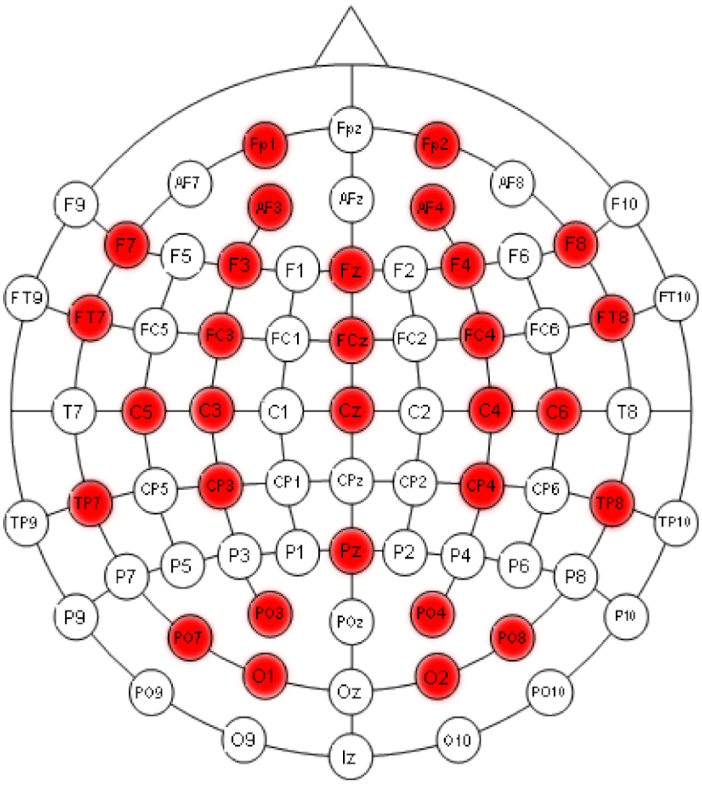


Figure 22: Illustration of the EEG electrode arrangement used in the study (channels highlighted in red). The channel AFz was used for grounding.

6.2.4 Procedure

The experiment took place in Bournemouth University's EEG lab. Participants were first screened for any allergy to alcohol or salt, both of which were subsequently used during the experiment. They were first asked to fill out a consent form, a games experience questionnaire and the Big Five Inventory (BFI). Each participant was trained in the game, regardless of their previous experience with video games, virtual reality or familiarity with tower defense games. Once participants reported feeling comfortable with handling the game, the EEG electrodes were carefully attached.

Then, a resting baseline of 60 seconds was recorded whilst the head-mounted display of PlayStation VR was donned, displaying the main menu of the console. Finally, the recordings were started, and the participants initialized the game. At the end of the game, participants were asked to fill out the Flow State Scale - 2 (FSS-2) questionnaire. The overall experiment, including setup, concluded after 90-120 minutes. A game voucher from a selection of PlayStation 4 games was given as compensation for participation. In line with the previous experiment, participants were not made aware that the game featured indomitable rounds and content that adapted to their performance before the experiment. This was done during the debriefing (i.e., the conclusion of the individual experimental session after all self-reports were filled out).

6.2.5 Statistical Analysis

A similar analysis to the equivalent statistical procedure of the previous study was adopted here (see section Chapter 5, **Statistical Analysis**). Each trial was segmented into three windows, whereby only the second window was used for the final analysis. A 3x7 within-subjects design with Condition (flow, anxiety, boredom) x Region of Interest (ROI; frontal-left, frontal-right, midline, central-left, central-right, posterior-left, posterior-right) was employed. Channels within each ROI were defined as follows (**Figure 23**): *frontal-left* (Fp1, AF3, F7, F3, FT7, FC3); *frontal-right* (Fp2, AF4, F8, F4, FT8, FC4); *midline* (Fz, FCz, Cz, Pz); *central-left* (C5, C3, TP7, CP3); *central-right* (C6, C4, TP8, CP4); *posterior-left* (PO7, PO3, O1); *posterior-right* (PO8, PO4, O2).

The ROIs follow a segmentation similar to that of Aftanas and Golocheikine (2001), except for the midline, which, due to the limited number of electrodes (four in this study versus eight in their study), formed a single cluster. However, several works have included frontal-to-parietal midline electrodes in their midline cluster analyses (e.g., Babiloni et al., 2004; Barry & De Blasio, 2017). Frontal-midline (Fz) has been associated with motor planning and parietal-midline (Pz) with perception; the pair Fz/Pz can present coherence, or functional connectivity (Walker et al., 2007). Flow's dimension of "merging of action and awareness" (Csikszentmihalyi, 1990) is embedded in this grouping.

The dependent variables were the relative power of three bands, namely theta, alpha and beta. The analysis was performed separately for each band. One participant (ID = 4) was dropped from the analyses as their trial data were corrupted. Hence, the final tested sample size was N = 31. For the self-report measure analyses, all the participants were considered (N = 32), except in the case of self-report class generation (i.e., low- and high-flow groups), which adhered to N = 31.

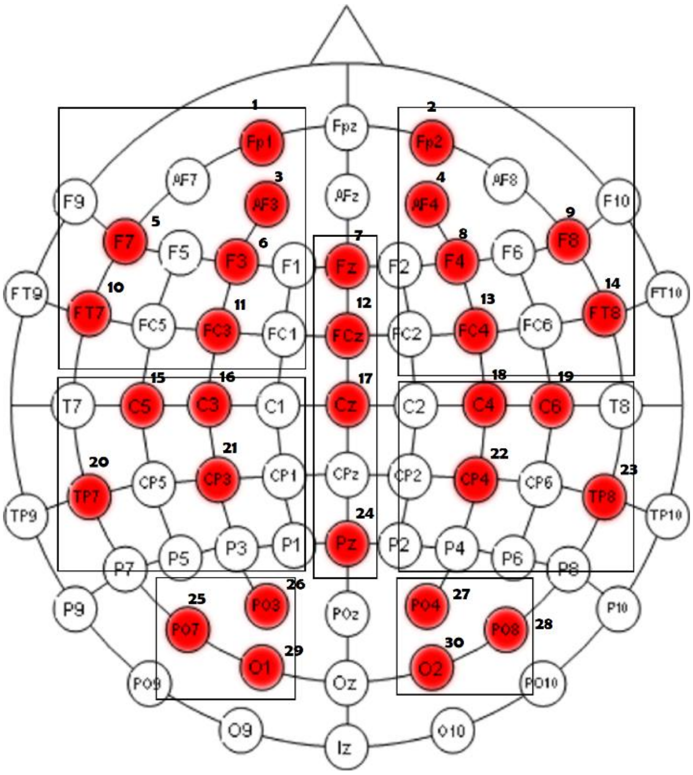


Figure 23: Electrode clusters used in the study. The nomenclature of each region of interest is given accordingly, and separated into left-frontal, right-frontal, midline, left-central, right-central, posterior-left (parietal and occipital sites) and posterior-right.

Alpha asymmetry was measured using the electrode sites Fp2-Fp1, AF4-AF3, F4-F3, and F8-F7, which are some of the most common pairs used in similar studies (Briesemeister et al., 2013; Fairclough et al., 2013). The absolute alpha power was extracted for each channel, using the approach detailed in **Electroencephalography Preprocessing and Features** section, and then the natural logarithm of the right channel was subtracted from the natural logarithm of the topographically equivalent left channel (Allen et al., 2004). Values approaching zero indicate symmetrical activity (Coan & Allen, 2004).

All post-hoc tests were adjusted with the Bonferroni method. Data normality was investigated using the Lilliefors normality test. Significance level was set to α = .05. Additionally, Greenhouse-Geisser correction was applied, and subsequently reported, where necessary. The statistical analyses were carried out using SPSS v23.0 (IBM Corporation, Armonk, NY). For data visualization, violin plots were constructed with a tool developed for MATLAB [7] and topographic maps for band activity were visualized with the Topographic EEG/MEG Tool on Mathworks File Exchange (Martínez-Cagigal, 2020).

6.2.6 Feature Extraction and Selection

Based on the statistical findings, we opted for reducing the number of channels needed for classifying the states of interest. The motivation behind this is to reduce the manufacturing costs that a real-time system would require and to omit redundant or less relevant channels (see e.g., Arvaneh et al., 2010; Lal et al., 2004). The trials were partitioned into seven windows to maintain comparable outcomes with the previous study (see sections **Data Segmentation** and **Data partitions and window length**).

Electrode minimization. Unlike the ECG and EOG features in chapter **5**, which were one-dimensional, EEG features are two-dimensional, involving a frequency band of interest as well as a corresponding channel. This is usually approached by constructing a feature space consisting of band × channel number of features, followed by a dimensionality reduction algorithm (e.g., Berta et al., 2013). However, common algorithms, such as principal component analysis, obtain the final components from the linearly combined contribution of the original variables (Francis & Wills, 1999). The resulting components are essentially composite variables (Song et al., 2013) that can be difficult to interpret (Rousson & Gasser, 2004). Given

[7] https://github.com/bastibe/Violinplot-Matlab

that the goal is to reduce the number of channels – thus, the number of the original variables required for measuring flow, anxiety and boredom, this approach was not considered suitable.

Feature selection methods. Other feature selection options can be classified into wrapper (or encapsulation) and filter methods (for a review, see Sánchez-Maroño et al., 2007). Feature selection that acts as a pre-processing step to classification can be thought of as a filter method (Belanche & González, 2011). The advantage is that the selection of the features is independent from the classifier, thereby potentially fostering the generalizability of the classification results. In contrast, wrapper methods select features through an iterative classification process; thus, the performance of the classification is used as an evaluation metric for the selection of the features (Wu & Wang, 2015). Albeit wrapper methods may be more robust to finding the optimal subset of features, their computational cost and lack of generalizable solutions are oftentimes a deterrent (Wang & Liu, 2016).

One such filter method is the minimum redundancy maximum relevance (MRMR) algorithm (Peng et al., 2005). This algorithm ranks the input variables based on their correlation with the target class (relevance), whilst aiming to minimize the correlation amongst each other (redundancy). Relevance is calculated using the F-statistic for continuous variables or mutual information for categorical variables, based on their joint probability distribution (Peng et al., 2005). On the other hand, redundancy is computed using the Pearson correlation coefficient for continuous variables or mutual information for categorical ones (Ding & Peng, 2005). MRMR has been used in EEG channel selection (e.g., Secerbegovic et al., 2017) and in the specific study of flow in video games (e.g., Mallapragada, 2018). The algorithm is considered one of the most widely applied and optimal feature selection methods (Ramírez-Gallego et al., 2017), with successful applications in EEG data (e.g., Liu et al., 2016; Wang et al., 2011), and hence chosen for this study.

Noteworthily, MRMR simply ranks variables; it does not eliminate unnecessary, noisy or redundant ones. Hence, further computations are needed to estimate the final *N* variables to be kept, via approaches referred to as hybrid, as they combine filter and wrapper methods (Jović et al., 2015). Examples of such methods include the well-known sequential feature selection algorithm, an iterative cross-validation with N variables at each iteration until classification performance no longer improves (e.g., Wu & Wang, 2015), or three-step approaches with k-means clustering as an intermediary step (e.g., Bins & Draper, 2001).

Using MRMR in this study. In this study, there are two flow definitions that we have explored, i.e., (a) the original game-design labels and (b) the flow experience as reported by the participants on the FSS-2 questionnaire. If the pursuit of an optimal subset of channels relies on the original game-design labels, we may compromise the ecological validity of the findings, thereby sacrificing their generalizability to other video games and applications. This is suggested by the poor performance of the classifiers without correcting for the latent inter-personal variables identified in the previous study (**Diagnosing the findings using an unsupervised learner**). Ergo, the optimal set of channels should also be derived from a more standardized definition of flow, which is the Flow State Scale. In doing so, we can explore the applicability of the selected channels whilst attempting to separate the classes of flow, anxiety, and boredom from our original design labels. In this study, we consider the intersection of the two channel selection methods using either class definition to best represent the flow experience.

The procedure using MRMR analysis was thus carried out using the groups (classes) of participants with low- and high-flow reported scores (a) and the game-design classes of flow, anxiety and boredom (b). The analysis was performed independently for each band, i.e., beta, theta, and alpha, each yielding its associated weights provided by the MRMR method. Thus, the 30 channels are ranked separately per frequency band. Next, the output weights corresponding to each band (3 bands × 30 weights per channel) were averaged separately for each channel, yielding 30 values, and later sorted in descending order that was indicative of each channel's importance.

For the classification approach, channels were progressively incorporated, based on their ranked order, as inputs to RBF-SVM classifiers with fixed parameters[6]. Three new features were added at each step of this process, whilst computing the accuracy of the 10-fold cross-validated classifier at each step. To clarify the process, the first iteration included theta, alpha and beta relative powers of the top contributing channel; at the second iteration, the theta, alpha and beta relative powers of the second-top contributing channel were added, and so on. Training data at each fold iteration were oversampled using SMOTE to balance the class sizes (refer to Chapter 3, **Balancing class sizes**). The acquired accuracy rates from all folds of consecutive classifiers were finally submitted to a paired-samples t-test to compare them statistically (α = .05).

6.3 Results

6.3.1 Statistical results

6.3.1.1 Self-report measures

Overall, the participants' perceived difficulty of the game concentrated around "Hard". 59.4% of the participants reported that they thought TD-VR was hard, 37.5% that it was of moderate difficulty and only 3.1% (equivalent to one participant) reported that it was easy, in response to the question "How difficult did you find the game you have just played?". A stepwise multiple linear regression between replay intention and the flow dimensions (independent variable) indicated that flow's self-reported autotelicity significantly predicted replay intention (i.e., "How likely would you play this game again in your free time?"), β = 1.21 (CI 95% [.821, 1.6]), R^2 = 57.4%, $p < .001$.

Participants were also asked after the game to estimate the time they thought had passed whilst they were playing, i.e., "How much time do you think passed since the beginning of the game?". Unsurprisingly, flow's dimension of time distortion presented a significant negative relationship, β = -4.46 (CI 95% [-8.32, -.688]), R^2 = 16.3%, $p < .05$. This suggests that the underestimation of time was higher as the self-reported time distortion increased. Loss of time perception during video game play ("When you play video games, do you lose track of time?") significantly correlated with enjoyment derived from video games ("Do you feel enjoyment when you play video games?"), $\chi^2(6) = 17.9$, $p < .01$. Flow's time distortion was also negatively associated with age, β = -0.63 (CI 95% [-.114, -.011]), R^2 = 16.8%, $p < .05$; self-reported time distortion decreased as age increased.

6.3.1.2 Comparison with Studies I and II

A one-way ANOVA was carried out to explore differences in the self-reported flow dimensions (FSS-2), personality scores (BFI) and other variables. All data from all studies were added as dependent variables. The following variables indicated significant differences: flow's autotelicity, $F(2, 82) = 6.094$, $p < .01$, flow's skills-challenge balance, $F(2, 82) = 3.438$, $p < .05$, personality trait of conscientiousness, $F(2, 82) = 7.284$, $p < .01$, Tower Defense VR final score, $F(2, 82) = 89.752$, $p < .001$, age, $F(2, 82) = 15.9$, $p < .001$, and years of games experience, $F(2, 82) = 6.917$, $p < .01$. Differences between the first and second study will not be detailed here, as they were mentioned in the previous chapter (see **Comparison with Study I**).

Pairwise comparisons with Bonferroni correction showed that autotelicity was significantly lower in this study than the first ($p = .036$). Significantly lower ratings were also observed for the dimension of skills-challenge balance ($p = .047$). The personality trait conscientiousness was found to be significantly lower from the second study ($p < .01$) as well as the years of experience ($p < .01$). The average age was significantly lower from the previous studies (first study: $p < .01$; second study: $p < .001$), indicating that the present study comprised the youngest sample across the three studies. Finally, the players in the present study performed better as evidenced by their Tower Defense VR score, which was significantly higher than the previous study ($p < .05$).

6.3.1.3 EEG band activity

The relative power from three frequency bands (details of processing in section Chapter 3, **Electroencephalography Preprocessing and Features**) was submitted to a repeated-measures ANOVA as a dependent variable. The model design included the factors Condition (flow, anxiety, boredom) and ROI (frontal-left, frontal-right, midline, central-left, central-right, posterior-left, posterior-right). Three separate analyses were carried out for each band (theta, alpha and beta). All topographic maps shown below are normalized by the local maximum power found at each band rather than the maximum value from the aggregated conditions.

Theta band. The main effects of ROI, $F(2.7, 82.10) = 41.32$, $\eta_p^2 = .579$, $p < .001$, Condition, $F(2, 60) = 9.35$, $\eta_p^2 = .238$, $p < .001$, and the interaction ROI × Condition, $F(4.58, 137.25) = 8.32$, $\eta_p^2 =$

.217, $p < .001$, were significant. For the ROIs, theta power increases were overall observed in the left-frontal (28.77 ± .876, $p < .001$), right-frontal (28.11 ± .971, $p < .001$) and midline (31.56 ± .973, $p < .001$), yielding significant differences with the remaining ROIs (**Figure 24**).

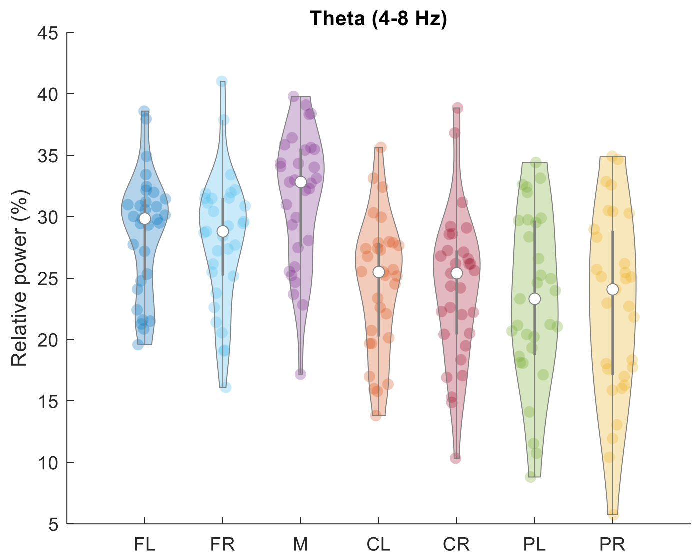


Figure 24: Relative theta power across seven regions of interest (ROIs). Theta power was most pronounced in the frontal-left, frontal-right and midline. FL = Frontal-Left, FR = Frontal-Right, M = midline, CL = Central-Left, CR = Central-Right, PL = Posterior-Left, PR = Posterior-Right.

Follow-up pairwise comparisons of the Condition factor indicated that relative theta power was significantly higher during the flow (26.57 ± .967, $p < .01$) and anxiety (26.59 ± .98, $p < .01$) than during boredom (25.37 ± 1.01). To obtain a better understanding of the Condition × ROI interaction, post-hoc tests were carried out and showed that flow and anxiety at left-frontal, right-frontal, midline and left-central ROIs modulated significantly higher theta power than boredom, $0.001 < p < .05$ (**Figure 25**). Right-central theta amplitude was also significantly higher in the anxiety versus boredom pair ($p < .05$), but not flow ($p = .134$).

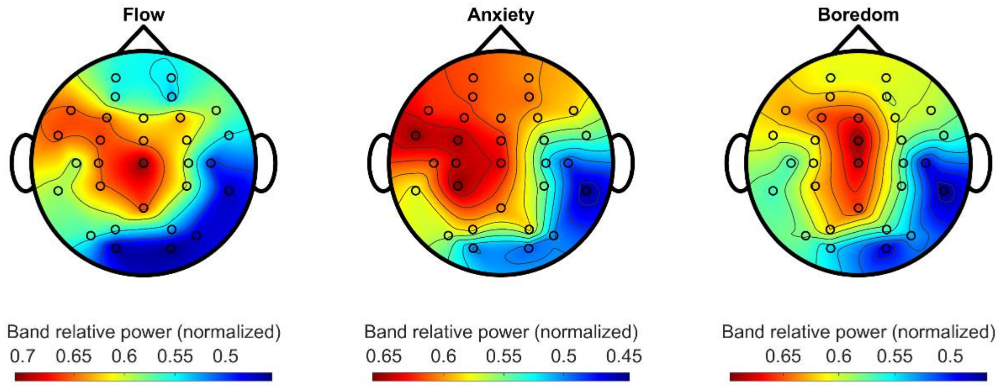


Figure 25: Average theta power distribution across the locations on the scalp visualized through the normalized relative band power from all participants. Theta power was significantly more pronounced in the flow and anxiety conditions in the bilateral-frontal and left-central sites. Theta power was also stronger during the anxiety condition in the right-central sites compared to boredom.

Alpha band. In the alpha range, the omnibus tests were significant for ROI, $F(2.42, 72.48) = 6.42$, $\eta_p^2 = .176$, $p < .01$, Condition, $F(1.51, 45.31) = 13.05$, $\eta_p^2 = .303$, $p < .001$ and the interaction ROI × Condition with a weak effect size, $F(4.99, 149.81) = 3.05$, $\eta_p^2 = .092$, $p < .05$. Alpha rhythm was significantly more pronounced in the midline and left-central sites compared to the bilateral frontal sites, $0.01 < p < .01$ (**Figure 26**). The left-posterior sites were also modulated by higher alpha activity compared to the right-frontal sites ($p < .01$). The post-hoc tests for the effect of Condition showed that boredom was characterized by higher alpha power than flow and anxiety ($p < .01$ for both comparisons).

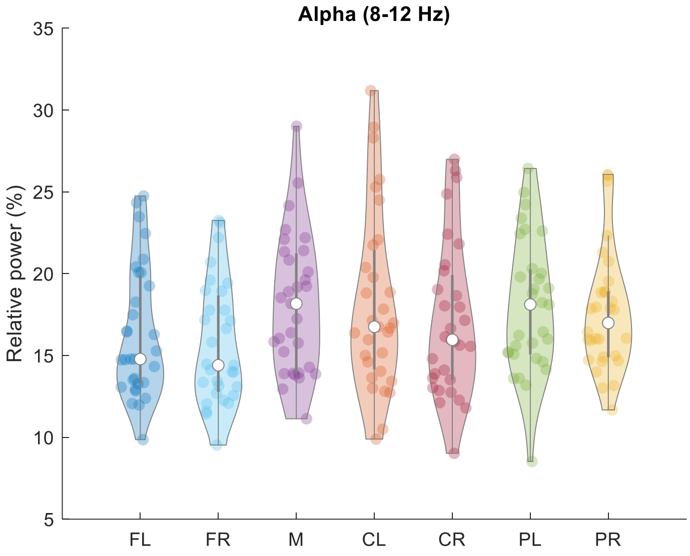


Figure 26: Relative alpha power across seven regions of interest (ROIs). Alpha power was best observed in the midline, central-left and posterior-left sites. FL = Frontal-Left, FR = Frontal-Right, M = midline, CL = Central-Left, CR = Central-Right, PL = Posterior-Left, PR = Posterior-Right.

The follow-up tests for the Condition × ROI interaction indicated that boredom was characterized by significantly higher alpha activity in the midline, left-central, right-central, and left-posterior sites compared to flow and anxiety (.05 < p < .001). In the right-posterior sites, boredom modulated significantly higher alpha activity than flow (p < .01), but not anxiety (p = .214). Furthermore, anxiety modulated higher alpha activity in the midline, in comparison to flow, p < .05 (**Figure 27**).

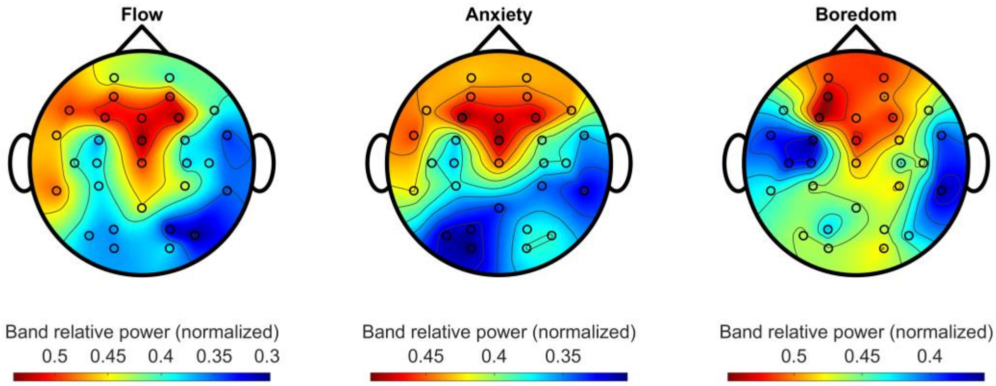


Figure 27: Average alpha power distribution across the locations on the scalp visualized through the normalized relative band power from all participants. Activity in the alpha band dominated during boredom over the midline, left-central, right-central and left-posterior sites. Furthermore, anxiety induced stronger alpha activity compared to flow in the midline.

Beta band. For the beta band, the effect of ROI was significant, $F(3.12, 93.7) = 33.24$, $p < .001$, $\eta_p{}^2 = .526$. Interestingly, the power in this range increased proportionally to the posteriority of the ROI it was extracted from, albeit not consistently (**Figure 28**). The lowest beta was observed in the midline (Fz, FCz, Cz and Pz) and was statistically significantly lower than any other ROI ($p < .01$). In contrast, beta activity at the right-central and bilateral posterior sites, was statistically significantly higher than left-frontal, right-frontal and midline ($p < .001$).

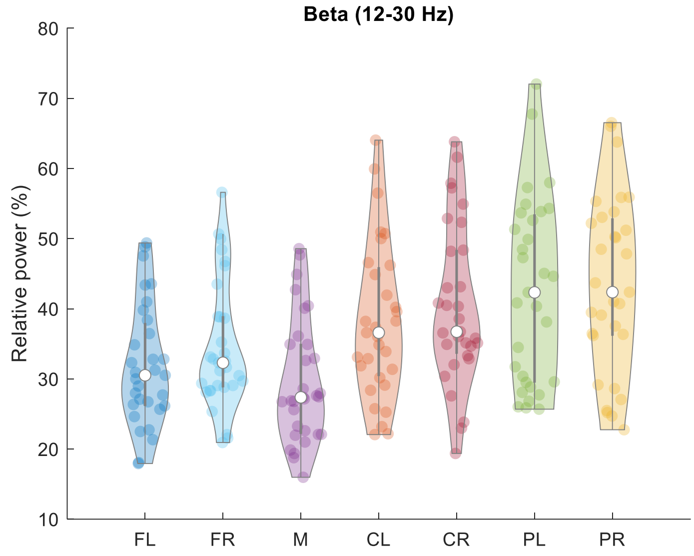


Figure 28: Relative beta power across seven regions of interest (ROIs). A significant modulation of band activity was shown to progressively increase based on the posteriority of each ROI. FL = Frontal-Left, FR = Frontal-Right, M = midline, CL = Central-Left, CR = Central-Right, PL = Posterior-Left, PR = Posterior-Right.

The main effect of Condition was also significant, $F(2, 60) = 3.80$, $\eta_p^2 = .112$, $p = .028$, as well as the interaction of Condition × ROI, $F(5.87, 176.16) = 4.89$, $\eta_p^2 = .140$, $p < .001$. Post-hoc tests indicated significantly higher right-central beta during flow (40.87 ± 1.95, $p < .05$) and anxiety (41.07 ± 2.20, $p < .01$) than during boredom (39.12 ± 2.07). The same trend was observed at the posterior sites where both flow and anxiety modulated significantly higher bilateral beta activity than boredom, $.01 < p < .05$, in all four comparisons (**Figure 29**).

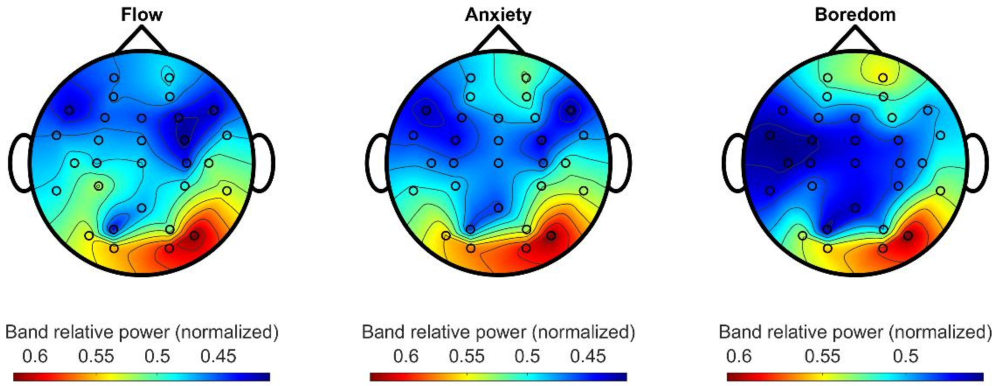


Figure 29: Average beta power distribution across the locations on the scalp visualized through the normalized relative band power from all participants. Right-central (C4, C6, CP4, TP8) and bilateral-posterior sites (PO7, PO3, PO4, PO8, O1, O2) produced stronger beta activity during flow and anxiety compared to boredom.

6.3.1.4 Alpha asymmetry

To address the specific case of alpha asymmetry, a preliminary 3 × 2 repeated-measures ANOVA was carried out with Condition (flow, anxiety, boredom) and ROI (left vs right channels) in its terms. The dependent variables were the averaged absolute power estimates of the alpha band for each participant at each condition across four pairs of channels, i.e., Fp2-Fp1, AF4-AF3, F4-F3, and F8-F7, collapsed into left (Fp1, AF3, F3, F7) and right (Fp2, AF4, F4, F8) as dependent variables. The data were extracted from each second window of each round in the game. The main effect of Condition was not significant, $F(1.55, 46.36) = 1.58$, $p = .215$. However, the main effect of ROI was significant, $F(1, 30) = 8.53$, $\eta_p^2 = .221$, $p < .01$, where the left hemisphere was characterized by significantly higher absolute alpha (6.41 ± .65) than the right hemisphere (5.42 ± .512 , $p < .01$). These results allude to an overall withdrawal tendency, when the experimental condition is not factored in.

Further, a second analysis (3 × 4 repeated-measures ANOVA) was carried out, testing for the main effects of Condition for each electrode pair (Fp2-Fp1, AF4-AF3, F4-F3, and F8-F7) using the difference in the log absolute power of the right channel subtracted from the contralateral channel (Allen et al., 2004). The analysis showed a significant main effect of Electrode Pair,

$F(1.49, 44.61) = 32.11$, $\eta_p^2 = .517$, $p < .001$, but the effect of Condition failed to attain significance ($p = .788$). Follow-up analysis on the Electrode Pair showed that the FAA index at the F8-F7 pair was significantly lower than all other pairs ($p < .01$ for all comparisons; **Figure 30**). FAA at the Fp2-Fp1, AF4-AF3 and F4-F3 paired comparisons was not significantly different, $p > .05$.

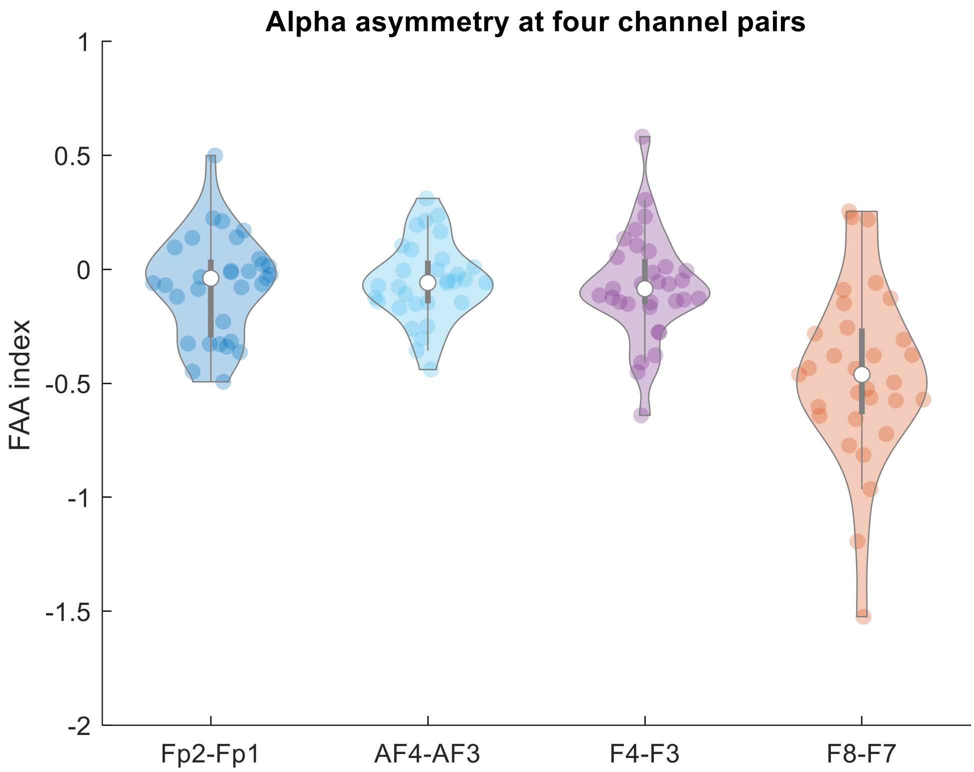


Figure 30: Alpha asymmetry at four channel pairs (Fp2-Fp1, AF4-AF3, F4-F3 and F8-F7) generated by the logarithm of the right channel (even-numbered channels) and subtracted from the logarithm of the left channel (odd-numbered channels, as per the standard nomenclature in the 10-20 international system). FAA index values (*y*-axis) close to zero indicates little to no asymmetry. Only FAA at the pair F8-F7 was significantly lower than the remaining pairs, indicating higher right-hemispheric activation.

Given the absence of a significant effect for condition, the three conditions were further collapsed at each of the four pairs, by averaging (condition-independent), and tested for their relationship with the ratings on each flow dimension and personality scores as reported through the FSS-2 and BFI questionnaires. The results indicated significant positive correlations between the FAA index and neuroticism at the F4-F3 pair, goal clarity at the F8-F7 pair, and autotelicity at the Fp2-Fp1 pair. However, a negative relationship arose for the personality dimension of extraversion, that was the result of a second explanatory model comprising both autotelicity and extraversion (**Table 11**).

Table 11: Results from stepwise linear regressions of the flow dimensions from Flow State Scale – 2 (FSS-2) and Big Five Inventory (BFI) onto alpha asymmetry index from four frontal pairs, averaged from all conditions (flow, anxiety and boredom). The reported findings were all significant at $p < .05$.

FSS-2 Dimension/BFI Trait	Fp2-Fp1	AF4-AF3	F4-F3	F8-F7
Extraversion	$\beta = -.136$, $R^2 = .306$, $p < .05$	$\beta = -.125$, $R^2 = .215$, $p < .01$		
Neuroticism			$\beta = .162$, $R^2 = .18$, $p < .05$	
Goal Clarity				$\beta = .28$, $R^2 = .285$, $p < .01$
Autotelicity	$\beta = .138$, $R^2 = .150$, $p < .05$			

6.3.1.5 Discussion of statistical findings

Self-report measures. An interesting finding emerged from the questionnaires used in the study, where the majority of players reported their self-perceived difficulty of the game as hard. In line with the second study, flow's dimension balance of challenge and skills was rated significantly lower than the first study. Adopting our conclusion from the second study, this is likely because the game's conditions were interwoven into a single game experience, which might have shifted the self-perceived difficulty rating. Noteworthily, the group of participants who rated the game as hard reported significantly higher time distortion in the FSS-2 questionnaire than the group that rated it as medium in difficulty. Furthermore, the performance of the players was statistically significantly higher in this study than any of the previous.

Unsurprisingly, autotelicity presented a significant positive correlation with the players' reported intention to replay the game. This confirms earlier studies that an intrinsically rewarding experience promotes one's willingness to repeat an activity (e.g., Keller et al., 2011b). Similarly, the reported time distortion in FSS-2 was highly correlated with the participants' estimation of the game's duration, indicating that they experienced loss of time perception. This has been shown to have a strong relationship with enjoyment (e.g., Agarwal & Karahanna, 2000).

Electrophysiology. The statistical findings of this study pointed to a few discernible spectral dynamics among the three experimental conditions. More specifically, flow and anxiety were associated with high theta band power in the fronto-central sites, whereas boredom was associated with dominant alpha power spread across multiple regions. Beta power demonstrated a significant increase in the right-central and bilateral posterior sites during flow and anxiety when compared to boredom.

Theta band. Our finding on power increases in the theta range corroborates a successful experimental manipulation, with flow and anxiety having the highest cognitive demand. This is confirmed by several studies, which showed increased frontal theta activity whilst being engaged in cognitive tasks, working memory, attention (Sauseng et al., 2007; Scheeringa et al., 2008), video game play (He et al., 2008) and virtual reality (Baka et al., 2018; Murphy & Higgins, 2019; Rogers et al., 2020). Notably, this finding poses a contradiction with an earlier study by Nacke and Lindley (2011). The experimental conditions used in their study were flow, immersion and boredom in a modified first-person shooter game. The authors found theta activity to be significantly increased in the immersion condition but not in the flow condition, which they specifically distinguished in their experimental design. They further postulated that theta may have underlay the architectural complexity featured in the immersion condition and suggested a possible link with spatial navigation.

However, our study did not feature spatial navigation to justify a similar conclusion. Instead, our results are akin to Ewing and colleagues (2016), who identified a quadratic trend in the theta power under the effect of game difficulty. The authors surmised that game difficulty modulated activity in this band as a result of cognitive load. Further, Katahira and colleagues (2018) identified high theta activity during flow and overload conditions, but not during boredom, which the authors deduced to be a neural signature of cognitive control and working memory load. A similar, yet slightly divergent, finding is that of Yelamanchili (2018), who also found increased frontal theta activity during a Tetris condition aimed to elicit flow. However, the author found significant theta power differences between the flow and anxiety conditions, with flow modulating higher power. This was not confirmed in the present study. Notably, the theta band power in our study was stronger in the flow and anxiety conditions in a wider network, including frontal and central sites; similar findings for flow are reported in the literature (Derbali & Frasson, 2010; Metin et al., 2017).

The distinction between anxiety and boredom conditions in the theta band can be attributed to the presence of an incentive about which the participants were informed during the training session. The right-frontal theta activity in particular may also relate to reward processing (Christie & Tata, 2009) and emotional activation (Grissmann et al., 2017; Stenberg, 1992). In the case of reward processing and in agreement with our experimental manipulation, the magnitude of the rewards in TD-VR was maximized during the flow and anxiety conditions. Related findings have shown fronto-central theta power to increase during high cognitive load in presence of a high incentive (Fairclough et al., 2013), which appears to be sustained even though the reward delivery is omitted (Hamel et al., 2018). In the case of emotional activation, Stenberg (1992) found that a task of emotional imagery with pleasant and unpleasant memories accentuated right-frontal theta activity than a neutral control task. The author suggested theta activity to be an indicator of emotional arousal. The same author further claimed that theta may be a more reliable biomarker for emotional arousal than alpha-band oscillations.

This is corroborated by the more recent study of Reuderink et al. (2013), who found increased right-frontal theta activity during periods of experimentally induced frustration in a Pacman game. The authors found a weak, but significant negative correlation between fronto-central theta activity and self-reported valence (also, Grissmann et al., 2017; Stenberg, 1992). Since the terms anxiety and frustration have been used interchangeably (e.g., Craveirinha & Roque, 2010; Jin, 2012; Knox et al., 2011), it is not a stretch to suggest that our findings are complementary to Reuderink and colleagues'. However, given this evidence, we are led to question whether the flow design of TD-VR was associated with negative emotional valence. Even if theta power increases could potentially allude to the existence of negative valence, Roest and Bakkes (2015) argued that this does not automatically compromise the engaging nature of a game. Perhaps, as argued by different authors, flow may not be uniquely linked to positive emotions during the experience (Csikszentmihalyi, 1990, 1996; Ghani & Deshpande, 1994; Walker, 2010). Nevertheless, the self-report measures used in our study cannot determine a possible link with affective valence adequately.

Alpha band. The alpha band was found to be predominantly related to the boredom condition, evincing a widespread power accentuation across the scalp. This was particularly noticeable in the posterior bilateral sites when contrasted with anxiety and flow rounds. Again, the results echo the findings of Ewing and colleagues (2016), who identified a decrease in posterior right alpha power when the game's demand increased in presence of an incentive. In TD-VR boredom design, the demand was high since the players were unable to eliminate the boss, but there was little monetary incentive compared to the anxiety and flow conditions. Thus, whereas the authors found reduction of right alpha power when an incentive was present, an increase of right alpha when the incentive was minimal was found in this study.

Parieto-occipital alpha activity has also been addressed in virtual reality settings. A recent study demonstrated that parieto-occipital alpha activity decreased during VR immersion, but increased during periods of inactivity (Magosso et al., 2019). A confirmatory finding from Wang and colleagues (2018b) suggested that decreases of alpha oscillations are observed in parieto-occipital areas in presence of cognitive and attentional demands. As increases of alpha power were identified in the present study, it can be concluded that cognitive demand during the boredom condition was significantly lower than the remaining conditions. Taken together, these results indicate that posterior alpha increases, as observed in the case of boredom in this study, may relate to phases of distractibility. Yelamanchili (2018) has also provided corroboratory evidence for this claim, where the flow condition was characterized by significantly lower alpha power compared to a resting condition during Tetris playing.

In support of this notion, a recent study identified enhanced alpha power in posterior sites when participants reported mind-wandering during a repetitive Stroop task (Compton et al., 2019). Although measures of mind-wandering were not employed in this work, it is a likely case if we consider that individuals are more prone to engaging in mind wandering when they feel bored (Critcher & Gilovich, 2010; Eastwood et al., 2012). Finally, Lagopoulos and colleagues (2009) also identified large alpha activity in posterior sites that the authors attributed to reduced cognitive processing of sensory integration. To conclude, possible sources of enhanced alpha power across the boredom condition include distractibility, idling and/or mind-wandering.

Beta band. The beta band presented statistically significant changes across the game levels and the experimental manipulation. These changes were observed in two ROIs, the right-side central sites and the bilateral posterior sites. In both ROIs, flow and anxiety trials were characterized by higher beta, relative to the total band power, compared to the boredom trials. Earlier research has shown that beta power increases in the occipital areas relate to visual attention and attentive visual performance, whereas diminution of beta may signify difficulty in sustaining attention (e.g., Gola et al., 2013). Moreover, occipital beta power increases have been observed in Zen and concentration meditation (Dunn et al., 1999; Huang & Lo, 2009); a similar pattern was mentioned earlier for the theta band. Parieto-occipital beta increases are also related to alertness (Kamiński et al., 2012). Taken together, the observed beta spectral dynamics in the parieto-occipital sites of this study may have related to sustained visual attention.

It is interesting that the oscillatory changes in this band were similar in the flow and anxiety conditions (no statistically significant differences observed) versus the boredom condition. These two conditions were mutually characterized by higher reward incentives. However, only flow rounds delivered rewards upon enemy elimination, since anxiety rounds could not be won. Thus, the common distinction made in the literature, that shows differential neural patterns between reward delivery and reward anticipation (e.g., Doñamayor et al., 2012) may explicate the findings here. Specifically, Alicart and colleagues (2015) found that events of reward delivery present a similar oscillatory activity to events of near-misses (almost winning). This means that the flow rounds which consistently delivered rewards when defeating the enemies through a fixed-schedule modulation and the anxiety rounds that paused the bosses' health before elimination (near-miss) evoked similar neural activity.

Perhaps, due to the rewarding mechanics underpinning the flow and anxiety rounds, attention was more sustained during playing.

Alpha asymmetry. The frontal alpha asymmetry index (FAA) failed to provide statistically significant differentiation across the three design conditions. Based on previous work, it was expected that FAA would be, on average, higher during the flow condition, as a marker of approach motivation, and lower during anxiety and boredom, as a marker of withdrawal tendency. Given that anxiety rounds featured high fiscal extrinsic rewards, the direction of this relationship with approach/withdrawal motivational direction was not clear a priori. Nevertheless, the hypotheses were not met.

However, FAA was found to be correlated with two flow dimensions (FSS-2) and two personality traits (BFI). Firstly, a positive correlation was observed between neuroticism and FAA for the pair F4-F3. Withdrawal or avoidance behavior may reflect emotional instability that characterizes higher levels of neuroticism (Huang et al., 2015; Uusberg et al., 2015), and is thus a relevant trait to FAA. Higher levels of neuroticism in this study were associated with higher relative left activation, or an implied approach motivation.

A corroboratory study is that of Farahi and colleagues (2019), wherein neuroticism showed a significant positive correlation with relative left activation at the F4-F3 sites, as well as the F8-F7 pair. However, the authors did not fully explain how these two are positively linked and recommended that subscales of neuroticism may perhaps clarify such a relationship. Similarly, Minnix and Kline (2004) reported a significant positive correlation of FAA with neuroticism at the F4-F3 sites. These findings are consistent with the results in the present work for the pairs F4-F3 and the direction of the relationship (cortical activity lateralized to the left). Minnix and Kline suggested that the positive relationship expresses regulatory mechanisms that individuals with high neuroticism may be employing to suppress negative affect, thereby switching continually between approach and avoidance strategies.

Secondly, a significant negative relationship between FAA and BFI's extraversion emerged at AF4-AF3, but also in a second stepwise linear model that included flow's autotelicity. According to the literature, reduction in FAA implies withdrawal tendency. The relationship of extraversion and FAA is controversial, with studies having reported no relationship (e.g., Hagemann et al., 1999; Schmidtke & Heller, 2004; Wacker et al., 2010) or a relationship opposite to what was observed in the present study (e.g., Wacker, 2018). Though Wacker (2018) found left-sided activation to be related to FAA, it was specifically found during an activity, but not during resting-state asymmetry.

This was a very important point in the author's work, given that previous research on personality focused on resting-state FAA. According to Coan and colleagues (2006)'s *capability model* of frontal EEG asymmetry, personality traits may be indexed via FAA, but it depends on certain conditions to best manifest said traits in a given situation. Thus, the capability model postulates that FAA is dependent on the context for the way a particular personality trait will be manifested. At first glance, this notion extricates us from the responsibility of making directed hypotheses on the FAA's relationship with different personality traits. However, what is meant by the authors is that different activities may have compatible or incompatible properties with a given personality trait, which can result in a shift of FAA index values (positive, negative or asymmetric). In the present case, the dominant right-sided activity for extraversion from the entire game recording is reminiscent of our previous study (see related **Table 8**), where extraversion consistently emerged as being negatively related to several data clusters irrespective of the game's experimental condition. As a result, it was postulated that the game may not have had appropriate game mechanics to satisfy an extravert's incentives for play.

Regarding the specific flow dimensions identified for the pairs Fp2-Fp1 and F8-F7, there is lacking evidence on the relationship between individual flow dimensions with alpha asymmetry. A natural assumption is that the flow dimensions would relate to an approach motivation, given that the task (or, the TD-VR game) is what triggers and helps sustain the flow experience. Increases observed in the reported autotelicity and goal clarity were positively linked to the FAA index, suggesting, according to the literature, an approach motivation. The case of autotelicity is interesting, in that it reflects intrinsic properties of an activity, and by extension intrinsic motivation (Abuhamdeh, 2020; Van den Hout et al., 2018), whereas the approach/withdrawal dichotomous theory has been associated mainly with extrinsic motivation.

Thus, this study is one of the few existing works to have found a link between the autotelic nature of an activity and frontal alpha asymmetry. Further studies are needed to provide a more detailed analysis on the specific intrinsic properties of play (see e.g., Schmid, 2011) associated with alpha asymmetry at the Fp2-Fp1 sites. The second FSS-2 dimension that emerged as being positively related to FAA was goal clarity at the F8-F7 sites. Although existing evidence for such a link is considerably limited, it may be inferred that familiarity with the game's goals entrains a sustained approach motivation toward the engagement with the activity.

Interestingly, this relationship was not shown during gameplay, but by the data averaged from the entire EEG recording per participant. This could be interpreted as autotelicity being more effective as a post-play metric rather than a during-play metric, thereby corroborating proponents of autotelicity viewing it as an outcome of the flow experience (e.g., Quinn, 2005; Rodríguez-Sánchez et al., 2011). Yet, autotelicity conceived as intrinsic motivation, which supposes that it exists prior to the engagement with the activity is an equally plausible case, but this relationship was not tested in the present study.

In summary, the electrophysiological responses were consistent with the previous study's findings using HRV and EBV as physiological measures. The connection of flow and anxiety to cognitive demand was similarly shown in this study through power increases in the theta band at the frontal sites, which also seem to carry assumptions on reward expectancy. This finding enriched our interpretations, considering that HRV and EBV were mostly interpreted as reflecting increased autonomic arousal. In addition, the regions of interest in which specific spectral dynamics of the alpha and beta band suggested the role of attention and distractibility, which was largely consistent with the aim of the TD-VR design as per the designated experimental conditions. In the following section, the results of the channel selection and the subsequent classification outcomes are detailed.

6.3.2 Classification results

6.3.2.1 EEG channel selection

Following the approach documented in **Feature Extraction and Selection**, a classification procedure was employed. Channels were first ranked according to their MRMR scores (**Table 12**), and then the three features of each channel (theta, alpha and beta; see **Table 3**) were progressively added at each iteration in decreasing order of the channels' relevance. Next, the out-of-sample accuracy was calculated at each iteration of the process, incorporating input information from each consecutive channel. In a similar manner to the previous study's exploratory phase (Chapter 5, **Classification results**), the reported performance measures were based off unoptimized classifiers but with fixed parameters (see corresponding footnotes later in the text). The classifiers were optimized only after the electrodes were reduced to a smaller set. The average classification accuracy of the game-design classes is displayed in **Figure 31**. The classification performance did not improve significantly with the addition of new channels (paired-samples t-test, α = .05, Bonferroni-corrected, normality accepted per Lilliefors test) beyond 14 features for both the game-design (flow, anxiety, boredom) and self-report (low-flow and high-flow) classes.

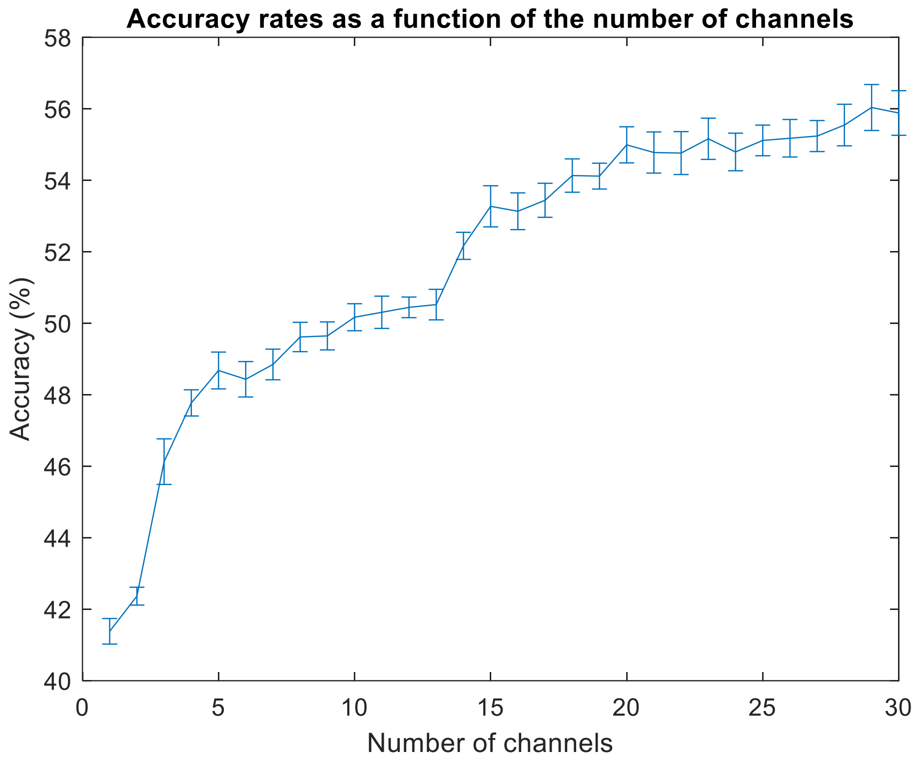


Figure 31: Average classification accuracy of the game-design classes (flow, anxiety and boredom) from cross-validated 10-fold multiclass SVMs with a gaussian kernel [8]. The input features comprised the theta, alpha and beta band power (thus $3 \times n$ channels) and added iteratively based on the importance of each channel as reported by the MRMR algorithm. Error bars are the standard errors of means. Baseline is at 33%, based on the probability of each class's occurrence.

Paired-samples t-tests performed on the fold accuracy rates resulting from the channel subset C_i against the channel subset C_{i-1}, it was observed that no significant differences occurred past 14 channels ($p < .05$). The differences were specifically tested such that the t-statistic indicated an improvement of mean accuracy when testing C_i with its immediately preceding channel subset C_{i-1}. The accuracy rate from C = 14 channels was 52.17% (*SEM* = 0.3785), which is comparable to the accuracy rates obtained in the previous study when using the game design

[8] SVM: Parameter $C = 1$, *Kernel* = RBF (optimal kernel scale determined via a heuristic procedure, see {'kernelScale', 'auto'} name-value pair parameter in Matlab's *fitcecoc* and *templateSVM*).

classes (see **Table 5**), as expected. It can be noticed that the more the channels the higher classification performance.

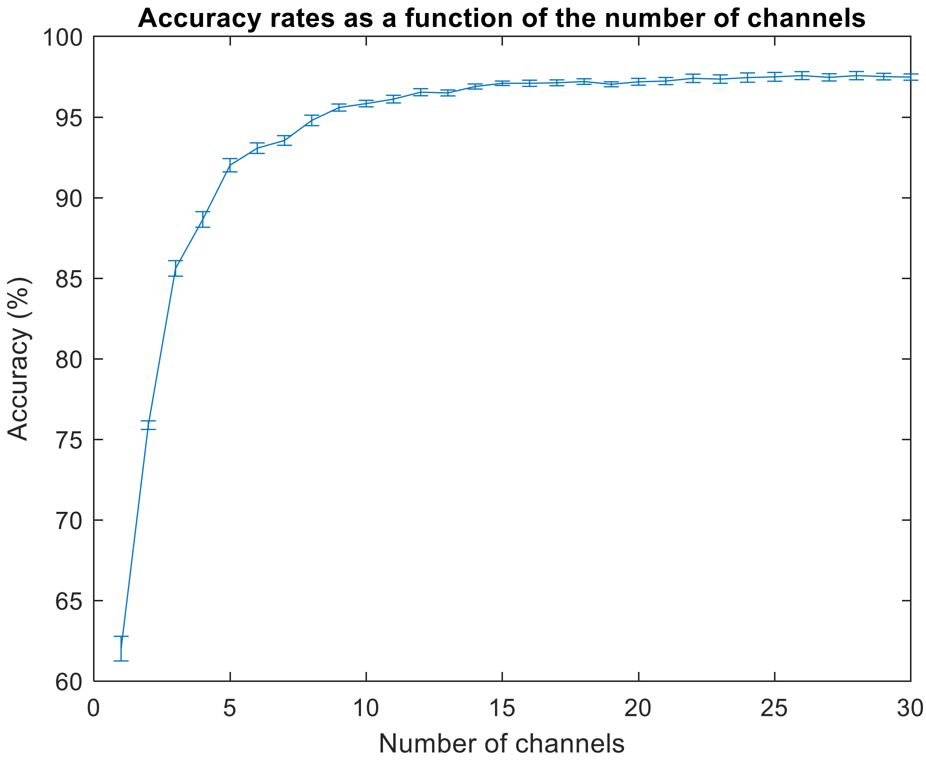


Figure 32: Average classification accuracy of the self-report groups (low- versus high-flow) from cross-validated 10-fold binary SVMs with a gaussian kernel [9]. The input features comprised the theta, alpha and beta band power and added iteratively based on the importance of each channel as reported by the MRMR algorithm. Error bars are the standard errors of the means. Baseline is at 50%.

For the self-report labels, a similar trend was observed (**Figure 32**), where 14 channels were found to be the optimal point, after which no significant improvement in the average fold accuracy was observed. The optimal point of 14 channels was equivalent to an accuracy rate of 96.9% (*SEM* = 0.1583), similarly expected as per the outcomes of the previous study (cf. **Table 6**).

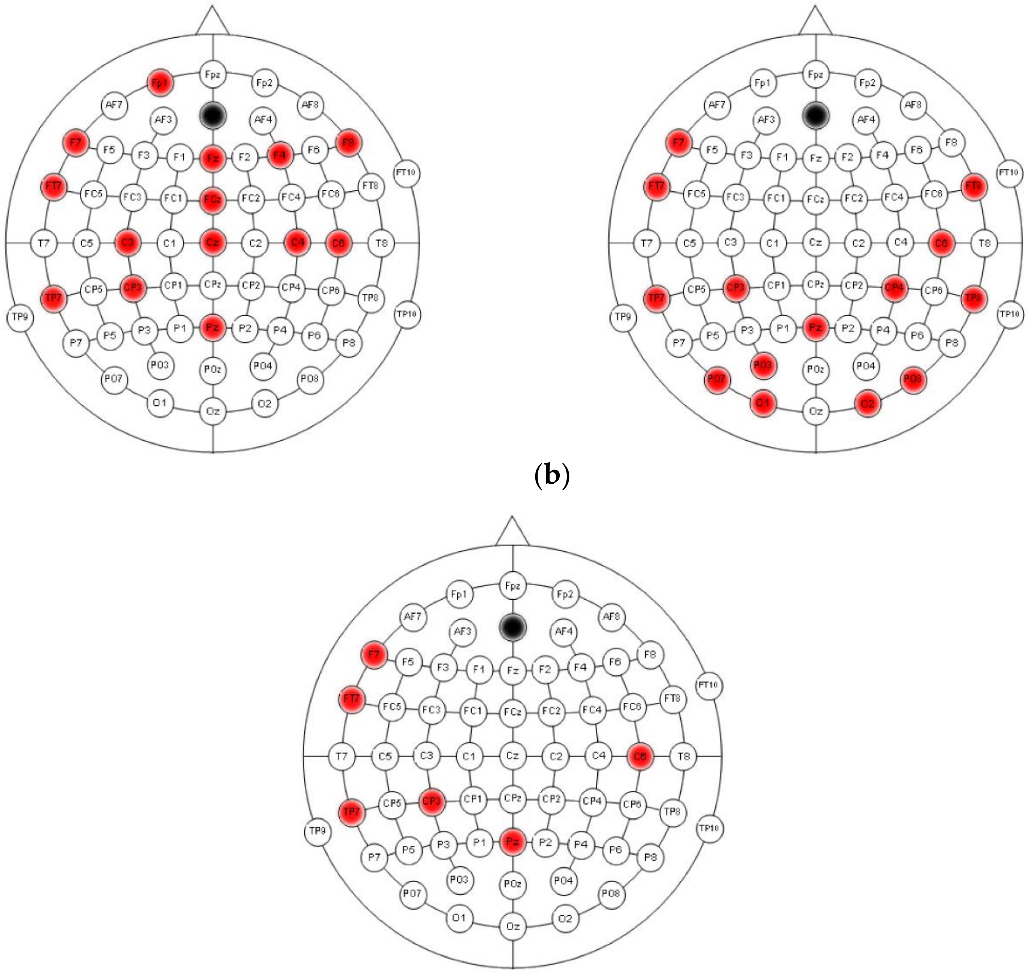
[9] SVM: Parameter $C = 1$, *Kernel* = RBF (optimal kernel scale determined via a heuristic procedure, see {'kernelScale', 'auto'} name-value pair parameter in Matlab's *fitcsvm*).

Table 12: Electrodes sites that best contributed to the resolution of flow, anxiety and boredom (original design classes) and low- and high-flow groups (self-report classes). The channels displayed are ranked in descending order based on their importance weight assigned by the MRMR algorithm. The six highlighted cells and in gray are the channels that were common to both class definitions. Importance weight scores reflect the Mutual Information Quotient (MIQ), which is derived from the division of relevance over redundancy of the Minimum Redundancy Maximum Relevance (MRMR) algorithm.

Final electrode sites			
Original design classes (C_D)	**Importance weight score**	**Self-report classes (C_S)**	**Importance weight score**
Fz	0.0034	TP8	0.0476
F7	0.0030	O1	0.0419
Pz	0.0030	O2	0.0299
C3	0.0027	Pz	0.0286
F8	0.0023	PO7	0.0211
CP3	0.0021	FT7	0.0207
Cz	0.0019	PO3	0.0195
C4	0.0018	C6	0.0185
Fp1	0.0017	FT8	0.0176
C6	0.0017	CP3	0.0152
FT7	0.0015	TP7	0.0151
FCz	0.0015	PO8	0.0138
F4	0.0015	CP4	0.0125
TP7	0.0015	F7	0.0081

The total number of unique channels for both class definitions are thus 14, a reduction of 53% from the original number of channels used in the study. A final consideration regards the overlap between the selected channels of each class definition (game-design and self-report classes).

The game mechanics can be thought of as embodying the antecedent qualities to trigger the flow experience that was subsequently reported by the participants. Thus, it is interesting to investigate whether the intersection of the final channels $C_D \cap C_S$ (**Figure 33**) reflects the optimal responses related to active playing and the self-reported flow experience. The channels that were commonly observed in both class definitions were **F7**, **FT7**, **C6**, **CP3**, **TP7** and **Pz**. We hypothesized that these channels make up a minimal, yet suboptimal, representation of both cognitive states expressed throughout the game and via the psychometric measures. Using these channels, we proceeded with the final step of the study detailed in the next section.



(**a**) (**b**) (**c**)

Figure 33: Channel selection as identified via MRMR and consecutive SVMs. The top-left panel displays the final channels when solving for the game-design classes (a), the top-right panel displays the final channels when solving for the self-report classes (b), and the bottom panel displays the intersection of the two solutions (c).

6.3.2.2 Understanding the neural representation of game design phases

To further understand the latent cognitive factors that may have underpinned experiential variability in the game, the same approach followed in **Chapter 5** for the ECG and EOG recordings was applied to the EEG data. Along the same line, the aim here is to identify the missing cognitive characteristics that help map responses from a minimal set of electrodes to the game design conditions.

First, the relative band powers of interest (i.e., theta, alpha and beta) were extracted specifically from the channels F7, FT7, C6, CP3, TP7 and Pz. The theta/beta ratio was computed for each individual channel across the intersectional channels and the final feature set was complemented with the alpha relative power. The features were submitted to the approach detailed in the **Diagnosing the findings using an unsupervised learner** section.

The results indicated that the FSS-2 dimensions of sense of control and feedback clarity were predicted by the number of player observations comprising each cluster (**Table 13**). These findings reveal that the intersectional channel selection bridged the game-design and self-report classes effectively. Unsurprisingly, due to the Davies-Bouldin index considering two clusters as an optimal *k* for the data segregation of each class, the regressions of the second cluster formed a slope direction opposite to the first cluster.

Based on the findings from each cluster, each observation was multiplied by the corresponding participant rating on the statistically significant FSS-2 dimensions of the regression analyses, depending on the cluster they originated from. The ratings were signed, based on the beta slope direction. The final feature set comprised the weighted observations and the alpha relative power. The addition of the alpha power provided improved classification rates mainly in the anxiety and boredom conditions, with nearly twice the true positives in the prediction rate of the boredom class.

Table 13: Results from separate stepwise linear regressions across the identified fuzzy clusters. For data of each design condition (columns), the proportion of participant observations in each Fuzzy C-means cluster (rows) demonstrated significant associations with flow-related dimensions. Double cells in designated rows indicate that more than one linear model was able to explain the data with a significant R^2 change ($p < .05$).

	Flow	**Anxiety**	**Boredom**
Cluster 1	Sense of Control, $\beta = 23.76$, $R^2 = .173$, $p < .05$	Flow's Sense of Control, $\beta = -9.45$, $R^2 = .144$, $p < .05$	Flow's Sense of Control, $\beta = -11.66$, $R^2 = .183$, $p < .05$
	Sense of Control and Feedback Clarity, $\beta_1 = 33.31$, $\beta_2 = -22.25$, $R^2 = .351$, $p < .01$, R^2 change $= .178$, $p < .05$		
Cluster 2	Sense of Control, $\beta = -23.78$, $R^2 = .173$, $p < .05$	Flow's Sense of Control, $\beta = 9.45$, $R^2 = .144$, $p < .05$	Flow's Sense of Control, $\beta = 11.66$, $R^2 = .183$, $p < .05$
	Sense of Control and Feedback Clarity, $\beta_1 = -33.31$, $\beta_2 = 22.25$, $R^2 = .351$, $p < .01$, R^2 change $= .178$, $p < .05$		

As seen in **Table 14**, the differences in the obtained confusion matrices, when training with only theta/beta ratios versus theta/beta ratios and alpha power, yielded a respective accuracy rate of 38.45% versus 67.15% for boredom.

Table 14: Confusion matrices from an optimized SVM classifier using the game-design classes. Each cell contains the number of observations predicted as belonging to the flow, anxiety and boredom classes through a 10-fold cross-validation procedure. The sum of each row is the original number of observations per class. The top table contained only the theta/beta ratio for the channels F7, FT7, C6, CP3, TP7 and Pz, whereas a clear improvement is observed on the bottom table, when adding the alpha power from the same channels to the input feature set.

Theta/beta ratios at feature level		Predicted Class		
		Flow	Anxiety	Boredom
Actual Class	Flow	3329	90	53
	Anxiety	300	806	413
	Boredom	285	650	584

Theta/beta ratios and alpha power at feature level		Predicted Class		
		Flow	Anxiety	Boredom
Actual Class	Flow	3394	47	31
	Anxiety	220	1056	243
	Boredom	161	338	1020

The results showed a clear mitigation of the misclassification rates when alpha power from the same channels was considered in the input features. Hence, the final classification was performed using theta/beta ratio and alpha power, making up a data set of 6510 samples × 12 features.

Noteworthily, no notes were taken regarding the participants' handedness. This is relevant in that handedness may evince functional lateralization, typically in a contralateral manner (e.g., Gut et al., 2007). The missing channels from the contralateral hemisphere are F8 (corresponding to *F7*), FT8 (*FT7*), C5 (*C6*), CP4 (*CP3*) and TP8 (*TP7*). To test for a possible improvement in classification performance, the same computations, based on the identified predictors, were performed on the feature set, but this time the electrodes from the opposite hemisphere were added to the feature space.

Table 15: Classification performance (%) metrics from three 10-fold cross-validated classifiers using the game-design classes. The feature set comprised three bands (theta, alpha and beta) × six channels (F7, FT7, C6, CP3, TP7 and Pz). Each data sample was scaled by the FSS-2 normalized rating of the corresponding participant on feedback clarity and sense of control. This was based on the conjunctional results of Fuzzy C-means clustering and the linear regressions on the membership count of each participant in each cluster.

Performance	Support Vector Machine	K-Nearest Neighbors	Bagged Ensemble of CART Decision Trees
Accuracy	85.73 (0.6091)	82.21 (0.3740)	83.30 (0.5022)
Specificity	92.90	91.23	91.47
Precision	82.00	77.55	79.07
Recall	81.55	77.43	77.91
F1-score	81.75	77.47	78.43

The resulting classification performance, when using the homologous channels from the opposite hemisphere, presented an improvement, but only in the cases of SVM and k-NN. For the BDT classifier, only precision was slightly improved. In short, combining the participant self-perceived feedback clarity and sense of control with the neural responses from a minimal set of electrodes, we were able to map them more precisely to the game design labels. These findings are discussed in the next section.

Table 16: Classification performance (%) metrics from three 10-fold cross-validated classifiers using the game-design classes. The feature set comprised three bands (theta/beta ratio and alpha relative power) × eleven channels (F7, F8, FT7, FT8, C5, C6, CP3, CP4, TP7, TP8 and Pz). Each data sample was scaled by the FSS-2 normalized rating of the corresponding participant on feedback clarity and sense of control. This was based on the conjunctional results of Fuzzy C-means clustering and the linear regressions on the membership count of each participant in each cluster, but also included the channels corresponding to the opposite hemispheric location.

Performance	Support Vector Machine	K-Nearest Neighbors	Bagged Ensemble of CART Decision Trees
Accuracy	86.96 (0.4849)	83.07 (0.4742)	82.90 (0.5232)
Specificity	93.33	91.47	90.61
Precision	83.86	79.46	79.65
Recall	83.04	79.63	76.74
F1-score	83.43	79.51	77.92

6.4 Discussion

The aim of this study was to identify a set of candidate electrodes for the detection of gameplay experiences, based on electrophysiological responses to the experimental manipulation of three design choices, flow, anxiety and boredom. The statistical results indicated an increase in fronto-central theta activity for flow and anxiety. Beta activity in the right-central and bilateral posterior sites increased during flow and anxiety versus boredom. There was also an increase of relative power in the alpha band observed in the boredom condition in the midline, central and posterior sites.

As per the classification results, we were able to discriminate the three design classes of TD-VR using six channels with the theta/beta ratio and alpha band powers at feature level. The effectiveness of combining fuzzy clustering with self-report ratings to improve inter-class separability was first demonstrated in the previous study (**Diagnosing the findings using an unsupervised learner**) and this is the second study to support the viability of the approach. Interestingly, through this approach, the maximum achieved accuracy rate reached 85.73% in the three-way classification of flow, anxiety and boredom. This suggests that the game design labels can be represented in a compact space, via a subset of the original electrodes, when the self-perceived sense of control and the immediate and unambiguous feedback are considered.

Specific to this study was the conceptualization of flow from two standpoints – the standardized definition, given by the FSS-2 scale according Csikszentmihalyi (1990)'s criteria of flow, and the game design's approach to eliciting flow, monitored via physiological changes. This is a novel approach in the literature that unifies two different class definitions to identify the EEG channels commonly representing the game design and the self-report ratings. The findings, using the averaged weights of the MRMR solutions, converged to six unique channels that contributed the most to the separation of the low- and high-flow groups and the three experimental conditions of the VR game – namely *F7*, *FT7*, *C6*, *CP3*, *TP7* and *Pz*.

These results indicated that the two definitions of flow are not independent from one another and demonstrated a 27.3% overlap of channel contribution ranked by relevance. Although video game playing is associated with widespread stimulation of brain activity (Chanel et al., 2011), the motivation behind choosing the common channels from the two definitions of flow was that the cognitive implications of the original flow definition must similarly be identified in the design of TD-VR. The idea is that different game mechanics employ a variety of cognitive functions which may not be facets of flow per se. Hence, the purpose was to lessen the risk of inheriting cognitive and affective "noise" in the final selection of the electrodes.

For the reasons explained above, it should be noted that those electrodes do not reflect an optimal subset in the present study. The first reason is that the intersection was taken between two seemingly incomparable solutions that were derived from a binary and a multiclass approach. The binary classification approach, which was based on the binary discretization of the self-report ratings, did not automatically assume either anxiety or boredom states in the low-flow group. On the other hand, the mapping of the responses to the designated game conditions in the three-way classification approach was non-stationary, and there existed an observable discrepancy in the classification performance of the two definitions.

The second reason, complementary to the first, is that the intersection of the channels does not best represent either flow definition, as 14 channels were originally identified per flow definition, later reduced to six. The game conditions have had an explicit aim in the experience they were designed to induce. Among those, anxiety and boredom may not be directly inferred from the self-reports. Ideally, this should be addressed in future works via self-reports that specifically target to quantify the three different states of flow, anxiety and boredom. Nevertheless, despite the less reliable performance of the multiclass approach, the overlap was of considerable extent. The final subset may have considered the anxiety and boredom conditions of the game design, given that the sense of control emerged as a latent factor in the regression approach. This factor was commonly found in the present and the previous study (see Chapter 5, **Classification results**) and was indeed a targeted dimension by the game design in TD-VR. The sensors that made up the final set have also had distinct roles in the regions of interest they belonged to, according to the statistical findings (section **EEG band activity**).

A surprising finding was the absence of the Big Five personality dimensions in the final classification procedure. Personality traits have been subject to controversy in the EEG literature. For example, Korjus and co-authors (2015) in their work entitled "*Personality cannot be predicted from the power of resting state EEG*" were unable to verify separable EEG spectra in the classification of the Big Five traits. This finding was challenged recently by Baumgartl and colleagues (2020) who were able to identify extraversion with an accuracy of 60.6%. The authors stated that their success may have related to a finer subdivision of a 99-power band spectrum.

Another recent study was able to discriminate across the five personality traits with individual accuracies ranging from 73% for openness to 86.5% for agreeableness (Klados et al., 2020). The authors mentioned that the resting-state EEG has so far yielded inconclusive results and thus they collected data during an emotion processing task. While the setting of our study was more similar to Klados and colleagues, only the alpha asymmetry showed connections to personality, at the sites Fp2-Fp1 for extraversion and AF4-AF3 for neuroticism. Neither of those electrodes were included in the subset of the channels, which may explain why personality did not emerge as a latent factor in the regression analyses following the clustering approach.

Notably, the sites that best represented the self-report classes (C_S in **Table 12**) present a considerable overlap with those of Mallapragada (2018) who identified a very similar set of channels to relate to the flow experience, namely parieto-occipital, occipital and central sites, with a particularly dense distribution in the parieto-occipital area. We further identified fronto-temporal and temporo-parietal regions to be of importance, indicating a wider distribution of areas to be involved in the self-reported experience of flow. However, although Mallapragada found that only alpha and beta were the most relevant frequency bands, we also identified theta to significantly contribute to the self-reported experience of flow.

The addition of the channels from the opposite hemisphere (C_N = 11) made a subtle improvement to the classification of the electrophysiological responses. However, this improvement, compared to the six electrodes, was not sufficient to justify the inclusion of five additional channels in the final set, highlighting a disadvantageous trade-off. Nevertheless, this omission may compromise the classification of other population samples, and, as such, more studies are needed to clarify the applicability and generalizability of these findings.

Future studies should investigate these channels across different game genres and non-VR applications.

Unfortunately, our study did not contribute to the expansion of knowledge over the brain structures involved in flow. The density of the EEG configuration, based on the number and spatial distribution of the electrodes used, was not high enough to perform source localization reliably (Lantz et al., 2003). Notably, Nuñez Castellar and colleagues (2016) performed EEG source localization in the context of flow using 31 electrodes. However, Michel & Brunet (2019) warned that, subject to electrode positioning, such low-density configurations might result in mislocalization and blurring. Nevertheless, future studies should consider this in settings involving a high-resolution setup, such as Yun et al. (2017) on EEG-based source localization of flow's substrates with 128 electrodes.

To summarize, this study demonstrated that EEG can be successfully combined with VR technologies to decode the flow experience. The obtained results were very similar to those reported for HRV and EBV. However, that study employed both physiological measures to reach similar classification rates and required more information from self-reports (BFI and FSS-2), whereas EEG utilized information from one self-report (FSS-2). Another difference is that the sensors employed for their measurement are spatially disconnected across the body (face, wrists and ankle), whereas EEG was able to provide slightly better classification outcomes from one local source (i.e., the scalp). However, this is not a limitation per se, considering that electrode placement around the mastoids can effectively extract HRV metrics (e.g., Ahn et al., 2019; Von Rosenberg et al., 2016). Similar proposals were made by Zhou et al. (2017) who used a wristband to record electrodermal activity, motion, ambient temperature, and heart rate. In our case, the head is elevated to a source of multiple biometrics that commercial VR can capitalize on, whilst keeping flow-aware technology production-friendly and of low design complexity.

7 General Discussion and Concluding Remarks

This thesis investigated the flow experience in the context of virtual reality game playing. Evidence from previous research has consistently emphasized the importance of flow-like episodes in the enjoyment of activities, including video game playing. At the time of writing, there is a shortage of commercial solutions that readily identify such episodes (see Kerous et al., 2018; Rashid et al., 2020). The motivation of this work, therefore, is that flow-aware technologies are still lacking, not only from an academic standpoint, but also as consumer-grade tools.

While there already are advancements in this area, literature leans toward concerns, study limitations and risks, which nevertheless aim to stimulate follow-up research. These include technical considerations, such as signal quality, or theoretical, such as the reliability of measurements with respect to the mental states the devices are purported to track (e.g., Lotte et al., 2018). Another deterrent, similar in nature, is that research is conducted with academic standards and equipment, which are rarely applicable or accessible to the industry. Much of the research in the field remains exploratory, which can translate to a significant burden on the industry. Yet, the industry may be reluctant in investing in such technologies, due to the inconsistency in methodological and theoretical considerations that abound existing research (e.g., Lin & Do, 2020; Nijholt et al., 2009; Parsons et al., 2020). As such, the complexity of flow as a theoretical and physiological construct highlights that identifying flow episodes is not merely limited by hardware considerations, but perhaps, primarily, by the interpretation of the contextual data.

To alleviate these drawbacks, this thesis devised a new approach to assess the effectiveness of a game design setting in eliciting flow. To this end, the information derived from the subjective experience of flow, as reported by video game players, was combined with their physiological responses to various experimental manipulations of a virtual reality game. The idea behind this combined approach was to bridge the subjective (self-reports) and objective (physiology) data to converge into a more robust representation of flow, that is nonetheless still challenging to articulate. It is hoped that this work will pave the way for new technologies that will advance the state-of-the-art methods for detecting the flow experience and other mental states typically accompanying immersive game play, while also enhance the practice of game design assessment.

7.1 Summary of findings

In a spate of three experiments, organized in the chapters 4, 5 and 6, flow and the neighboring states of anxiety and boredom were investigated through manipulations of the game's design. The first experiment (**Chapter 4**) was a pilot study that sought to confirm whether a custom-designed game for virtual reality could effectively elicit flow. The research goal (objective 1; see **Research Goals**) addressed in this study was to understand the relationship between personality measures and the subjective flow experience as a result of VR gaming. Two important findings helped define the course of this work. First, the three experimental conditions that were purely based on perceived task difficulty (easy, moderate, and hard) did not demonstrate significant differences in the flow ratings of the participants, even though easy and hard modes were expected to produce low flow ratings, as a result of anxiety and boredom respectively. However, the average global flow rating (sum of all flow's dimensions on the Flow State Scale – 2 questionnaire) indicated that participants rated their flow experience highly, regardless of the game's difficulty. Interestingly, this led to the consideration that the way difficulty was conceived at the development stage was arbitrary, but also that the overall design of the game was compatible with the flow experience.

Next, we identified correlations between the flow dimensions and the personality ratings, as captured through the Big Five Inventory scale. The most significant predictor was the trait of agreeableness, that explained more than 20% of the individual ratings on flow. Previous research showed that agreeable individuals tend to be resistant to annoyance (Johnson et al., 2012), that underlies the emotional states of anxiety and boredom (Eastwood et al., 2007; Monahan, 1981), elevating personality as a significant moderator of the flow ratings across the three experimental conditions. Under this consideration, it was speculated that personality might be a usable parameter in the detection of flow through physiology.

In the second study (**Chapter 5**), the second goal of this research was addressed (objective 2; see **Research Goals**). Initially, the TD-VR game underwent design modifications in order to maximize the differences between the easy and hard modes (what we originally thought to reflect boredom and anxiety respectively). Most importantly, a dynamic difficulty adjustment mechanic was implemented so that the difficulty would not have been as arbitrarily defined as in the previous study (see Chapter 5, **Rounds and Dynamic Difficulty Adjustment**). In addition, the conditions of flow, anxiety and boredom were merged into one holistic game session, thus experienced by all players instead of allocating each player randomly to a single difficulty mode. The gameplay experience was investigated with heart rate and eye blink variability as objective measures of flow, anxiety and boredom.

In line with the second research goal, physiological patterns were shown to have underlain the different experimental conditions. The results demonstrated increased heart rate, eye blink rates and eye blink durations during boredom compared to flow and anxiety, whereas eye blink amplitudes were on average maximal during anxiety than during boredom (see Chapter 5, **Statistical results**). The analysis workflow starts by submitting specialized features extracted from these physiological recordings to three competitive classifiers.

Interestingly, the performance of the classifier strongly depended on the class definition. When using the original design classes of flow, anxiety and boredom, based on the game round each observation was extracted from, the classifier was able to discriminate among the classes with a maximum accuracy of 56.47%, achieved with BDT. These results were comparable to those reported by Chanel and colleagues (2008) in a similar setting involving physiological analysis of game-induced flow, anxiety, and boredom. Contrarily, when bisecting the participants' self-reported experience of flow into two groups (low- and high-flow ratings), the BDT classifier, which scored the maximum accuracy among the three classifiers, achieved an accuracy of 89.16%. Similar observations were made in Chanel and co-authors (2006). In the context of emotion, they suggested that self-evaluation be preferred over predefined labels, as they also reported higher classification rates with the former. However, in the case of this work, both flow definitions were needed to approach the research goals.

The classification outcomes of the two flow definitions (subjective ratings versus game design) led us to devise a new approach for assessing the effectiveness of the game design. A clustering approach with probabilistic assignment of observations to each cluster was used to elucidate the origins of this inconsistency. From the clustering solution and subsequent linear regression analyses, it was revealed that the personality traits of agreeableness, extraversion and conscientiousness explained a large portion of each designated cluster. These traits were further embedded in the sample data through a simple element-wise computation and the classifier was retrained and cross-validated with the modified data, maintaining however the original game design classes as the prime flow definition. A 10-fold cross-validation yielded classification rates that were comparable to the labels generated from the self-reports. Taken together, the approach of the study proposed a novel avenue for evaluating the compatibility of a game with specific personality traits. This is different from Benlamine and colleagues (2017) who incorporated the Big Five traits directly as inputs to classifiers.

In the third and final study of the thesis (**Chapter 6**), the last two research goals were addressed (objectives 3 & 4; **Research Goals**). As per the third goal, electrophysiological patterns emerged throughout the three experimental conditions of flow, anxiety and boredom. The statistical analyses revealed that activity from three frequency bands – theta (4-8 Hz), alpha (8-12 Hz) and beta (12-30 Hz) – emerged relative to the type of round the participants played (see Chapter 6, **EEG band activity**). However, the electrophysiological responses during flow and anxiety were largely homogenous, as also observed in **Chapter 5**.

Both flow and anxiety were modulated by high fronto-central theta activity, as well as high right-central and bilateral posterior beta. These results were significant when contrasted with the boredom condition. On the other hand, boredom exhibited increased power in the alpha rhythm across several regions of interest. Asymmetry in the alpha band for inferring approach/withdrawal motivation did not seem to be sufficiently variable across the different experimental conditions to be considered a useful metric in the specific design of TD-VR. This again contrasts Benlamine and colleagues (2017), who were able to classify a facet of the GameFlow questionnaire, namely Mastery versus Performance, with an F1-score of 75% (SVM) and 81% (Random Forests).

The fourth goal was to select a subset of the original 32 electrodes that are maximally involved in the discrimination of the three experimental conditions in TD-VR. The results indicated that F7, FT7, C6, CP3, TP7 and Pz were commonly involved in the classification of the experimentally induced states and the self-reported experience of flow. Following the clustering and linear regression approach that was employed in **Chapter 5**, the cross-validated decoding accuracy was similar to the previous study, which employed two physiological modalities (HRV and EBV) instead of one (EEG).

Specific aspects of these findings were discussed in the corresponding chapters, whilst more general implications are discussed in the next section.

7.2 Theoretical and practical implications

The present work was founded on existing theoretical and experimental advancements in games research and the flow state. The flow literature is far from a saturation point; this is not only evident from the diversity of experimental methods, but also from the different aspects of flow that are being evaluated. While the richness thereof is indicative of a vibrant and active research area, it also veers away from efforts to systematize the measurement of flow[10].

7.2.1 Theoretical considerations

Standardizing the measurement of flow. In this work, two definitions of flow were utilized. The first definition is by Jackson and Eklund (2002) and it is directly derived from the original theory of flow (Csikszentmihalyi, 1975, 1990). This definition was measured via the Flow State Scale – 2 (FSS-2), which aims to capture salient properties of flow across nine dimensions, based on the self-reported perception of the flow state. The second definition is an implementation of the first definition, but adapted in the context of video games, using literature-driven design techniques to elicit the flow state, as well as anxiety and boredom. From the physiological responses collected from participants, the overlap of electrodes

[10] The interested reader may refer to Swann et al. (2018) who raise interesting points on the state of flow research.

toward resolving the two different flow definitions as target classes was interpreted as a more robust representation of flow. In our view, the approach presented here contributes toward the standardization of experimental approaches in the measurement of flow that is worth refining further. We also consider important that future studies provide a quantification measure of the extent the player-reported experience overlaps physiologically with the researchers' intended experience.

Post-assessment of flow. Earlier, the potential caveats of a retrospective assessment of the subjective flow experience were illustrated (see Chapter 2, **The challenges in measuring flow in video games**). In this work, for a game session to have endured approximately 30 minutes on average, it is puzzling how well the physiological data were mapped onto the self-reported experience of flow (chapters **5** and **6)** through classification processes. This is interesting because the flow questionnaire used in this work (FSS-2) is an explicit measure of one's subjective experience, which automatically assumes that the person can retrospectively access their flow experience. Similar observations, but in the context of self-reports for human motives, were made in Müller et al. (2018).

The articulation and/or evaluation of one's flow experience, when asked to report it in structured interviews or questionnaires, does allude to the presence of meta-awareness, i.e., the ability to make conscious judgments about and monitor the experience (e.g., Baird et al., 2019; Dunne et al., 2019). The flow experience in FSS-2 is reported through statements that are majorly written in first person. However, this notion contrasts Chin and Schooler (2010), who suggested that meta-awareness is disruptive to flow. The successful mapping of the physiological responses during the TD-VR game and the reported experience of flow, as well as the overlap between the two definitions of flow, suggest that this may not be entirely true.

Yet, it is not known whether the appraisal of one's flow experience undergoes evaluation during the *experience* of flow or during the *recollection* of the experience. Schooler and Mauss (2010) argued that the absence of self-reflection during flow can impair people's ability to remember their experience. However, in an earlier work of the first author, it was suggested that some aspects of the experience can be made aware to oneself (Schooler, 2002). Furthermore, according to Chun and Turk-Browne (2007), attention and memory encoding are intertwined. If one is able to recall their flow experience, then a portion of their attention might be involved in the mnemonic encoding of the experience.

An interesting future direction would be to evaluate how far in time the memory of one's flow experience is sustained to assess its impact on the self-reported ratings. Perhaps, the involvement of self-reflection and meta-awareness during flow is more complex than originally assumed. These accounts likely point to the existence of a brain structure that makes memory formation possible during flow. For example, studies have shown that the temporoparietal junction and the medial prefrontal cortex are incumbent upon one's ability to and frequency of recalling dreams (Eichenlaub et al., 2014; Vallat et al., 2018). Both of those structures have similarly been reported to be involved in the experience of flow and self-reflection (Gusnard et al., 2001; Klasen et al., 2012; Ulrich et al., 2014).

7.2.2 Applications

The idea of electroencephalography embedded in virtual reality head-mounted displays has gained traction in the recent years since the inception of the present work. A number of recent studies aimed to evaluate the quality of EEG signal in VR and found that the signal quality remained largely unaffected (e.g., Cattan et al., 2018; Hertweck et al., 2019; Sá et al., 2020; Tauscher et al., 2019). Our EEG study (**Chapter 6**) corroborates these findings. However, the participants were instructed to minimize their movements, hence one cannot confidently extrapolate to other settings, such as exergames (Yoo & Kay, 2016). The potential productisation of wearable EEG and VR would also entail several design considerations.

First, the integration of dry sensors is a standard and intuitive practice in the industry, even though their ability to produce comparable quality to wet sensors is somewhat controversial (Lopez-Gordo et al., 2014). Factors that can significantly affect the sensors' quality and impedance are the size and material of the electrode (Griss et al., 2002). A second consideration is whether the EEG/VR-HMD is going to be wired or wireless, given the market's increasing adoption of wireless technologies (Huang et al., 2018). This can affect the speed in which the data is transmitted to the processing unit, placing an accumulated latency strain on the application that depends on the EEG input. Nevertheless, studies have shown promising results whereby wireless EEG with dry sensors can perform comparably well to a wired wet system – as used in the present work (Hinrichs et al., 2020; Kam et al., 2019; Radüntz, 2018).

Among the most well-known consumer-grade mobile EEG products are MindWave Mobile 2 (NeuroSky), Enobio 8 (Neuroelectrics), Epoc-X (EMOTIV), and Muse 2/S (InteraXon Inc.). Notable examples of EEG meant for VR integration are Looxid Link (Looxid Labs), DSI VR300 (Wearable Sensing), and the upcoming Galea (OpenBCI). Gaming is the most targeted application of consumer-grade EEG (Sawangjai et al., 2019; Todd et al., 2012), suggesting a considerable interest in the area. The potential of these technologies in video games is not exclusive to passive monitoring of mental states, referred to as stealth assessment (Shute et al., 2009), but also to brain-controlled game playing (Sawangjai et al., 2019). Many efforts are in the area of adaptive gaming, whereby the user's mental state informs the game and forces it to adapt in order to maintain the level of a desired state to a target threshold (e.g., Carofiglio et al., 2019; Fernández et al., 2017). While games should already adapt to the player to produce flow-like episodes (as argued in Chapter 2, **Achieving flow by achieving balance**), using brain activity as additional input could help tailor dynamic difficulty adjustment systems that eliminate the need for fixed increments/decrements depending on the player's performance. For example, the game's difficulty could adapt to the alpha rhythm of the user (e.g., Dey et al., 2019).

The current work focused on the integration of physiological tracking in VR applications. The supporting aim was to assess the game design by helping developers evaluate the appeal of their game during mass-testing. Furthermore, stakeholders can utilize this data to forecast the revenue of their title. However, the applications are not limited to gaming. Similar technologies can be used for marketing purposes, such as understanding the consumers' behavior and preferences (e.g., Alvino et al., 2019), education to maximize learning outcomes (Chiang et al., 2018), exposure therapy (Stolz et al., 2019), neurodevelopmental disorders research (Lau-Zhu et al., 2019) and healthcare in general, improvement of assistive technologies (Kumar et al., 2017), training (Roos et al., 2017), and driving (Gao et al., 2019). Another interesting application would be to expose individuals interested in career orientation to different routine work scenarios and guide their choices based on the virtual scenarios that maximized their flow experience.

In summary, the potential of creating better products by tracking flow, anxiety, boredom, and related mental states during activity engagement, using physiological signals, is consistently showcased in the literature. The experiments comprising this work provide support and highlight that the detection of these states is not a trivial venture, while a general metric may be hard to obtain (Ye et al., 2020). Indeed, this is not only due to the difficulty in capturing a stable definition of flow, but also because of various limitations in research works. The limitations of this work are discussed in the next section.

7.3 Limitations and future work

The approach employed in the present work utilized a self-report measure of personality, the Big Five Inventory, as a means of evaluating the personality types most compatible with the game design of TD-VR. The overarching goal was to devise an approach for effectively assessing the game design from the flow experience angle, as a preliminary step for informing the design in future works. Self-report measures were instrumental to decoding the game-design levels, suggesting the need for design modifications that reconcile personality-based preferences and sense of control. However, important topics can still be queried. For instance, it remains unclear how the game design should be modified to be made compatible with a broader range of personality types. Indeed, this is challenging to answer based on the present work alone, as the game design was evaluated indirectly through the lens of personality traits and flow assessments collected after the game. This is the most intriguing direction that we will be investigating in future work.

The game design phases were treated as a fixed stimulus and we observed how the physiological responses changed as a function of personality traits in each experimental condition. To have been able to answer the question above, a modified methodological approach would be needed, where one should record the difficulty adjustment over the course of the game, employ different reward schedules (e.g., randomization of reward acquisition times) and magnitudes of the rewards, as well as types of rewards. In doing so, the game would be treated as a dynamic stimulus, able to provide a map of (interacting) game mechanics and personality types.

One could then estimate optimal modifications to improve the game experience for specific personality types or to derive an average model that encapsulates multiple personality types at the same time. Future studies should elaborate on this crucial topic, as well as expand on the battery of self-reports to uncover additional factors that enhance the learning quality. Not only will game developers be able to assess the appeal of their game, but they will also be able to know what modifications they should perform to widen their player audience. Notably, a growing body of research has advocated the integration of questionnaires on player experience within the virtual environment, commonly known as inVRQ (Alexandrovsky et al., 2020; Putze et al., 2020; Regal et al., 2019), which is an interesting direction for future studies using VR or non-VR technologies.

Another limitation relates to the assumptions behind the way the data were grouped. For example, it can be argued that the first round in the game, which was grouped under the flow label, might be too premature to have successfully induced a flow episode. Similarly, the dynamic difficulty adjustment mechanic was not activated upon the inception of the first round, and thus the skills-difficulty balance was *assumed* rather than *observed*. Unfortunately, the way data should be grouped is often not straightforward, and hence the motivation of this study. Grouping the data based on the design specifications of each round was considered a reasonable approach, albeit future work is needed to validate this approach in other settings. More trials should have been introduced for the anxiety and boredom conditions, even if this was intended. For example, De Manzano and colleagues (2010) suggested that five trials for boredom may be insufficient. However, we did not want to overpower the adaptive gameplay. After all, games are generally meant to be immersive and fun. Hence, it should be acknowledged that the experimental setting of the TD-VR game does not translate well to commercial games.

Critically, it was assumed that the players would alternate between three discrete states, namely flow, anxiety and boredom. This assumption disregarded the possibility that other mental states could have taken place. For example, a state of neutrality akin to the baseline or resting state was unaccounted for, even though it could have been a possible candidate during the game. Additionally, our analytical approach took for granted that all the hypothesized states (flow, anxiety and boredom) would be experienced, to an extent, and/or physiologically manifested by all of the players.

The ambiguity regarding the grouping of data suggested the need for a fuzzy clustering method (Chapter 3, **Unsupervised learner: Fuzzy C-means**), which is agnostic to our original assumptions, yet considers the uncertainty in the clustering assignment. Our future studies will seek to standardize the assumptions on different mental states in the experimental game (i.e., the annotation process). For example, most studies regard the flow condition as a function of task difficulty and user skills (e.g., Chen, 2007), but the time the task difficulty is manipulated, relative to the task's inception, is rarely considered.

With the inclusion of time as an additive dimension, one might be able to better illustrate the temporal evolution of flow that explains whether it should be treated equally across the duration of the experiment or across consecutive experiments in longitudinal studies. In these complex settings, state-of-the-art analytical tools could contribute to a better understanding of flow's elicitation by design. For instance, deep learners specialized in representing high-dimensional features in compact latent spaces, inverting such mapping (e.g., Goodfellow et al., 2014, Donahue et al., 2016), could foster promising approaches toward the detection and manipulation of flow in real-time settings.

8 References

Abuhamdeh, S. (2020). Investigating the "flow" experience: Key conceptual and operational issues. *Frontiers in Psychology*, *11*, 158.

Adams, E. (2009). *Fundamentals of game design* (2nd ed.). Indianapolis, IN: New Riders Publishing.

Adams, E., & Dormans, J. (2012). *Game mechanics: advanced game design*. New Riders.

Aftanas, L. I., & Golocheikine, S. A. (2002). Non-linear dynamic complexity of the human EEG during meditation. *Neuroscience letters*, *330*(2), 143-146.

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS quarterly*, 665-694.

Ahn, J. W., Ku, Y., & Kim, H. C. (2019). A novel wearable EEG and ECG recording system for stress assessment. *Sensors*, *19*(9), 1991.

Aidman, E., Chadunow, C., Johnson, K., & Reece, J. (2015). Real-time driver drowsiness feedback improves driver alertness and self-reported driving performance. *Accident Analysis & Prevention*, *81*, 8-13.

Akiduki, H., Nishiike, S., Watanabe, H., Matsuoka, K., Kubo, T., & Takeda, N. (2003). Visual-vestibular conflict induced by virtual reality in humans. *Neuroscience letters*, *340*(3), 197-200.

Akın, A. (2012). The relationships between Internet addiction, subjective vitality, and subjective happiness. *Cyberpsychology, Behavior, and Social Networking*, *15*(8), 404-410.

Al-Azawi, R., Ayesh, A., & Obaidy, M. A. (2013). Generic evaluation framework for games development methodology. In *2013 Third International Conference on Communications and Information Technology (ICCIT)* (pp. 55-60). IEEE.

Albrecht, K., Abeler, J., Weber, B., & Falk, A. (2014). The brain correlates of the effects of monetary and verbal rewards on intrinsic motivation. *Frontiers in neuroscience*, *8*, 303.

Alexandrovsky, D., Putze, S., Bonfert, M., Höffner, S., Michelmann, P., Wenig, D., ... & Smeddinck, J. D. (2020). Examining design choices of questionnaires in vr user studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-21).

Alicart, H., Cucurell, D., Mas-Herrero, E., & Marco-Pallarés, J. (2015). Human oscillatory activity in near-miss events. *Social cognitive and affective neuroscience*, *10*(10), 1405-1412.

Allen, J. J., Coan, J. A., & Nazarian, M. (2004). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biological psychology*, *67*(1-2), 183-218.

Allen, J. J., Harmon-Jones, E., & Cavender, J. H. (2001). Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG. *Psychophysiology*, *38*(4), 685-693.

Allison, F., Carter, M., & Gibbs, M. (2015). Good frustrations: the paradoxical pleasure of fearing death in DayZ. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* (pp. 119-123).

Alvino, L., van der Lubbe, R., Joosten, R. A., & Constantinides, E. (2019). Which wine do you prefer? An analysis on consumer behaviour and brain activity during a wine tasting experience. *Asia Pacific Journal of Marketing and Logistics*.

Anderson, C. A., & Ford, C. M. (1986). Affect of the game player: Short-term effects of highly and mildly aggressive video games. *Personality and social psychology bulletin*, *12*(4), 390-402.

Andrade, G., Ramalho, G., Gomes, A. S., & Corruble, V. (2006). Dynamic Game Balancing: An Evaluation of User Satisfaction. *AIIDE*, *6*, 3-8.

Andreassen, C. S., Hetland, J., Molde, H., & Pallesen, S. (2011). 'Workaholism'and potential outcomes in well-being and health in a cross-occupational sample. *Stress and health*, *27*(3), e209-e214.

Ang, C. S. (2006). Rules, gameplay, and narratives in video games. *Simulation & Gaming*, *37*(3), 306-325.

Angelidis, A., Hagenaars, M., van Son, D., van der Does, W., & Putman, P. (2018). Do not look away! Spontaneous frontal EEG theta/beta ratio as a marker for cognitive control over attention to mild and high threat. *Biological psychology*, *135*, 8-17.

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425-438.

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of general psychology*, *10*(3), 229-240.

Apperley, T. H. (2006). Genre and game studies: Toward a critical approach to video game genres. *Simulation & Gaming*, *37*(1), 6-23.

Arvaneh, M., Guan, C., Ang, K. K., & Quek, H. C. (2010). EEG channel selection using decision tree in brain-computer interface. In *Proceedings of the Second APSIPA Annual Summit and Conference* (pp. 225-230).

Arzate, C., and Ramirez, J. A. U. (2017). Player-centered game AI from a flow perspective: towards a better understanding of past trends and future directions. *Entertain. Comput.* 20, 11–24. doi: 10.1016/j.entcom.2017.02.003

Ashmore, C., & Nitsche, M. (2007). The Quest in a Generated World. In *DiGRA Conference*.

Atroszko, P. A., Andreassen, C. S., Griffiths, M. D., & Pallesen, S. (2015). Study addiction—A new area of psychological study: Conceptualization, assessment, and preliminary empirical findings. *Journal of behavioral addictions*, *4*(2), 75-84.

Augustine, A. A., & Larsen, R. J. (2015). *Personality, affect, and affect regulation.* In M. Mikulincer, P. R. Shaver, M. L. Cooper, & R. J. Larsen (Eds.), *APA handbooks in psychology. APA handbook of personality and social psychology, Vol. 4. Personality processes and individual differences* (p. 147–165). American Psychological Association. https://doi.org/10.1037/14343-007

Avery, P., Togelius, J., Alistar, E., & Van Leeuwen, R. P. (2011). Computational intelligence and tower defence games. In *2011 IEEE Congress of Evolutionary Computation (CEC)* (pp. 1084-1091). IEEE.

Babiker, A., Baashar, Y., Alkahtani, A. A., Faye, I., & Alkawsi, G. (2021). Towards Detection of Interest Using Physiological Sensors. *Applied Sciences*, *11*(3), 1318.

Babiloni, C., Miniussi, C., Babiloni, F., Carducci, F., Cincotti, F., Del Percio, C., ... & Rossini, P. M. (2004). Sub-second "temporal attention" modulates alpha rhythms. A high-resolution EEG study. *Cognitive Brain Research*, *19*(3), 259-268.

Baijal, S., & Srinivasan, N. (2010). Theta activity and meditative states: spectral changes during concentrative meditation. *Cognitive processing*, *11*(1), 31-38.

Baird, B., Riedner, B. A., Boly, M., Davidson, R. J., & Tononi, G. (2019). Increased lucid dream frequency in long-term meditators but not following mindfulness-based stress reduction training. *Psychology of Consciousness: Theory, Research, and Practice*, *6*(1), 40.

Baka, E., Stavroulia, K. E., Magnenat-Thalmann, N., & Lanitis, A. (2018). An EEG-based evaluation for comparing the sense of presence between virtual and physical environments. In *Proceedings of Computer Graphics International 2018* (pp. 107-116).

Bakker, A. B. (2005). Flow among music teachers and their students: The crossover of peak experiences. *Journal of vocational behavior*, *66*(1), 26-44.

Bakker, A. B., Van Der Zee, K. I., Lewig, K. A., & Dollard, M. F. (2006). The relationship between the big five personality factors and burnout: A study among volunteer counselors. *The Journal of social psychology*, *146*(1), 31-50.

Baldassarre, G. (2011). What are intrinsic motivations? A biological perspective. In *2011 IEEE international conference on development and learning (ICDL)* (Vol. 2, pp. 1-8). IEEE.

Banakou, D., Chorianopoulos, K., & Anagnostou, K. (2009). Avatars' appearance and social behavior in online virtual worlds. In *2009 13th Panhellenic Conference on Informatics* (pp. 207-211). IEEE.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American psychologist*, *37*(2), 122.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press, 1998).

Banerjee, A., Pal, M., Datta, S., Tibarewala, D. N., & Konar, A. (2014). Eye movement sequence analysis using electrooculogram to assist autistic children. *Biomedical Signal Processing and Control*, *14*, 134-140.

Barathi, S. C., Proulx, M., O'Neill, E., & Lutteroth, C. (2020). Affect Recognition using Psychophysiological Correlates in High Intensity VR Exergaming. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia*, *35*(10), 1373-1380.

Barnacle, G. E., Schaefer, A., Tsivilis, D., & Talmi, D. (2016). *Understanding Emotional Memory: Cognitive Factors* (Doctoral dissertation). University of Manchester, United Kingdom.

Barry, R. J., & De Blasio, F. M. (2018). EEG frequency PCA in EEG-ERP dynamics. *Psychophysiology*, *55*(5), e13042.

Bartle, R. A. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *Journal of MUD research*, *1*(1), 19.

Bartle, R. A. (2004). *Designing virtual worlds*. Berkeley, CA: New Riders.

Bassi, M., Steca, P., Monzani, D., Greco, A., & Delle Fave, A. (2014). Personality and optimal experience in adolescence: Implications for well-being and development. *Journal of Happiness Studies*, *15*(4), 829-843.

Baştuğ, E., Bennis, M., Médard, M., & Debbah, M. (2017). Toward interconnected virtual reality: Opportunities, challenges, and enablers. *IEEE Communications Magazine*, *55*(6), 110-117.

Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine learning*, *36*(1-2), 105-139.

Baumann, N., & Scheffer, D. (2011). Seeking flow in the achievement domain: The achievement flow motive behind flow experience. *Motivation and Emotion*, *35*(3), 267-284.

Baumgartl, H., Bayerlein, S., & Buettner, R. (2020). Measuring extraversion using EEG data. In *NeuroIS Retreat* (pp. 259-265). Springer, Cham.

Baumgartner, T., Speck, D., Wettstein, D., Masnari, O., Beeli, G., and Jäncke, L. (2008). Feeling present in arousing virtual reality worlds: prefrontal brain regions differentially orchestrate presence experience in adults and children. *Front. Hum. Neurosci.* 2:8. doi: 10.3389/neuro.09.008.2008

Bavelier, D., Achtman, R. L., Mani, M., & Föcker, J. (2012). Neural bases of selective attention in action video game players. *Vision research*, *61*, 132-143.

Bean, A., & Groth-Marnat, G. (2016). Video gamers and personality: A five-factor model to understand game playing style. *Psychology of Popular Media Culture*, *5*(1), 27.

Beard, K. W., & Wolf, E. M. (2001). Modification in the proposed diagnostic criteria for Internet addiction. *CyberPsychology & Behavior*, *4*(3), 377-383.

Beaumont, S. L. (2009). Identity processing and personal wisdom: An information-oriented identity style predicts self-actualization and self-transcendence. *Identity: An International Journal of Theory and Research*, *9*(2), 95-115.

Beck, C. W., & Bliwise, N. G. (2014). Interactions are critical. *CBE—Life Sciences Education*, *13*(3), 371-372.

Bednarik, R., Koskinen, J., Vrzakova, H., Bartczak, P., & Elomaa, A. P. (2018). Blink-Based Estimation of Suturing Task Workload and Expertise in Microsurgery. In *Proceedings: 31st IEEE International Symposium on Computer-Based Medical Systems-CBMS 2018*.

Belanche, L. A., & González, F. F. (2011). Review and evaluation of feature selection algorithms in synthetic problems. *arXiv preprint arXiv:1101.2320*.

Bench, S., & Lench, H. (2013). On the function of boredom. *Behavioral sciences*, *3*(3), 459-472.

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation research part F: traffic psychology and behaviour*, *14*(3), 199-208.

Benlamine, M. S., Dombouya, R., Dufresne, A., & Frasson, C. (2017). Game Experience and Brain based Assessment of Motivational Goal Orientations in Video Games. In *International Conference on Brain Function Assessment in Learning* (pp. 118-132). Springer, Cham.

Benner, P. (1984). *From novice to expert: Excellence and Power in Clinical Nursing Practice*. California, Menlo Park: Addison-Wesley.

Benson, H., Beary, J. F., & Carol, M. P. (1974). *The Relaxation Response*. *Psychiatry*, *37*(1), 37–46. doi:10.1080/00332747.1974.11023785

Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill.

Bernadine, C., Nanda, N., Vijayalakshmi, R., Thilaga, M., Naga, D., & Nabaraj, D. (2013). Breaking the Camel's Back: Can Cognitive Overload be Quantified in the Human Brain?. *Procedia-Social and Behavioral Sciences*, *97*, 21-29.

Berta, R., Bellotti, F., De Gloria, A., Pranantha, D., & Schatten, C. (2013). Electroencephalogram and physiological signal analysis for assessing flow in games. *IEEE Transactions on Computational Intelligence and AI in Games*, *5*(2), 164-175.

Besedeš, T., Deck, C., Sarangi, S., & Shor, M. (2015). Reducing choice overload without reducing choices. *Review of Economics and Statistics*, *97*(4), 793-802.

Bessière, K., Seay, A. F., & Kiesler, S. (2007). The ideal elf: Identity exploration in World of Warcraft. *Cyberpsychology & behavior*, *10*(4), 530-535.

Bezdek, J. C. (1981). *Pattern Recognition With Fuzzy Objective Function Algorithms*. Plenum Press, New York, NY.

Bhowan, U., Johnston, M., Zhang, M., & Yao, X. (2012). Evolving diverse ensembles using genetic programming for classification with unbalanced data. *IEEE Transactions on Evolutionary Computation*, *17*(3), 368-386.

Bian, Y., Yang, C., Gao, F., Li, H., Zhou, S., Li, H., ... & Meng, X. (2016). A framework for physiological indicators of flow in VR games: construction and preliminary evaluation. *Personal and Ubiquitous Computing*, *20*(5), 821-832.

Bianchi-Berthouze, N., Cairns, P., Cox, A., Jennett, C., & Kim, W. W. (2006). On posture as a modality for expressing and recognizing emotions. In *Emotion and HCI workshop at BCS HCI London*.

Biercewicz, K., Borawski, M., & Duda, J. (2020). Method for Selecting an Engagement Index for a Specific Type of Game Using Cognitive Neuroscience. *International Journal of Computer Games Technology*, *2020*.

Bilgihan, A., Cobanoglu, C., Nusair, K., Okumus, F., & Bujisic, M. (2013). A quantitative study exploring the difference between gaming genre preferences. *The computer games journal*, *2*(1), 19-40.

Bins, J., & Draper, B. A. (2001). Feature selection from huge feature sets. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001* (Vol. 2, pp. 159-165). IEEE.

Blaszczynski, A. (2008). Commentary: A response to "Problems with the concept of video game "addiction": Some case study examples". *International Journal of Mental Health and Addiction*, *6*(2), 179-181.

Bodala, I. P., Li, J., Thakor, N. V., & Al-Nashash, H. (2016). EEG and eye tracking demonstrate vigilance enhancement with challenge integration. *Frontiers in human neuroscience*, *10*, 273.

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature reviews neuroscience*, *12*(12), 752.

Bond, M., & Beale, R. (2009). What makes a good game?: using reviews to inform design. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (pp. 418-422). British Computer Society.

Boniwell, I. (2008). *Positive psychology in a nutshell: A balanced introduction to the science of optimal functioning*. New York: Personal Well-Being Centre.

Bonometti, V., Ringer, C., Ruiz, M., Wade, A., & Drachen, A. (2020). From Theory to Behaviour: Towards a General Model of Engagement. *arXiv preprint arXiv:2004.12644*.

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition?. *Frontiers in psychology*, *2*, 226.

Borderie, J., & Michinov, N. (2016). Identifying Flow in Video Games: Towards a New Observation-Based Method. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS), 8*(3), 19-38.

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: pleasure and arousal in memory. *Journal of experimental psychology: Learning, Memory, and Cognition*, *18*(2), 379.

Brandtzæg, P. B., Følstad, A., & Heim, J. (2018). Enjoyment: lessons from Karasek. In *Funology 2* (pp. 331-341). Springer, Cham.

Braun, B., Stopfer, J. M., Müller, K. W., Beutel, M. E., & Egloff, B. (2016). Personality and video gaming: Comparing regular gamers, non-gamers, and gaming addicts and differentiating between game genres. *Computers in Human Behavior*, *55*, 406-412.

Braun, M. H. (2019). *Enhancing User Experience with Olfaction in Virtual Reality* (Doctoral dissertation). City University of London, London, United Kingdom.

Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*(2), 123-140.

Brich, J., Rogers, K., Frommel, J., Weidhaas, M., Bruckner, A., Mirabile, S., ... & Weber, M. (2015). LiverDefense: using a tower defense game as a customisable research tool. In *2015 7th International Conference on Games and Virtual Worlds for Serious Applications (VS-Games)* (pp. 1-8). IEEE.

Briesemeister, B. B., Tamm, S., Heine, A., & Jacobs, A. M. (2013). Approach the good, withdraw from the bad—a review on frontal alpha asymmetry measures in applied psychological research. *Psychology*, *4*(03), 261.

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

Brodal, P. (2004). *The central nervous system: structure and function*. Oxford University Press.

Brouwer, A. M., Hogervorst, M. A., Holewijn, M., & van Erp, J. B. (2014a). Evidence for effects of task difficulty but not learning on neurophysiological variables associated with effort. *International Journal of Psychophysiology*, *93*(2), 242-252.

Brouwer, A. M., Van Schaik, M. G., Korteling, J. H., van Erp, J. B., & Toet, A. (2014b). Neuroticism, extraversion, conscientiousness and stress: Physiological correlates. *IEEE transactions on affective computing*, *6*(2), 109-117.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1297-1300). ACM.

Bryce, J., & Rutter, J. (2002). Spectacle of the deathmatch: Character and narrative in first-person shooters. *ScreenPlay: Cinema/videogames/interfaces*, 66-80.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of cognition*, *2*(1), 16.

Buckley, D., Chen, K., & Knowles, J. (2013). Predicting skill from gameplay input to a first-person shooter. In *2013 IEEE Conference on Computational Inteligence in Games (CIG)* (pp. 1-8). IEEE.

Buckley, K. E., & Anderson, C. A. (2006). A theoretical model of the effects and consequences of playing video games. *Playing video games: Motives, responses, and consequences*, 363-378.

Bunian, S., Canossa, A., Colvin, R., & El-Nasr, M. S. (2018). Modeling individual differences in game behavior using HMM. *arXiv preprint arXiv:1804.00245*.

Burn, C. C. (2017). Bestial boredom: A biological perspective on animal boredom and suggestions for its scientific investigation. *Animal Behaviour*, *130*, 141-151.

Buschman, T. J., and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862. doi: 10.1126/science.1138071

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of personality and Social Psychology*, *76*(5), 839.

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of personality and social psychology*, *50*(2), 260.

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European journal of applied physiology*, *89*(3-4), 319-325.

Canento, F., Fred, A., Silva, H., Gamboa, H., & Lourenço, A. (2011). Multimodal biosignal sensor data handling for emotion recognition. In *SENSORS, 2011 IEEE* (pp. 647-650). IEEE.

Cairns, P. (2016). Engagement in Digital Games. In *Why Engagement Matters*, eds H. O'Brien and P. Carins (New York, NY: Springer International Publishing), 81–104.

Cairns, P., Cox, A., and Nordin, A. I. (2014). Immersion in digital games: review of gaming experience research. *Handb. Digit. Games* 339:784. doi: 10.1002/9781118796443.ch12

Caldwell, Y. T., & Steffen, P. R. (2018). Adding HRV biofeedback to psychotherapy increases heart rate variability and improves the treatment of major depressive disorder. *International journal of Psychophysiology*, *131*, 96-101.

Calleja, G. (2010). Digital games and escapism. *Games and Culture*, *5*(4), 335-353.

Calleja, G. (2011). *In-game: From immersion to incorporation*. Massachusetts: MIT Press.

Calleja, G. (2013). Narrative involvement in digital games. In *Conference Proceedings from Foundations of Digital Games. Chania, Crete, Greece*.

Calvillo-Gámez, E. H., Cairns, P., & Cox, A. L. (2015). Assessing the core elements of the gaming experience. In *Game user experience evaluation* (pp. 37-62). Springer, Cham.

Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior Analyst*, *24*(1), 1-44.

Cao, F., & Su, L. (2007). Internet addiction among Chinese adolescents: prevalence and psychological features. *Child: care, health and development*, *33*(3), 275-281.

Çardak, M. (2013). Psychological well-being and Internet addiction among university students. *Turkish Online Journal of Educational Technology-TOJET*, *12*(3), 134-141.

Cardamone, L., Yannakakis, G. N., Togelius, J., & Lanzi, P. L. (2011). Evolving interesting maps for a first person shooter. In *European Conference on the Applications of Evolutionary Computation* (pp. 63-72). Springer, Berlin, Heidelberg.

Carofiglio, V., De Carolis, B. N., & D'Errico, F. (2019). A BCI-based Assessment of a Player's State of Mind for Game Adaptation. In *GHITALY@ CHItaly*.

Carpentier, J., Mageau, G. A., & Vallerand, R. J. (2012). Ruminations and flow: Why do people with a more harmonious passion experience higher well-being?. *Journal of Happiness studies*, *13*(3), 501-518.

Carroll, M., Osborne, E., & Yildirim, C. (2019). Effects of VR Gaming and Game Genre on Player Experience. In *2019 IEEE Games, Entertainment, Media Conference (GEM)* (pp. 1-6). IEEE.

Carson, D. (2000). Environmental storytelling: Creating immersive 3D worlds using lessons learned from the theme park industry. *Gamasutra*, Available from: https://www.gamasutra.com/view/feature/131594/environmental_storytelling_.php.

Carson, H. J., & Collins, D. (2016). The fourth dimension: A motoric perspective on the anxiety–performance relationship. *International Review of Sport and Exercise Psychology*, *9*(1), 1-21.

Caruana, R., & Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on Machine learning* (pp. 161-168). ACM.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports*, *100*(2), 126.

Castronova, E. (2005). *Synthetic worlds: The business and culture of online games*. University of Chicago Press.

Cattan, G., Andreev, A., Mendoza, C., & Congedo, M. (2018). The impact of passive head-mounted virtual reality devices on the quality of EEG signals. In *Workshop on Virtual Reality Interaction and Physical Simulation*. The Eurographics Association.

Causse, M., Chua, Z., Peysakhovich, V., Campo, N., and Matton, N. (2017). Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS. *Sci. Rep.* 7:5222. doi: 10.1038/s41598-017-05378-x

Cawley, G. C., & Talbot, N. L. (2007). Preventing over-fitting during model selection via Bayesian regularisation of the hyper-parameters. *Journal of Machine Learning Research*, *8*(Apr), 841-861.

Ceja, L., & Navarro, J. (2011). Dynamic patterns of flow in the workplace: Characterizing within-individual variability using a complexity science approach. *Journal of Organizational Behavior*, *32*(4), 627-651.

Chanel, G., Kronegg, J., Grandjean, D., & Pun, T. (2006). Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals. In *International workshop on multimedia content representation, classification and security* (pp. 530-537). Springer, Berlin, Heidelberg.

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2008). Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. In *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era* (pp. 13-17). ACM.

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2011). Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *41*(6), 1052-1063.

Chang, C. C., Liang, C., Chou, P. N., & Lin, G. Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior*, *71*, 218-227.

Charles, D., & Black, M. (2004). Dynamic player modeling: A framework for player-centered digital games. In *Proc. of the International Conference on Computer Games: Artificial Intelligence, Design and Education* (pp. 29-35).

Charles, R., & Nixon, J. (2017). Blink counts can differentiate between task type and load. *Proceedings of Contemporary Ergonomics & Human Factors* (pp. 257-265).

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of artificial intelligence research*, *16*, 321-357.

Chen, C. H., Sun, C. T., & Hsieh, J. (2008). Player guild dynamics and evolution in massively multiplayer online games. *CyberPsychology & Behavior*, *11*(3), 293-301.

Chen, H. (2006). Flow on the net–detecting Web users' positive affects and their flow states. *Computers in human behavior*, *22*(2), 221-233.

Chen, J. (2007). Flow in games (and everything else). *Communications of the ACM*, *50*(4), 31-34.

Chen, S., & Epps, J. (2014). Using task-induced pupil diameter and blink rate to infer cognitive load. *Human–Computer Interaction*, *29*(4), 390-413.

Chen, S., Epps, J., Ruiz, N., & Chen, F. (2011). Eye activity as a measure of human mental effort in HCI. In *Proceedings of the 16th international conference on Intelligent user interfaces* (pp. 315-318). ACM.

Chesham, A., Wyss, P., Müri, R. M., Mosimann, U. P., & Nef, T. (2017). What older people like to play: genre preferences and acceptance of casual games. *JMIR serious games*, *5*(2), e8.

Chess, S., Evans, N. J., & Baines, J. J. (2017). What does a gamer look like? Video games, advertising, and diversity. *Television & New Media*, *18*(1), 37-57.

Cheung, G., & Huang, J. (2011). Starcraft from the stands: understanding the game spectator. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 763-772). ACM.

Chiang, H. S., Hsiao, K. L., & Liu, L. C. (2018). EEG-based detection model for evaluating and improving learning attention. *Journal of Medical and Biological Engineering*, *38*(6), 847-856.

Child, D. (1994). *Psychology and the Teacher* (5th ed.) London: Cassell.

Chin, J., & Schooler, J. W. (2010). Meta-awareness. *Encyclopedia of consciousness*, *2*, 33-41.

Chisholm, J. D., & Kingstone, A. (2012). Improved top-down control reduces oculomotor capture: The case of action video game players. *Attention, Perception, & Psychophysics*, *74*(2), 257-262.

Cho, Y. (2021). Rethinking Eye-blink: Assessing Task Difficulty through Physiological Representation of Spontaneous Blinking. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

Choi, A., & Shin, H. (2018). Quantitative analysis of the effect of an ectopic beat on the heart rate variability in the resting condition. *Frontiers in physiology*, *9*, 922.

Choi, D., & Kim, J. (2004). Why people continue to play online games: In search of critical design factors to increase customer loyalty to online contents. *CyberPsychology & behavior*, *7*(1), 11-24.

Chomboon, K., Chujai, P., Teerarassamee, P., Kerdprasop, K., & Kerdprasop, N. (2015). An empirical study of distance metrics for k-nearest neighbor algorithm. In *Proceedings of the 3rd international conference on industrial application engineering* (pp. 1-6).

Chory, R. M., & Goodboy, A. K. (2011). Is basic personality related to violent and non-violent video game play and preferences?. *Cyberpsychology, Behavior, and Social Networking*, *14*(4), 191-198.

Chou, T. J., & Ting, C. C. (2003). The role of flow experience in cyber-game addiction. *CyberPsychology & Behavior*, *6*(6), 663-675.

Christie, G. J., & Tata, M. S. (2009). Right frontal cortex generates reward-related theta-band oscillatory activity. *Neuroimage*, *48*(2), 415-422.

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current opinion in neurobiology*, *17*(2), 177-184.

Cisler, D., Greenwood, P. M., Roberts, D. M., McKendrick, R., & Baldwin, C. L. (2019). Comparing the relative strengths of EEG and low-cost physiological devices in modeling attention allocation in semiautonomous vehicles. *Frontiers in human neuroscience*, *13*, 109.

Clarke, A. R., Barry, R. J., Karamacoska, D., & Johnstone, S. J. (2019). The EEG theta/beta ratio: a marker of arousal or cognitive processing capacity?. *Applied psychophysiology and biofeedback*, *44*(2), 123-129.

Claypool, M., Claypool, K., & Damaa, F. (2006). The effects of frame rate and resolution on users playing first person shooter games. In *Multimedia Computing and Networking 2006* (Vol. 6071, p. 607101). International Society for Optics and Photonics.

Clearwater, D. (2011). What defines video game genre? Thinking about genre study after the great divide. *Loading...*, *5*(8).

Clemente, M., Rey, B., Rodríguez-Pujadas, A., Barros-Loscertales, A., Baños, R. M., Botella, C., et al. (2013). An fMRI study to analyze neural correlates of presence during virtual reality experiences. *Interact. Comp.* 26, 269–284. doi: 10.1093/iwc/iwt037

Clyde, J., & Thomas, C. (2008), Building an information literacy first-person shooter. *Reference Services Review*, *36*(4), 366-380.

Coan, J. A., & Allen, J. J. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological psychology*, *67*(1-2), 7-50.

Coan, J. A., Allen, J. J., & McKnight, P. E. (2006). A capability model of individual differences in frontal EEG asymmetry. *Biological psychology*, *72*(2), 198-207.

Cohen, B., Dai, M., Yakushin, S. B., & Cho, C. (2019). The neural basis of motion sickness. *Journal of neurophysiology*, *121*(3), 973-982.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT Press.

Collins, A. L., Sarkisian, N., & Winner, E. (2009). Flow and happiness in later life: An investigation into the role of daily and weekly flow experiences. *Journal of Happiness Studies*, *10*(6), 703-719.

Compton, R. J., Gearinger, D., & Wild, H. (2019). The wandering mind oscillates: EEG alpha power is enhanced during moments of mind-wandering. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(5), 1184-1191.

Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, *35*(3), 124-129.

Consalvo, M., & Harper, T. (2009). The sexi(e)st of all: Avatars, gender, and online games. In *Virtual Social Networks* (pp. 98-113). Palgrave Macmillan, London.

Conway, J. C., & Rubin, A. M. (1991). Psychological predictors of television viewing motivation. *Communication Research*, *18*, 443–463.

Cortes, C., Jackel, L. D., & Chiang, W. P. (1995). Limits on learning machine accuracy imposed by data quality. In *Advances in Neural Information Processing Systems* (pp. 239-246).

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Costa Jr, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of personality assessment*, *64*(1), 21-50.

Costa Jr, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and individual Differences*, *12*(9), 887-898.

Cottrell, N. D., & Barton, B. K. (2013). The role of automation in reducing stress and negative affect while driving. *Theoretical Issues in Ergonomics Science*, *14*(1), 53-68.

Cowley, B., Charles, D., Black, M., & Hickey, R. (2008). Toward an understanding of flow in video games. *Computers in Entertainment (CIE)*, *6*(2), 20.

Cox, A., Cairns, P., Shah, P., & Carroll, M. (2012). Not doing but thinking: the role of challenge in the gaming experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 79-88). ACM.

Craveirinha, R., & Roque, L. (2010). Looking for the heart of interactive media: reflections on video games' emotional expression. In *Proceedings of the 3rd International Conference on Fun and Games* (pp. 8-17). ACM.

Critcher, C. R., & Gilovich, T. (2010). Inferring attitudes from mindwandering. *Personality and Social Psychology Bulletin*, *36*(9), 1255-1266.

Critchley, H. D., Corfield, D. R., Chandler, M. P., Mathias, C. J., & Dolan, R. J. (2000). Cerebral correlates of autonomic cardiovascular arousal: a functional neuroimaging investigation in humans. *The Journal of physiology*, *523*(1), 259-270.

Cruz, C., Hanus, M. D., & Fox, J. (2017). The need to achieve: Players' perceptions and uses of extrinsic meta-game reward systems for video game consoles. *Computers in Human Behavior*, *71*, 516-524.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. San Francisco: Jossey-Bass.

Csikzentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikzentmihalyi & I. S. Csikzentmihalyi (Eds.), *Optimal experience* (pp. 15-35). Cambridge, UK: Cambridge University Press.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.

Csikszentmihalyi, M. (1996). *Flow and the Psychology of Discovery and Invention.* New York, NY: Harper Collins.

Csikszentmihalyi, M. (1997a). Finding flow. New York: Basic Books.

Csikszentmihalyi, M. (1997b). Flow and the psychology of discovery and invention. *Harper Perennial, New York*, *39*.

Csikszentmihalyi, M. (1997c). Happiness and creativity going with the flow. *The Futurist*, *31*, 8–12.

Csikszentmihalyi, M. (1999). If we are so rich, why aren't we happy?. *American psychologist*, *54*(10), 821.

Csikszentmihalyi, M. (2014). Play and intrinsic rewards. In *Flow and the foundations of positive psychology* (pp. 135-153). Springer, Dordrecht.

Csikszentmihalyi, M., & Hunter, J. (2003). Happiness in everyday life: The uses of experience sampling. *Journal of happiness studies*, *4*(2), 185-199.

Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of personality and social psychology*, *56*(5), 815.

Cummings, A. H. (2007). The evolution of game controllers and control schemes and their effect on their games. In *The 17th annual university of southampton multimedia systems conference* (Vol. 21).

Czarnecki, W. M., Podlewska, S., & Bojarski, A. J. (2015). Robust optimization of SVM hyperparameters in the classification of bioactive compounds. *Journal of cheminformatics*, *7*(1), 38.

D'Mello, S. S., Chipman, P., & Graesser, A. (2007). Posture as a predictor of learner's affective engagement. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 29, No. 29).

D'Zurilla, T. J., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2011). Predicting social problem solving using personality traits. *Personality and individual Differences*, *50*(2), 142-147.

Da Cruz, J. N., Wan, F., Wong, C. M., & Cao, T. (2015). Adaptive time-window length based on online performance measurement in SSVEP-based BCIs. *Neurocomputing*, *149*, 93-99.

Dale, G., & Green, C. S. (2017). The changing face of video games and video gamers: Future directions in the scientific study of video game play and cognitive performance. *Journal of Cognitive Enhancement*, *1*(3), 280-294.

Danckert, J., Hammerschmidt, T., Marty-Dugas, J., & Smilek, D. (2018). Boredom: Under-aroused and restless. *Consciousness and cognition*, *61*, 24-37.

Darvishi, S., Ridding, M. C., Abbott, D., & Baumert, M. (2013). Investigation of the trade-off between time window length, classifier update rate and classification accuracy for restorative brain-computer interfaces. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1567-1570). IEEE.

Davidson, R. J. (1984). Affect, cognition, and hemispheric specialization. In C. E. Izard, J. Kagan, & R. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 320-365). Cambridge University Press.

Davidson, R. J. (1988). EEG measures of cerebral asymmetry: Conceptual and methodological issues. *International Journal of Neuroscience*, *39*, 71-89. doi:10.3109/00207458808985694

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

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

Davidson, R. J., & Fox, N. A. (1982). Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants. *Science*, *218*(4578), 1235-1237.

Davidson, R. J., Marshall, J. R., Tomarken, A. J., & Henriques, J. B. (2000). While a phobic waits: Regional brain electrical and autonomic activity in social phobics during anticipation of public speaking. *Biological psychiatry*, *47*(2), 85-95.

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *2*, 224-227.

De Greef, T., Lafeber, H., van Oostendorp, H., & Lindenberg, J. (2009). Eye movement as indicators of mental workload to trigger adaptive automation. In *International Conference on Foundations of Augmented Cognition* (pp. 219-228). Springer, Berlin, Heidelberg.

De Manzano, Ö., Cervenka, S., Jucaite, A., Hellenäs, O., Farde, L., & Ullén, F. (2013). Individual differences in the proneness to have flow experiences are linked to dopamine D2-receptor availability in the dorsal striatum. *Neuroimage*, *67*, 1-6.

De Manzano, Ö., Theorell, T., Harmat, L., & Ullén, F. (2010). The psychophysiology of flow during piano playing. *Emotion*, *10*(3), 301.

De Sampaio Barros, M. F., Araújo-Moreira, F. M., Trevelin, L. C., & Radel, R. (2018). Flow experience and the mobilization of attentional resources. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(4), 810-823.

De Schutter, B. (2011). Never too old to play: The appeal of digital games to an older audience. *Games and Culture*, *6*(2), 155-170.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology*, *18*(1), 105.

Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum

Deci, E.L., Koestner, R., & Ryan, R. M. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, *125*(6), 627-668. doi: 10.1037/0033-2909.125.6.627

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Delle Fave, A., Massimini, F., & Bassi, M. (2011a). Optimal experience and meditation: western and asian approaches to well-being. In *Psychological selection and optimal experience across cultures* (pp. 111-126). Springer, Dordrecht.

Delle Fave, A., Massimini, F., & Bassi, M. (2011b). The phenomenology of optimal experience in daily life. In *Psychological Selection and Optimal Experience Across Cultures* (pp. 89-110). Springer, Dordrecht.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

Demerouti, E. (2006). Job characteristics, flow, and performance: The moderating role of conscientiousness. *Journal of occupational health psychology*, *11*(3), 266.

Denden, M., Tlili, A., Essalmi, F., & Jemni, M. (2018). Does personality affect students' perceived preferences for game elements in gamified learning environments?. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)* (pp. 111-115). IEEE.

Denisova, A., & Cairns, P. (2015a). Adaptation in digital games: The effect of challenge adjustment on player performance and experience. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (pp. 97-101). ACM.

Denisova, A., & Cairns, P. (2015b). First person vs. third person perspective in digital games: do player preferences affect immersion?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 145-148). ACM.

Denisova, A., Nordin, A. I., & Cairns, P. (2016). The Convergence of Player Experience Questionnaires. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (pp. 33-37). ACM.

Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and brain sciences*, *22*(3), 491-517.

Derbali, L., Chalfoun, P., & Frasson, C. (2011). A theoretical and empirical approach in assessing motivational factors: From serious games to an ITS. In *Twenty-Fourth International FLAIRS Conference*.

Derbali, L., & Frasson, C. (2010). Prediction of players motivational states using electrophysiological measures during serious game play. In *2010 10th IEEE International Conference on Advanced Learning Technologies* (pp. 498-502). IEEE.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In *CHI'04 extended abstracts on Human factors in computing systems*(pp. 1509-1512). ACM.

Dey, A., Chatburn, A., & Billinghurst, M. (2019). Exploration of an EEG-based cognitively adaptive training system in virtual reality. In *2019 ieee conference on virtual reality and 3d user interfaces (vr)* (pp. 220-226). IEEE.

Di Bono, M. G., & Zorzi, M. (2008). Decoding Cognitive States from fMRI Data Using Support Vector Regression. *PsychNology Journal*, *6*(2), 189-201.

Di Domenico, S. I., & Ryan, R. M. (2017). The emerging neuroscience of intrinsic motivation: a new frontier in self-determination research. *Frontiers in Human Neuroscience*, *11*, 145.

Dickey, M. D. (2007). Game design and learning: A conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development*, *55*(3), 253-273.

Dickinson, A., & Dearing, M. F. (1979). Appetitive-aversive interactions and inhibitory processes. In Dickinson, A., & Boakes, R. A. (Eds.), *Mechanisms of learning and* motivation (pp. 203-231). New Jersey: Hillsdale.

Dickmeis, A., & Roe, K. (2019). Genres matter: Video games as predictors of physical aggression among adolescents. *Communications*, *44*(1), 105-129.

Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affect. *Journal of personality and social psychology*, *47*(5), 1105.

Diener, E., & Larsen, R. J. (1993). The experience of emotional well-being. In M. Lewis & J. M. Haviland (Eds.), Handbook of emotions (pp. 405–415). New York: Guilford.

Diener, E., Sandvik, E., & Pavot, W. (2009). Happiness is the frequency, not the intensity, of positive versus negative affect. In *Assessing well-being* (pp. 213-231). Springer, Dordrecht.

Dietrich, A. (2004). Neurocognitive mechanisms underlying the experience of flow. *Conscious Cogn.* 13, 746–761. doi: 10.1016/j.concog.2004.07.002

Dietterich, T. G. (2000). An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine learning*, *40*(2), 139-157.

Ding, C., & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology*, *3*(02), 185-205.

Doherty, S. M., Keebler, J. R., Davidson, S. S., Palmer, E. M., & Frederick, C. M. (2018). Recategorization of Video Game Genres. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 2099-2103). Sage CA: Los Angeles, CA: SAGE Publications.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron*, *42*(5), 855-863.

Donahue, J., Krähenbühl, P., & Darrell, T. (2016). Adversarial feature learning. *arXiv preprint arXiv:1605.09782*.

Doñamayor, N., Schoenfeld, M. A., & Münte, T. F. (2012). Magneto-and electroencephalographic manifestations of reward anticipation and delivery. *Neuroimage*, *62*(1), 17-29.

Dong, J. G. (2016). The role of heart rate variability in sports physiology. *Experimental and therapeutic medicine*, *11*(5), 1531-1536.

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

Drachen, A., Sifa, R., Bauckhage, C., & Thurau, C. (2012). Guns, swords and data: Clustering of player behavior in computer games in the wild. In *2012 IEEE conference on Computational Intelligence and Games (CIG)* (pp. 163-170). IEEE.

Dreher, J. C., Kohn, P., Kolachana, B., Weinberger, D. R., & Berman, K. F. (2009). Variation in dopamine genes influences responsivity of the human reward system. *Proceedings of the National Academy of Sciences*, *106*(2), 617-622.

Droit-Volet, S., & Gil, S. (2009). The time–emotion paradox. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1943-1953.

Drucker, H., & Cortes, C. (1996). Boosting decision trees. In *Advances in Neural Information Processing Systems*, *8*, 479-485, MIT Press.

Ducheneaut, N., Yee, N., Nickell, E., & Moore, R. J. (2006). Alone together?: exploring the social dynamics of massively multiplayer online games. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*(pp. 407-416). ACM.

Dunn, B. R., Hartigan, J. A., & Mikulas, W. L. (1999). Concentration and mindfulness meditations: unique forms of consciousness?. *Applied psychophysiology and biofeedback*, *24*(3), 147-165.

Dunn, R. A., & Guadagno, R. E. (2012). My avatar and me–Gender and personality predictors of avatar-self discrepancy. *Computers in Human Behavior*, *28*(1), 97-106.

Dunne, J. D., Thompson, E., & Schooler, J. (2019). Mindful meta-awareness: Sustained and non-propositional. *Current Opinion in Psychology*, *28*, 307-311.

Eastwood, J. D., Cavaliere, C., Fahlman, S. A., & Eastwood, A. E. (2007). A desire for desires: Boredom and its relation to alexithymia. *Personality and Individual Differences*, *42*(6), 1035-1045.

Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The unengaged mind: Defining boredom in terms of attention. *Perspectives on Psychological Science*, *7*(5), 482-495.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of personality and social psychology*, *83*(6), 1441.

Eichenlaub, J. B., Nicolas, A., Daltrozzo, J., Redouté, J., Costes, N., & Ruby, P. (2014). Resting brain activity varies with dream recall frequency between subjects. *Neuropsychopharmacology*, *39*(7), 1594-1602.

Ekman, P. (1970). Universal facial expressions of emotion. *California Mental Health Research Digest*, *8*(4), 151-158.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*(4616), 1208 –1210.

Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and emotion*, *30*(2), 111-116.

Elpidorou, A. (2014). The bright side of boredom. *Frontiers in psychology*, *5*, 1245.

Emerson, R. M. (1962). Power-dependence relations. *American sociological review*, 31-41.

Emes, C. E. (1997). Is Mr Pac Man eating our children? A review of the effect of video games on children. *The Canadian Journal of Psychiatry*, *42*(4), 409-414.

Emmons, R. A. (1999). *The psychology of ultimate concerns: Motivation and spirituality in personality*. New York: Guilford Press.

Engeser, S. (2012). *Advances in flow research*. Springer Science & Business Media.

Engeser, S., & Schiepe-Tiska, A. (2012). Historical lines and an overview of current research on flow. In S. Engeser (Ed.), *Advances in Flow Research* (pp. 1-22). New York: Springer

Epstein, L. H. (1990). Perception of activity in the zygomaticus major and corrugator supercilii muscle regions. *Psychophysiology*, *27*(1), 68-72.

Ermi, L., & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. *Worlds in play: International perspectives on digital games research*, *37*(2), 37-53.

Estabrooks, A., Jo, T., & Japkowicz, N. (2004). A multiple resampling method for learning from imbalanced data sets. *Computational intelligence*, *20*(1), 18-36.

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*, 223.

Eysenck, M. W. (1981). Learning, memory and personality. In *A model for personality* (pp. 169-209). Springer, Berlin, Heidelberg.

Faber, N. M., & Rajko, R. (2007). How to avoid over-fitting in multivariate calibration—The conventional validation approach and an alternative. *Analytica Chimica Acta*, *595*(1-2), 98-106.

Fabricatore, C. (2007). Gameplay and Game Mechanics: A Key to Quality in Videogames. In *ENLACES (MINEDUC Chile) OECD Expert Meeting on Videogames and Education*. Santiago de Chile, Chile.

Fairclough, S. H., Gilleade, K., Ewing, K. C., & Roberts, J. (2013). Capturing user engagement via psychophysiology: measures and mechanisms for biocybernetic adaptation. *International Journal of Autonomous and Adaptive Communications Systems*, *6*(1), 63-79.

Faisal, A., & Peltoniemi, M. (2015). Establishing video game genres using data-driven modeling and product databases. *Games and Culture*, *13*(1), 20-43.

Fang, X., & Zhao, F. (2009). Sensation seeking, self forgetfulness, and computer game enjoyment. In *Symposium on Human Interface* (pp. 632-641). Springer, Berlin, Heidelberg.

Fang, X., & Zhu, M. (2011). Extraversion and computer game play: who plays what games?. In *International Conference on Human-Computer Interaction* (pp. 659-667). Springer, Berlin, Heidelberg.

Farahi, S. M. M., Ebrahimabad, M. J. A,, Gorji, A., Bigdeli, I., & Farahi, S. M. M. M. (2019). Neuroticism and frontal EEG asymmetry correlated with dynamic facial emotional processing in adolescents. *Frontiers in psychology*, *10*, 175.

Farmer, R., & Sundberg, N. D. (1986). Boredom proneness--the development and correlates of a new scale. *Journal of personality assessment*, *50*(1), 4-17.

Farooq, A., Evreinov, G., & Raisamo, R. (2019). Enhancing Multimodal Interaction for Virtual Reality Using Haptic Mediation Technology. In *International Conference on Applied Human Factors and Ergonomics* (pp. 377-388). Springer, Cham.

Federoff, M. A. (2002). *Heuristics and usability guidelines for the creation and evaluation of fun in video games* (Doctoral dissertation). Indiana University, Bloomington.

Feigenson, N. (2006). Too real? The future of virtual reality evidence. *Law & Policy*, *28*(2), 271-293.

Feist, G. J. (1999). Autonomy and independence. *Encyclopedia of creativity*, *1*, 157-163.

Feldon, D. F. (2007). Cognitive load and classroom teaching: The double-edged sword of automaticity. *Educational Psychologist*, *42*(3), 123-137.

Feng, H., Golshan, H. M., & Mahoor, M. H. (2018). A wavelet-based approach to emotion classification using EDA signals. *Expert Systems with Applications*, *112*, 77-86.

Ferguson, C. J. (2010). Blazing angels or resident evil? Can violent video games be a force for good?. *Review of general psychology*, *14*(2), 68-81.

Ferguson, C. J., & Olson, C. K. (2013). Friends, fun, frustration and fantasy: Child motivations for video game play. *Motivation and Emotion*, *37*(1), 154-164.

Fernandes, L. V., Castanho, C. D., & Jacobi, R. P. (2018). A Survey on Game Analytics in Massive Multiplayer Online Games. In *2018 17th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)* (pp. 21-2109). IEEE.

Fernández, H. D. B., Mikami, K., & Kondo, K. (2017). Adaptable Game Experience Based on Player's Performance and EEG. In *2017 Nicograph International (NicoInt)* (pp. 1-8). IEEE.

Ferster, C.B., & Skinner, B.F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.

Finneran, C. M., & Zhang, P. (2003). A person–artefact–task (PAT) model of flow antecedents in computer-mediated environments. *International Journal of Human-Computer Studies*, *59*(4), 475-496.

Fisher, C. D. (1993). Boredom at work: A neglected concept. *Human Relations*, *46*(3), 395-417.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition (2nd ed.)*. New York: McGraw-Hill

Fong, C. J., Zaleski, D. J., & Leach, J. K. (2015). The challenge–skill balance and antecedents of flow: A meta-analytic investigation. *The Journal of Positive Psychology*, *10*(5), 425-446.

Forster, S. (2013). Distraction and mind-wandering under load. *Frontiers in psychology*, *4*, 283.

Fortin-Côté, A., Chamberland, C., Parent, M., Tremblay, S., Jackson, P., Beaudoin-Gagnon, N., ... & Lefebvre, L. (2018). Predicting Video Game Players' Fun from Physiological and Behavioural Data. In *Future of Information and Communication Conference* (pp. 479-495). Springer, Cham.

Fox, N. A. (1991). If it's not left, it's right. Electroencephalograph asymmetry and the development of emotion. *American Psychologist*, *46*, 863–872.

Francis, P. J., & Wills, B. J. (1999). Introduction to principal components analysis. *arXiv preprint astro-ph/9905079*.

Freeman, N., & Muraven, M. (2010). Don't interrupt me! Task interruption depletes the self's limited resources. *Motivation and emotion*, *34*(3), 230-241.

Freimuth, M., Moniz, S., & Kim, S. R. (2011). Clarifying exercise addiction: differential diagnosis, co-occurring disorders, and phases of addiction. *International journal of environmental research and public health*, *8*(10), 4069-4081.

Freund, Y., & Schapire, R. E. (1996). Experiments with a new boosting algorithm, *Machine Learning: Proceedings of the Thirteenth International Conference*, 148–156.

Freund, Y., Schapire, R. E., & Abe, N. (1999). A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, *14*(771-780), 1612.

Frijda, N. H. (1986). *The emotions*. Cambridge University Press.

Fu, F. L., Su, R. C., & Yu, S. C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, *52*(1), 101-112.

Fu, Y., Cao, L., Guo, G., & Huang, T. S. (2008). Multiple feature fusion by subspace learning. In *Proceedings of the 2008 international conference on Content-based image and video retrieval* (pp. 127-134).

Fukuda, K., & Matsunaga, K. (1983). Changes in blink rate during signal discrimination tasks. *Japanese Psychological Research*, *25*(3), 140-146.

Fukuda, K., Stern, J. A., Brown, T. B., & Russo, M. B. (2005). Cognition, blinks, eye-movements, and pupillary movements during performance of a running memory task. *Aviation, space, and environmental medicine*, *76*(7), C75-C85.

Fullagar, C. J., & Kelloway, E. K. (2009). Flow at work: An experience sampling approach. *Journal of occupational and organizational psychology*, *82*(3), 595-615.

Fullagar, C. J., Knight, P. A., & Sovern, H. S. (2013). Challenge/skill balance, flow, and performance anxiety. *Applied Psychology*, *62*(2), 236-259.

Funk, J. B. (1993). Reevaluating the impact of video games. *Clinical pediatrics*, *32*(2), 86-90.

Fuster, H., Chamarro, A., Carbonell, X., & Vallerand, R. J. (2014). Relationship between passion and motivation for gaming in players of massively multiplayer online role-playing games. *Cyberpsychology, Behavior, and Social Networking*, *17*(5), 292-297.

Gabbiadini, A., & Greitemeyer, T. (2017). Uncovering the association between strategy video games and self-regulation: A correlational study. *Personality and Individual Differences*, *104*, 129-136.

Gackenbach, J., & Bown, J. (2011). Mindfulness and video game play: A preliminary inquiry. *Mindfulness*, *2*(2), 114-122.

Gaggioli, A., Cipresso, P., Serino, S., & Riva, G. (2013). Psychophysiological correlates of flow during daily activities. *Annual Review of Cybertherapy and Telemedicine*, *191*, 65-69.

Gaillard, A.W.K. (1993). Comparing the concepts of mental load and stress. *Ergonomics*, *36*(9), 991–1006.

Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., & Zuo, S. (2019). EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation. *IEEE transactions on neural networks and learning systems*, *30*(9), 2755-2763.

Gargiulo, G., Bifulco, P., Cesarelli, M., McEwan, A., Moeinzadeh, H., O'Loughlin, A., ... & Thiagalingam, A. (2018). On the Einthoven Triangle: A Critical Analysis of the Single Rotating Dipole Hypothesis. *Sensors*, *18*(7), 2353.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & gaming*, *33*(4), 441-467.

Gastorf, J. W. (1981). Physiologic reaction of Type A's to objective and subjective challenge. *Journal of Human Stress*, *7*(1), 16-27.

Geisler, B. (2002). *An Empirical Study of Machine Learning Algorithms Applied to Modeling Player Behavior in a" First Person Shooter" Video Game* (Doctoral dissertation). University of Wisconsin-Madison, Madison, WI, United States.

Gelbart, M. A., Snoek, J., & Adams, R. P. (2014). Bayesian optimization with unknown constraints. *arXiv preprint arXiv:1403.5607*.

Gentil, A. F., Eskandar, E. N., Marci, C. D., Evans, K. C., & Dougherty, D. D. (2009). Physiological responses to brain stimulation during limbic surgery: further evidence of anterior cingulate modulation of autonomic arousal. *Biological psychiatry*, *66*(7), 695-701.

George, J. M. (2011). Dual tuning: A minimum condition for understanding affect in organizations. *Organizational Psychology Review*, *1*, 147-164.

Ghani, J. A. (1995). Flow in human computer interactions: test of a model. *Human factors in information systems: Emerging theoretical bases*, 291-311.

Ghani, J. A., & Deshpande, S. P. (1994). Task characteristics and the experience of optimal flow in human—computer interaction. *The Journal of psychology*, *128*(4), 381-391.

Ghani, J. A., Supnick, R., & Rooney, P. (1991). The Experience of Flow in Computer-mediated and in Face-to-face Groups. In *ICIS* (Vol. 91, No. 6, pp. 229-237).

Ghorai, S., & Ghosh, D. (2017). Arrhythmia classification by nonlinear kernel-based ECG signal modeling. *Computer, Communication and Electrical Technology. CRC Press*, 325-330.

Giakoumis, D., Vogiannou, A., Kosunen, I., Moustakas, K., Tzovaras, D., & Hassapis, G. (2010). Identifying Psychophysiological Correlates of Boredom and Negative Mood Induced during HCI. In *B-Interface* (pp. 3-12).

Gilleade, K. M., & Dix, A. (2004). Using frustration in the design of adaptive videogames. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 228-232). ACM.

Giluk, T. L. (2009). Mindfulness, Big Five personality, and affect: A meta-analysis. *Personality and Individual Differences*, *47*(8), 805-811.

Gilyadov, A. (2017). PS4 Hits 70 Million Sales, PSVR Hits 2 Million. Retrieved from https://uk.ign.com/articles/2017/12/07/ps4-hits-70-million-sales-psvr-hits-2-million. Accessed on 27/12/2019.

Glas, R. (2015). Vicarious play: Engaging the viewer in Let's Play videos. *Empedocles: European Journal for the Philosophy of Communication*, *5*(1-2), 81-86.

Goetz, T., & Frenzel, A. C. (2006). Phänomenologie schulischer Langeweile [Phenomenology of boredom at school]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *38*(4), 149–153. doi:10.1026/0049-8637.38.4.149.

Goetz, T., Frenzel, A. C., Hall, N. C., Nett, U. E., Pekrun, R., & Lipnevich, A. A. (2014). Types of boredom: An experience sampling approach. *Motivation and Emotion*, *38*(3), 401-419.

Gola, M., Magnuski, M., Szumska, I., & Wróbel, A. (2013). EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects. *International Journal of Psychophysiology*, *89*(3), 334-341.

Golbeck, J., Robles, C., & Turner, K. (2011). Predicting personality with social media. In *CHI'11 extended abstracts on human factors in computing systems* (pp. 253-262). ACM.

Goldberg, I. I., Harel, M., and Malach, R. (2006). When the brain loses its self: prefrontal inactivation during sensorimotor processing. *Neuron* 50, 329–339. doi: 10.1016/j.neuron.2006.03.015

Goldberg, Y. K., Eastwood, J. D., LaGuardia, J., & Danckert, J. (2011). Boredom: An emotional experience distinct from apathy, anhedonia, or depression. *Journal of Social and Clinical Psychology*, *30*(6), 647-666.

Goldstein, D. S., Bentho, O., Park, M. Y., & Sharabi, Y. (2011). Low‐frequency power of heart rate variability is not a measure of cardiac sympathetic tone but may be a measure of modulation of cardiac autonomic outflows by baroreflexes. *Experimental physiology*, *96*(12), 1255-1261.

Goldstein, R., Walrath, L. C., Stern, J. A., & Strock, B. D. (1985). Blink activity in a discrimination task as a function of stimulus modality and schedule of presentation. *Psychophysiology*, *22*(6), 629-635.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint arXiv:1406.2661*.

Gordon, A., Wilkinson, R., McGown, A., & Jovanoska, S. (1997). The psychometric properties of the Boredom Proneness Scale: An examination of its validity. *Psychological Studies*.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in personality*, *37*(6), 504-528.

Grace, L. (2005). Game type and game genre. Retrieved from http://www.lgrace.com/documents/game_types_and_genres.pdf.

Graham, L. T., & Gosling, S. D. (2013). Personality profiles associated with different motivations for playing World of Warcraft. *Cyberpsychology, Behavior, and Social Networking*, *16*(3), 189-193.

Granic, I., Lobel, A., and Engels, R. C. (2014). The benefits of playing video games. *Am. Psychol.* 69, 66. doi: 10.1037/a0034857.

Gratton, G. (1998). Dealing with artifacts: The EOG contamination of the event-related brain potential. *Behavior Research Methods, Instruments, & Computers*, *30*(1), 44-53.

Gray, J. A. (1994). Three fundamental emotion systems. In Ekman, P., Davidison, R.J. (Eds.), *The Nature of Emotion: Fundamental Questions* (pp. 243–247). New York, NY, US: Oxford University Press.

Gray, J. A., & McNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system* (No. 33). Oxford university press.

Graziano, W. G., & Tobin, R. M. (2009). Agreeableness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 46-61). New York, NY, US: The Guilford Press.

Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person× situation perspective. *Journal of personality and social psychology*, *93*(4), 583.

Greco, A., Valenza, G., Citi, L., & Scilingo, E. P. (2016). Arousal and valence recognition of affective sounds based on electrodermal activity. *IEEE Sensors Journal*, *17*(3), 716-725.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534.

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, *14*(4), 311-327.

Greenaway, K. H., Cruwys, T., Haslam, S. A., & Jetten, J. (2016). Social identities promote well-being because they satisfy global psychological needs. *European Journal of Social Psychology*, *46*(3), 294-307.

Greenberg, B. S., Sherry, J., Lachlan, K., Lucas, K., & Holmstrom, A. (2010). Orientations to video games among gender and age groups. *Simulation & Gaming*, *41*(2), 238-259.

Greenson, R. (1953). On boredom. *Journal of the American Psychoanalytic Association*, *1*, 7–21.

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use?. *Psychological Bulletin*, *83*(2), 314.

Griffiths, M. D. (2000). Does Internet and computer" addiction" exist? Some case study evidence. *CyberPsychology and Behavior*, *3*(2), 211-218.

Griffiths, M. D. (2002). The educational benefits of videogames. *Education and health*, *20*(3), 47-51.

Griffiths, M. D., & Hunt, N. (1998). Dependence on computer games by adolescents. *Psychological reports*, *82*(2), 475-480.

Griffiths, M. D., & Nuyens, F. (2017). An overview of structural characteristics in problematic video game playing. *Current addiction reports*, *4*(3), 272-283.

Griffiths, M. D., & Wood, R. T. (2000). Risk factors in adolescence: The case of gambling, videogame playing, and the Internet. *Journal of gambling studies*, *16*(2-3), 199-225.

Griss, P., Tolvanen-Laakso, H. K., Merilainen, P., & Stemme, G. (2002). Characterization of micromachined spiked biopotential electrodes. *IEEE Transactions on Biomedical Engineering*, *49*(6), 597-604.

Grissmann, S., Faller, J., Scharinger, C., Spüler, M., & Gerjets, P. (2017). Electroencephalography based analysis of working memory load and affective valence in an n-back task with emotional stimuli. *Frontiers in Human Neuroscience*, *11*, 616.

Grodal, T. (2000). Video games and the pleasures of control. *Media entertainment: The psychology of its appeal*, 197-213.

Gunn, S. R. (1998). Support vector machines for classification and regression. *ISIS technical report*, *14*(1), 5-16.

Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). KNN model-based approach in classification. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"* (pp. 986-996). Springer, Berlin, Heidelberg.

Guo, X., Yin, Y., Dong, C., Yang, G., & Zhou, G. (2008). On the class imbalance problem. In *2008 Fourth international conference on natural computation* (Vol. 4, pp. 192-201). IEEE.

Guo, Z., Xiao, L., Chanyoung, S., and Lai, Y. (2012). *Flow Experience and Continuance Intention Toward Online Learning: An Integrated Framework.* Available at: http://aisel.aisnet.org/icis2012/ proceedings/ISCurriculum/6/

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4259–4264. doi: 10.1073/pnas.071043098

Gut, M., Urbanik, A., Forsberg, L., Binder, M., Rymarczyk, K., Sobiecka, B., ... & Grabowska, A. (2007). Brain correlates of right-handedness. *Acta neurobiologiae experimentalis*, *67*(1), 43.

Gyurkovics, M., Kotyuk, E., Katonai, E. R., Horvath, E. Z., Vereczkei, A., & Szekely, A. (2016). Individual differences in flow proneness are linked to a dopamine D2 receptor gene variant. *Consciousness and cognition*, *42*, 1-8.

Hafeez, T., Saeed, S. M. U., Arsalan, A., Anwar, S. M., Ashraf, M. U., & Alsubhi, K. (2021). EEG in game user analysis: A framework for expertise classification during gameplay. *bioRxiv*.

Hagemann, D., Hewig, J., Walter, C., and Naumann, E. (2008). Skull thickness and magnitude of EEG alpha activity. *Clinical Neurophysiology*, *119*, 1271–1280. doi: 10.1016/j.clinph.2008.02.010

Hagemann, D., Naumann, E., Lürken, A., Becker, G., Maier, S., & Bartussek, D. (1999). EEG asymmetry, dispositional mood and personality. *Personality and Individual Differences*, *27*(3), 541-568.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–151.

Hallford, N., & Hallford, J. (2001). *Swords & circuitry: A designer's guide to computer role playing games*. Roseville, CA: Prima Tech.

Hamari, J., Alha, K., Järvelä, S., Kivikangas, J. M., Koivisto, J., & Paavilainen, J. (2017). Why do players buy in-game content? An empirical study on concrete purchase motivations. *Computers in Human Behavior*, *68*, 538-546.

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional flow scale-2. *Computers in Human Behavior*, *40*, 133-143.

Hamari, J., & Sjöblom, M. (2017). What is eSports and why do people watch it?. *Internet research*, *27*(2), 211-232.

Hamel, R., Savoie, F. A., Lacroix, A., Whittingstall, K., Trempe, M., & Bernier, P. M. (2018). Added value of money on motor performance feedback: Increased left central beta-band power for rewards and fronto-central theta-band power for punishments. *Neuroimage*, *179*, 63-78.

Hamill, R. W., & Shapiro, R. E. (2004). Peripheral autonomic nervous system. In *Primer on the autonomic nervous system* (pp. 20-28). Academic press.

Hammoud, S., Karam, R., Mourad, R., Saad, I., & Kurdi, M. (2019). Stress and Heart Rate Variability during University Final Examination among Lebanese Students. *Behavioral Sciences*, *9*(1), 3.

Hansen, L. K., & Salamon, P. (1990). Neural network ensembles. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *10*, 993-1001.

Harbluk, J. L., Noy, Y. I., & Eizenman, M. (2002). *The impact of cognitive distraction on driver visual behaviour and vehicle control* (No. TP# 13889 E).

Hardie, E., & Tee, M. Y. (2007). Excessive Internet use: The role of personality, loneliness and social support networks in Internet Addiction. *Australian Journal of Emerging Technologies & Society*, *5*(1).

Hardy, S., Kern, A., Dutz, T., Weber, C., Göbel, S., & Steinmetz, R. (2014). What makes games challenging?: Considerations on how to determine the challenge posed by an exergame for balance training. In *Proceedings of the 2014 ACM International Workshop on Serious Games* (pp. 57-62). ACM.

Harmat, L., de Manzano,Ö, Theorell, T., Högman, L., Fischer, H., and Ullén, F. (2015). Physiological correlates of the flow experience during computer game playing. *Int. J. Psychophysiol.* 97, 1–7. doi: 10.1016/j.ijpsycho.2015.05.001

Harmon-Jones, E. (2003). Clarifying the emotive functions of asymmetrical frontal cortical activity. *Psychophysiology*, *40*(6), 838-848.

Harmon-Jones, E., & Allen, J. J. (1998). Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of personality and social psychology*, *74*(5), 1310.

Harris, D. J., Vine, S. J., and Wilson, M. R. (2017a). Is flow really effortless? The complex role of effortful attention. *Sport Exerc. Perform. Psychol.* 6:103. doi: 10.1037/spy0000083

Harris, D. J., Vine, S. J., & Wilson, M. R. (2017b). Neurocognitive mechanisms of the flow state. *Progress in Brain Research*.

Harris, M. B. (2000). Correlates and Characteristics of Boredom Proneness and Boredom 1. *Journal of Applied Social Psychology*, *30*(3), 576-598.

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

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

He, H., & Ma, Y. (2013). Imbalanced learning: foundations, algorithms, and applications. John Wiley & Sons, New York.

He, J., Becic, E., Lee, Y. C., & McCarley, J. S. (2011). Mind wandering behind the wheel: performance and oculomotor correlates. *Human factors*, *53*(1), 13-21.

Heere, B. (2018). Embracing the sportification of society: defining e-sports through a polymorphic view on sport. *Sport Management Review*, *21*(1), 21-24.

Hefner, D., Klimmt, C., & Vorderer, P. (2007). Identification with the player character as determinant of video game enjoyment. In *International conference on entertainment computing* (pp. 39-48). Springer, Berlin, Heidelberg.

Hejjel, L., & Roth, E. (2004). What is the adequate sampling interval of the ECG signal for heart rate variability analysis in the time domain?. *Physiological Measurement*, *25*(6), 1405.

Heller, K., Bullerjahn, C., & von Georgi, R. (2015). The relationship between personality traits, flow-experience, and different aspects of practice behavior of amateur vocal students. *Frontiers in psychology*, *6*, 1901.

Heller, W. (1990). The neuropsychology of emotion: Developmental patterns and implications for psychopathology. In N. Stein, B. L. Leventhal, T. Trabasso (Eds.), *Psychological and Biological Approaches to Emotion* (pp. 167–211). NJ: Hillsdale.

Hellman, M., Schoenmakers, T. M., Nordstrom, B. R., & van Holst, R. J. (2013). Is there such a thing as online video game addiction? A cross-disciplinary review. *Addiction Research & Theory*, *21*(2), 102-112.

Heo, J., Lee, Y., Pedersen, P. M., and McCormick, B. P. (2010). Flow experience in the daily lives of older adults: an analysis of the interaction between flow, individual differences, serious leisure, location, and social context. *Can. J. Aging* 29, 411–423. doi: 10.1017/S0714980810000395

Heo, J., Yoon, H., & Park, K. S. (2017). A Novel Wearable Forehead EOG Measurement System for Human Computer Interfaces. *Sensors*, *17*(7), 1485.

Hertweck, S., Weber, D., Alwanni, H., Unruh, F., Fischbach, M., Latoschik, M. E., & Ball, T. (2019). Brain activity in virtual reality: Assessing signal quality of high-resolution eeg while using head-mounted displays. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 970-971). IEEE.

Hettinger, L. J., & Riccio, G. E. (1992). Visually induced motion sickness in virtual environments. *Presence: Teleoperators & Virtual Environments*, *1*(3), 306-310.

Hidi, S. (2000). An interest researcher's perspective: The effects of extrinsic and intrinsic factors on motivation. In *Intrinsic and extrinsic motivation* (pp. 309-339). Academic Press.

Hild, J., Voit, M., & Peinsipp-Byma, E. (2020). Estimating Immersed User States from Eye Movements: A Survey. In *International Conference on Human-Computer Interaction* (pp. 337-342). Springer, Cham.

Hilgard, J., Engelhardt, C. R., & Bartholow, B. D. (2013). Individual differences in motives, preferences, and pathology in video games: the gaming attitudes, motives, and experiences scales (GAMES). *Frontiers in psychology*, *4*, 608.

Hill, A. B., & Perkins, R. E. (1985). Towards a model of boredom. *British Journal of Psychology*, *76*(2), 235-240.

Hilla, Y., von Mankowski, J., Föcker, J., & Sauseng, P. (2020). Faster Visual Information Processing in Video Gamers Is Associated With EEG Alpha Amplitude Modulation. *Frontiers in Psychology*, *11*, 3333.

Hinrichs, H., Scholz, M., Baum, A. K., Kam, J. W., Knight, R. T., & Heinze, H. J. (2020). Comparison between a wireless dry electrode EEG system with a conventional wired wet electrode EEG system for clinical applications. *Scientific reports*, *10*(1), 1-14.

Hirsh, J. B., Kang, S. K., & Bodenhausen, G. V. (2012). Personalized persuasion: Tailoring persuasive appeals to recipients' personality traits. *Psychological science*, *23*(6), 578-581.

Hodent, C. (2016). The Gamer's Brain, Part 2: UX of Onboarding and Player Engagement. In: Speech Presented at *Game Developer Conference 2016* (n.d.). Retrieved April 30, 2019, from https://celiahodent.com/the-gamers-brain-part-2-gdc16/.

Hoffman, D. L., & Novak, T. P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of marketing*, *60*(3), 50-68.

Hoffman, D. L., & Novak, T. P. (2009). Flow online: lessons learned and future prospects. *Journal of interactive marketing*, *23*(1), 23-34.

Hogervorst, M. A., Brouwer, A. M., & Van Erp, J. B. (2014). Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Frontiers in neuroscience*, *8*, 322.

Holzwarth, V., Schneider, J., Handali, J., Gisler, J., Hirt, C., Kunz, A., & vom Brocke, J. (2021). Towards estimating affective states in virtual reality based on behavioral data. *Virtual reality*, 1-14.

Homan, R. W. (1988). The 10-20 electrode system and cerebral location. *American Journal of EEG Technology*, *28*(4), 269-279.

Hooker, R., Wasko, M., Paradice, D., Teigland, R., & Hofacker, C. (2019). Beyond gaming: Linking flow, brand attitudes, and purchase intent in realistic and emergent three-dimensional virtual environments. *Information Technology & People*, *32*(6), 1397-1422. doi: https://doi.org/10.1108/ITP-11-2017-0407.

Horslen, B. C., & Carpenter, M. G. (2011). Arousal, valence and their relative effects on postural control. *Experimental brain research*, *215*(1), 27.

Hotho, A., Maedche, A., & Staab, S. (2002). Ontology-based text document clustering. *KI*, *16*(4), 48-54.

Houle, M. S., & Billman, G. E. (1999). Low-frequency component of the heart rate variability spectrum: a poor marker of sympathetic activity. *American Journal of Physiology-Heart and Circulatory Physiology*, *276*(1), H215-H223.

Hsu, C. C., & Chen, E. C. H. (2009). Exploring the Elements and Design Criteria of Massively-Multiplayer Online Role-Playing Game (MMORPG) Interfaces. In *International Conference on Human-Computer Interaction* (pp. 325-334). Springer, Berlin, Heidelberg.

Hsu, C. L., & Lu, H. P. (2004). Why do people play on-line games? An extended TAM with social influences and flow experience. *Information & management*, *41*(7), 853-868.

Hsu, S. H., Wen, M. H., & Wu, M. C. (2007). Exploring design features for enhancing players' challenge in strategy games. *CyberPsychology & Behavior*, *10*(3), 393-397.

Hsu, S. H., Wen, M. H., & Wu, M. C. (2009). Exploring user experiences as predictors of MMORPG addiction. *Computers & Education*, *53*(3), 990-999.

Hu, L. Y., Huang, M. W., Ke, S. W., & Tsai, C. F. (2016). The distance function effect on k-nearest neighbor classification for medical datasets. *SpringerPlus*, *5*(1), 1304.

Huang, H. Y., & Lo, P. C. (2009). EEG dynamics of experienced Zen meditation practitioners probed by complexity index and spectral measure. *Journal of medical engineering & technology*, *33*(4), 314-321.

Huang, V. K., Pang, Z., Chen, C. J. A., & Tsang, K. F. (2018). New trends in the practical deployment of industrial wireless: From noncritical to critical use cases. *IEEE Industrial Electronics Magazine*, *12*(2), 50-58.

Huang, Y., Zhou, R., Cui, H., Wu, M., Wang, Q., Zhao, Y., & Liu, Y. (2015). Variations in resting frontal alpha asymmetry between high-and low-neuroticism females across the menstrual cycle. *Psychophysiology*, *52*(2), 182-191.

Hubert-Wallander, B., Green, C. S., & Bavelier, D. (2011). Stretching the limits of visual attention: the case of action video games. *Wiley interdisciplinary reviews: cognitive science*, *2*(2), 222-230.

Huizinga, J. (1950). *Homo ludens: A study of the play element in culture*. Boston: Beacon.

Hull, D. C., Williams, G. A., & Griffiths, M. D. (2013). Video game characteristics, happiness and flow as predictors of addiction among video game players: A pilot study. *Journal of behavioral addictions*, *2*(3), 145-152.

Hunicke, R. (2005). The case for dynamic difficulty adjustment in games. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 429-433). ACM.

Hurd, S. (2005). Autonomy and the distance language learner. In Holmberg et al. (eds.), 1–19.

Iacovides, I., Cox, A., Kennedy, R., Cairns, P., & Jennett, C. (2015). Removing the HUD: the impact of non-diegetic game elements and expertise on player involvement. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (pp. 13-22). ACM.

IJsselsteijn, W., De Kort Y., Poels, K., Jurgelionis, A., Belotti, F. (2007) Characterising and measuring user experiences. In: *Proceedings of the international Conference on Advances in Computer Entertainment Technology* (Salzburg, Austria, 13–15 June 2007). ACE '07, Vol. 203, ACM, New York.

IJsselsteijn, W., Poels, K., & De Kort, Y. A. (2008). The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games. *TU Eindhoven, Eindhoven, The Netherlands*, 42.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. *Journal of sleep research*, *15*(1), 47-53.

Ip, B., & Adams, E. (2002). From casual to core: A statistical mechanism for studying gamer dedication. *Available online at* https://www.gamasutra.com/view/feature/131397/from_casual_to_core_a_statistical_.php.

Ishaque, S., Rueda, A., Nguyen, B., Khan, N., & Krishnan, S. (2020). Physiological Signal Analysis and Classification of Stress from Virtual Reality Video Game. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 867-870). IEEE.

Ishiguro, T., Suzuki, C., Nakakoji, H., Funagira, Y., & Takao, M. (2019). Immersive experience influences eye blink rate during virtual reality gaming. *Polish Psychological Bulletin*, *50*(1), 49-53.

Iversen, S. D., & Iversen, L. L. (2007). Dopamine: 50 years in perspective. *Trends in neurosciences*, *30*(5), 188-193.

Izkara, J. L., Pérez, J., Basogain, X., & Borro, D. (2007). Mobile augmented reality, an advanced tool for the construction sector. In *Proceedings of the 24th W78 conference, Maribor, Slovenia* (pp. 190-202).

Jackson, S. A. (1996). Toward a conceptual understanding of the flow experience in elite athletes. *Research quarterly for exercise and sport*, *67*(1), 76-90.

Jackson, S. A., & Csikszentmihalyi, M. (1999). *Flow in Sports.* Champaign IL: Human Kinetics.

Jackson, S. A., & Eklund, R. C. (2002). Assessing flow in physical activity: the flow state scale-2 and dispositional flow scale-2. *Journal of sport & exercise psychology*, *24*(2).

Jackson, S. A., Ford, S. K., Kimiecik, J. C., & Marsh, H. W. (1998). Psychological correlates of flow in sport. *Journal of Sport and exercise Psychology*, *20*(4), 358-378.

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of sport and exercise psychology*, *18*(1), 17-35.

Jackson, S. A., Thomas, P. R., Marsh, H. W., & Smethurst, C. J. (2001). Relationships between flow, self-concept, psychological skills, and performance. *Journal of applied sport psychology*, *13*(2), 129-153.

Jackson, S. A., Ford, S. K., Kimiecik, J. C., & Marsh, H. W. (1998). Psychological correlates of flow in sport. *Journal of Sport and exercise Psychology*, *20*(4), 358-378.

Jackson, S. A., Thomas, P. R., Marsh, H. W., & Smethurst, C. J. (2001). Relationships between flow, self-concept, psychological skills, and performance. *Journal of applied sport psychology*, *13*(2), 129-153.

Janecek, A., Gansterer, W., Demel, M., & Ecker, G. (2008). On the relationship between feature selection and classification accuracy. In *New challenges for feature selection in data mining and knowledge discovery* (pp. 90-105).

Järvelä, S., Ekman, I., Kivikangas, J. M., & Ravaja, N. (2015). Stimulus games. In *Game Research Methods* (pp. 193-205). ETC Press.

Jasper, H. (1958). Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology*, *10*(2), 370-375. doi: https://doi.org/10.1016/0013-4694(58)90053-1

Jatupaiboon, N., Pan-ngum, S., & Israsena, P. (2013). Real-time EEG-based happiness detection system. *The Scientific World Journal*, *2013*.

Jennett, C. I. (2010). *Is Game Immersion Just Another Form of Selective Attention? An Empirical Investigation of Real World Dissociation in Computer Game Immersion* (Doctoral dissertation). University College London, London.

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.

Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *Journal of personality*, *69*(2), 323-362.

Jimenez, C. O. S., Mesa, H. G. A., Rebolledo-Mendez, G., & de Freitas, S. (2011). Classification of cognitive states of attention and relaxation using supervised learning algorithms. In *2011 IEEE International Games Innovation Conference (IGIC)* (pp. 31-34). IEEE.

Jin, S. A. A. (2012). "Toward integrative models of flow": Effects of performance, skill, challenge, playfulness, and presence on flow in video games. *Journal of Broadcasting & Electronic Media*, *56*(2), 169-186.

Johannes, N., Vuorre, M., & Przybylski, A. K. (2021). Video game play is positively correlated with well-being. *Royal Society open science*, *8*(2), 202049.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, *2*(1999), 102-138.

Johnson, D., Gardner, J., & Sweetser, P. (2016). Motivations for videogame play: Predictors of time spent playing. *Computers in Human Behavior*, *63*, 805-812.

Johnson, D., Klarkowski, M., Vella, K., Phillips, C., McEwan, M., & Watling, C. N. (2018). Greater rewards in videogames lead to more presence, enjoyment and effort. *Computers in Human Behavior*, *87*, 66-74.

Johnson, D., Nacke, L. E., & Wyeth, P. (2015). All about that base: differing player experiences in video game genres and the unique case of MOBA games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2265-2274). ACM.

Johnson, D., Wyeth, P., Sweetser, P., & Gardner, J. (2012). Personality, genre and videogame play experience. In *Proceedings of the 4th International Conference on Fun and Games* (pp. 117-120). ACM.

Johnson, J. A., Keiser, H. N., Skarin, E. M., & Ross, S. R. (2014). The Dispositional flow scale–2 as a measure of autotelic personality: An examination of criterion-related validity. *Journal of personality assessment*, *96*(4), 465-470.

Johnson, J. A., & Ostendorf, F. (1993). Clarification of the five-factor model with the Abridged Big Five Dimensional Circumplex. *Journal of personality and social psychology*, *65*(3), 563.

Jørgensen, K. (2008). Left in the dark: Playing computer games with the sound turned off. In Collins, K. (Ed.), *From Pac-Man to pop music: Interactive audio in games and new media* (pp. 163–176). Aldershot, Hampshire: Ashgate Publishing Group.

Jorna, P. G. A. M. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological psychology*, *34*(2-3), 237-257.

Jorna, P. G. A. M. (1993). Heart rate and workload variations in actual and simulated flight. *Ergonomics*, *36*(9), 1043-1054.

Jović, A., Brkić, K., & Bogunović, N. (2015). A review of feature selection methods with applications. In *2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO)* (pp. 1200-1205). IEEE.

Joyner, L. A., & TerKeurst, J. (2006). Accounting for user needs and motivations in game design. *Interactive Convergence: Critical Issues in Multimedia*, *31*.

Jurnet, I. A., Beciu, C. C., & Maldonado, J. G. (2005). Individual differences in the sense of presence. In *Proceedings of Presence 2005: The 8th International Workshop on Presence* (pp. 133-142).

Juul, J. (2009). Fear of failing? The many meanings of difficulty in video games. *The video game theory reader*, *2*, 237-252.

Juul, J. (2013). *The art of failure: An essay on the pain of playing video games*. MIT press.

Kahneman, D. (1999). Objective happiness. *Well-being: The foundations of hedonic psychology*, *3*(25), 1-23.

Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: Foundations of hedonic psychology*. New York: Russell Sage Foundation.

Kahneman, D., Kahneman, D., & Tversky, A. (2003). Experienced utility and objective happiness: A moment-based approach. *The psychology of economic decisions*, *1*, 187-208.

Kam, J. W., Griffin, S., Shen, A., Patel, S., Hinrichs, H., Heinze, H. J., ... & Knight, R. T. (2019). Systematic comparison between a wireless EEG system with dry electrodes and a wired EEG system with wet electrodes. *NeuroImage*, *184*, 119-129.

Kamiński, J., Brzezicka, A., Gola, M., & Wróbel, A. (2012). Beta band oscillations engagement in human alertness process. *International Journal of Psychophysiology*, *85*(1), 125-128.

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological science*, *18*(7), 614-621.

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*(5), 348-354.

Kang, J., Ko, I., & Ko, Y. (2009). The impact of social support of guild members and psychological factors on flow and game loyalty in MMORPG. In *2009 42nd Hawaii International Conference on System Sciences* (pp. 1-9). IEEE.

Kannegieser, E., & Atorf, D. (2020). A Study to Further Understand the Link Between Immersion and Flow. In *Interactivity and the Future of the Human-Computer Interface* (pp. 114-122). IGI Global.

Kannetis, T., Potamianos, A., & Yannakakis, G. N. (2009). Fantasy, curiosity and challenge as adaptation indicators in multimodal dialogue systems for preschoolers. In *Proceedings of the 2nd Workshop on Child, Computer and Interaction* (p. 1). ACM.

Kao, K. C. D. (2012). *Reward Preference in Video Games* (Doctoral dissertation). Princeton University, New Jersey, United States.

Kashdan, T. B., & Silvia, P. J. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. *Oxford handbook of positive psychology*, *2*, 367-374.

Katahira, K., Yamazaki, Y., Yamaoka, C., Ozaki, H., Nakagawa, S., & Nagata, N. (2018). EEG correlates of the flow state: A combination of increased frontal theta and moderate frontocentral alpha rhythm in the mental arithmetic task. *Frontiers in psychology*, *9*, 300.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review*, *19*(4), 429.

Katz-Leurer, M., & Shochina, M. (2005). Heart Rate Variability (HRV) parameters correlate with motor impairment and aerobic capacity in stroke patients. *NeuroRehabilitation*, *20*(2), 91-95.

Kavzoglu, T. (2009). Increasing the accuracy of neural network classification using refined training data. *Environmental Modelling & Software*, *24*(7), 850-858.

Kaye, L. K., Monk, R. L., Wall, H. J., Hamlin, I., and Qureshi, A. W. (2018). The effect of flow and context on in-vivo positive mood in digital gaming. *Int. J. Hum.Comp. Stud.* 110, 45–52. doi: 10.1016/j.ijhcs.2017.10.005

Kearns, M., & Ron, D. (1999). Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural computation*, *11*(6), 1427-1453.

Kekecs, Z., Szekely, A., & Varga, K. (2016). Alterations in electrodermal activity and cardiac parasympathetic tone during hypnosis. *Psychophysiology*, *53*(2), 268-277.

Keller, J., & Bless, H. (2008). Flow and regulatory compatibility: An experimental approach to the flow model of intrinsic motivation. *Personality and social psychology bulletin*, *34*(2), 196-209.

Keller, J., Bless, H., Blomann, F., & Kleinböhl, D. (2011a). Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology*, *47*(4), 849-852.

Keller, J., Ringelhan, S., & Blomann, F. (2011b). Does skills–demands compatibility result in intrinsic motivation? Experimental test of a basic notion proposed in the theory of flow-experiences. *The Journal of Positive Psychology*, *6*(5), 408-417.

Kennedy, D. O., & Scholey, A. B. (2000). Glucose administration, heart rate and cognitive performance: effects of increasing mental effort. *Psychopharmacology*, *149*(1), 63-71.

Kern, M. L., & Friedman, H. S. (2008). Do conscientious individuals live longer? A quantitative review. *Health psychology*, *27*(5), 505.

Kerous, B., Skola, F., & Liarokapis, F. (2018). EEG-based BCI and video games: a progress report. *Virtual Reality*, *22*(2), 119-135.

Khanmohammadi, S., Adibeig, N., & Shanehbandy, S. (2017). An improved overlapping k-means clustering method for medical applications. *Expert Systems with Applications*, *67*, 12-18.

Kiili, K., De Freitas, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science*, *15*, 78-91.

Kiili, K., Lainema, T., de Freitas, S., & Arnab, S. (2014). Flow framework for analyzing the quality of educational games. *Entertainment computing*, *5*(4), 367-377.

Kim, H. C., Pang, S., Je, H. M., Kim, D., & Bang, S. Y. (2002). Support vector machine ensemble with bagging. In *International Workshop on Support Vector Machines* (pp. 397-408). Springer, Berlin, Heidelberg.

Kim, J., Oh, H., Kim, W., Choi, S., Son, W., & Lee, S. (2020). A Deep Motion Sickness Predictor Induced by Visual Stimuli in Virtual Reality. *IEEE Transactions on Neural Networks and Learning Systems*.

Kim, Y., & Ross, S. D. (2006). An exploration of motives in sport video gaming. *International Journal of Sports Marketing and Sponsorship*, *8*(1), 28-40.

King, D. L., & Delfabbro, P. H. (2009). Understanding and assisting excessive players of video games: A community psychology perspective. *Australian Community Psychologist*, *21*(1), 62-74.

King, D. L., Delfabbro, P. H., & Griffiths, M. D. (2010). The role of structural characteristics in problem video game playing: A review. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *4*(1).

Kirriemuir, J., & McFarlane, A. (2003). Use of Computer and Video Games in the Classroom. In *DiGRA Conference*.

Kirsch, S. J. (2006). *Children, adolescents, and media violence*. Thousand Oaks, CA: Sage.

Kirschner, D., & Williams, J. P. (2014). Measuring video game engagement through gameplay reviews. *Simulation & Gaming*, *45*(4-5), 593-610.

Kivikangas, J. M. (2006). Psychophysiology of flow experience: An explorative study (Master Thesis). University of Helsinki: Helsinki, Finland.

Kivikangas, J. M., & Ravaja, N. (2013). Emotional responses to victory and defeat as a function of opponent. *IEEE Transactions on Affective Computing*, *4*(2), 173-182.

Klados, M. A., Konstantinidi, P., Dacosta-Aguayo, R., Kostaridou, V. D., Vinciarelli, A., & Zervakis, M. (2020). Automatic Recognition of Personality Profiles Using EEG Functional Connectivity During Emotional Processing. *Brain sciences*, *10*(5), 278.

Klarkowski, M., Johnson, D., Wyeth, P., Smith, S., & Phillips, C. (2015). Operationalising and measuring flow in video games. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* (pp. 114-118). ACM.

Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A., and Mathiak, K. (2012). Neural contributions to flow experience during video game playing. *Soc. Cogn. Affect. Neurosci.* 7, 485–495. doi: 10.1093/scan/nsr021

Klimesch, W., Schack, B., & Sauseng, P. (2005). The functional significance of theta and upper alpha oscillations. *Experimental psychology*, *52*(2), 99-108.

Klimmt, C. (2005). *Computerspielen als Handlung: Dimensionen und Determinanten des Erlebens interaktiver Unterhaltung*. [*Computer game play as action: Dimensions and determinants of the experience of interactive entertainment*]. Köln: von Halem.

Klimmt, C., & Hartmann, T. (2006). Effectance, self-efficacy, and the motivation to play video games. *Playing video games: Motives, responses, and consequences*, 133-145.

Klimmt, C., Hartmann, T., & Frey, A. (2007). Effectance and control as determinants of video game enjoyment. *Cyberpsychology & behavior*, *10*(6), 845-848.

Klimmt, C., Blake, C., Hefner, D., Vorderer, P., & Roth, C. (2009). Player performance, satisfaction, and video game enjoyment. In *International Conference on Entertainment Computing* (pp. 1-12). Springer, Berlin, Heidelberg.

Klingberg, T. (2009). *The overflowing brain: Information overload and the limits of working memory*. Oxford University Press.

Knierim, M. T., Rissler, R., Dorner, V., Maedche, A., & Weinhardt, C. (2018). The psychophysiology of flow: A systematic review of peripheral nervous system features. In *Information Systems and Neuroscience* (pp. 109-120). Springer, Cham.

Knox, M., Lentini, J., Cummings, T. S., McGrady, A., Whearty, K., & Sancrant, L. (2011). Game-based biofeedback for paediatric anxiety and depression. *Mental health in family medicine*, *8*(3), 195.

Knyazev, G. G. (2007). Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neuroscience & Biobehavioral Reviews*, *31*(3), 377-395.

Ko, S. M., & Ji, Y. G. (2018). How we can measure the non-driving-task engagement in automated driving: comparing flow experience and workload. *Applied ergonomics*, *67*, 237-245.

Kobayashi, R., Honda, T., Hashimoto, J., Kashihara, S., Iwasa, Y., Yamamoto, K., ... & Haraguchi, Y. (2020). Resting-state theta/beta ratio is associated with distraction but not with reappraisal. *Biological Psychology*, *155*, 107942.

Kong, X., & Wilson, G. F. (1998). A new EOG-based eyeblink detection algorithm. *Behavior Research Methods, Instruments, & Computers*, *30*(4), 713-719.

Korjus, K., Uusberg, A., Uusberg, H., Kuldkepp, N., Kreegipuu, K., Allik, J., ... & Aru, J. (2015). Personality cannot be predicted from the power of resting state EEG. *Frontiers in human neuroscience*, *9*, 63.

Kozhevnikov, M., Li, Y., Wong, S., Obana, T., & Amihai, I. (2018). Do enhanced states exist? Boosting cognitive capacities through an action video-game. *Cognition*, *173*, 93-105.

Kramer, A. F. (1990). Physiological metrics of mental workload: A review of recent progress. In D. Damos (Eds.), *Multiple-task Performance* (pp. 279-328). London: Taylor & Francis.

Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, *42*, 853–863.

Krekhov, A., Cmentowski, S., & Krüger, J. (2018). VR Animals: Surreal Body Ownership in Virtual Reality Games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*(pp. 503-511). ACM.

Kron, A., Pilkiw, M., Banaei, J., Goldstein, A., & Anderson, A. K. (2015). Are valence and arousal separable in emotional experience?. *Emotion*, *15*(1), 35.

Krygier, J. R., Heathers, J. A., Shahrestani, S., Abbott, M., Gross, J. J., & Kemp, A. H. (2013). Mindfulness meditation, well-being, and heart rate variability: a preliminary investigation into the impact of intensive Vipassana meditation. *International Journal of Psychophysiology*, *89*(3), 305-313.

Kunc, N. (1992). The need to belong: Rediscovering Maslow's hierarchy of needs. In R. A. Villa, J. S. Thousand, W. Stainback, & S. Stainback (Eds.), *Restructuring for caring and effective education: An administrative guide to creating heterogeneous schools* (pp. 25-39). Baltimore, MD, England: Paul H. Brookes Publishing.

Kumar, D., & Sharma, A. (2016). Electrooculogram-based virtual reality game control using blink detection and gaze calibration. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2358-2362). IEEE.

Kumar, P., Saini, R., Sahu, P. K., Roy, P. P., Dogra, D. P., & Balasubramanian, R. (2017). Neuro-phone: An assistive framework to operate Smartphone using EEG signals. In *2017 IEEE Region 10 Symposium (TENSYMP)* (pp. 1-5). IEEE.

Lafrenière, M. A. K., Vallerand, R. J., Donahue, E. G., & Lavigne, G. L. (2009). On the costs and benefits of gaming: The role of passion. *CyberPsychology & Behavior*, *12*(3), 285-290.

Lagopoulos, J., Xu, J., Rasmussen, I., Vik, A., Malhi, G. S., Eliassen, C. F., ... & Davanger, S. (2009). Increased theta and alpha EEG activity during nondirective meditation. *The Journal of Alternative and Complementary Medicine*, *15*(11), 1187-1192.

Lahey, B. B. (2009). Public health significance of neuroticism. *American Psychologist*, *64*(4), 241.

Lal, T. N., Schroder, M., Hinterberger, T., Weston, J., Bogdan, M., Birbaumer, N., & Scholkopf, B. (2004). Support vector channel selection in BCI. *IEEE transactions on biomedical engineering*, *51*(6), 1003-1010.

Lall, U., & Sharma, A. (1996). A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resources Research*, *32*(3), 679-693.

Landhäußer, A., & Keller, J. (2012). Flow and its affective, cognitive, and performance-related consequences. In *Advances in flow research* (pp. 65-85). New York : Springer.

Landolfi, E. (2013). Exercise addiction. *Sports medicine*, *43*(2), 111-119.

Lang, P. J. (1995). The emotion probe: studies of motivation and attention. *American psychologist*, *50*(5), 372.

Lantz, G., De Peralta, R. G., Spinelli, L., Seeck, M., & Michel, C. M. (2003). Epileptic source localization with high density EEG: how many electrodes are needed?. *Clinical neurophysiology*, *114*(1), 63-69.

Larche, C. J., & Dixon, M. J. (2020). The relationship between the skill-challenge balance, game expertise, flow and the urge to keep playing complex mobile games. *Journal of Behavioral Addictions*.

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time?. *Journal of personality and social psychology*, *81*(4), 684–696.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*(5), 776-785.

Larsen, R. J. (2000). Toward a science of mood regulation. *Psychological Inquiry*, *11*(3), 129-141.

Latham, A. J., Patston, L. L., and Tippett, L. J. (2013). Just how expert are "expert" video-game players? Assessing the experience and expertise of video-game players across "action" video-game genres. *Front. Psychol.* 4:941. doi: 10.3389/fpsyg.2013.00941

Lau-Zhu, A., Lau, M. P., & McLoughlin, G. (2019). Mobile EEG in research on neurodevelopmental disorders: Opportunities and challenges. *Developmental cognitive neuroscience*, *36*, 100635.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences*, *9*(2), 75-82.

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annu. Rev. Psychol.*, *58*, 317-344.

Leary, M. R., Adams, C. E., & Tate, E. B. (2006). Hypo-egoic self-regulation: Exercising self-control by diminishing the influence of the self. *Journal of personality*, *74*(6), 1803-1832.

Lee, D., & LaRose, R. (2007). A socio-cognitive model of video game usage. *Journal of Broadcasting & Electronic Media*, *51*(4), 632-650.

Lee, P. M., Jheng, S. Y., & Hsiao, T. C. (2014). Towards automatically detecting whether student is in flow. In *International Conference on Intelligent Tutoring Systems* (pp. 11-18). Springer, Cham.

Léger, P. M., Davis, F. D., Cronan, T. P., and Perret, J. (2014). Neurophysiological correlates of cognitive absorption in an enactive training context. *Comp. Hum. Behav.* 34, 273–283. doi: 10.1016/j.chb.2014.02.011

Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2009). Development and validation of a game addiction scale for adolescents. *Media psychology*, *12*(1), 77-95.

Leroy, A., & Cheron, G. (2020). EEG dynamics and neural generators of psychological flow during one tightrope performance. *Scientific reports*, *10*(1), 1-13.

Li, B., Fu, H., Wen, D., & Lo, W. (2018). Etracker: A mobile gaze-tracking system with near-eye display based on a combined gaze-tracking algorithm. *Sensors*, *18*(5), 1626.

Li, L., Yang, Y., & Mingxin, L. (2006). The Relationship Between Adolescents' Neuroticism, Internet Service Preference, and Internet Addiction. *Acta Psychologica Sinica*, *3*, 008.

Li, M., & Lu, B. L. (2009). Emotion classification based on gamma-band EEG. In *2009 Annual International Conference of the IEEE Engineering in medicine and biology society* (pp. 1223-1226). IEEE.

Li, S., Leider, A., Qiu, M., Gai, K., & Liu, M. (2017). Brain-based computer interfaces in virtual reality. In *2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud)* (pp. 300-305). IEEE.

Liao, L. D., Chen, C. Y., Wang, I. J., Chen, S. F., Li, S. Y., Chen, B. W., ... & Lin, C. T. (2012). Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors. *Journal of neuroengineering and rehabilitation*, *9*(1), 5.

Lim, C. K. A., & Chia, W. C. (2015). Analysis of single-electrode EEG rhythms using MATLAB to elicit correlation with cognitive stress. *International Journal of Computer Theory and Engineering*, *7*(2), 149.

Limperos, A. M., Schmierbach, M. G., Kegerise, A. D., & Dardis, F. E. (2011). Gaming across different consoles: exploring the influence of control scheme on game-player enjoyment. *Cyberpsychology, Behavior, and Social Networking*, *14*(6), 345-350.

Lin, C. T., & Do, T. T. N. (2020). Direct-Sense Brain-Computer Interfaces and Wearable Computers. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Liu, C. C. (2017). A model for exploring players flow experience in online games. *Information Technology & People*, *30*(1), 139-162.

Liu, H. J., & Shiue, Y. C. (2014). Influence of Facebook game players' behavior on flow and purchase intention. *Social Behavior and Personality: an international journal*, *42*(1), 125-133.

Liu, J., Meng, H., Nandi, A., & Li, M. (2016). Emotion detection from EEG recordings. In *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* (pp. 1722-1727). IEEE.

Liu, T., & Csikszentmihalyi, M. (2020). Flow among introverts and extraverts in solitary and social activities. *Personality and Individual Differences*, *167*, 110197.

Llorens, S., Salanova, M., & Rodríguez, A. M. (2013). How is flow experienced and by whom? Testing flow among occupations. *Stress and Health*, *29*(2), 125-137.

Loewy, A. D., & Spyer, K. M. (1990). *Central regulation of autonomic functions*. Oxford University Press.

Longstreet, P., & Brooks, S. (2017). Life satisfaction: A key to managing internet & social media addiction. *Technology in Society*, *50*, 73-77.

Lopez-Gordo, M. A., Sanchez-Morillo, D., & Valle, F. P. (2014). Dry EEG electrodes. *Sensors*, *14*(7), 12847-12870.

Lora, A. T., Riquelme, J. C., Ramos, J. L. M., Santos, J. M. R., & Expósito, A. G. (2003). Influence of kNN-based load forecasting errors on optimal energy production. In *Portuguese Conference on Artificial Intelligence* (pp. 189-203). Springer, Berlin, Heidelberg.

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of neural engineering*, *4*(2), R1.

Lotte, F., Jeunet, C., Mladenović, J., N'Kaoua, B., & Pillette, L. (2018). A BCI challenge for the signal processing community: considering the user in the loop.

Løvoll, H. S., & Vittersø, J. (2014). Can balance be boring? A critique of the "challenges should match skills" hypotheses in flow theory. *Social indicators research*, *115*(1), 117-136.

Lyons, E. J. (2015). Cultivating engagement and enjoyment in exergames using feedback, challenge, and rewards. *Games for health journal*, *4*(1), 12-18.

Ma, J., Zhang, Y., Cichocki, A., & Matsuno, F. (2015). A novel EOG/EEG hybrid human–machine interface adopting eye movements and ERPs: Application to robot control. *IEEE Transactions on Biomedical Engineering*, *62*(3), 876-889.

Maffei, A., & Angrilli, A. (2019). Spontaneous blink rate as an index of attention and emotion during film clips viewing. *Physiology & behavior*, *204*, 256-263.

Magliacano, A., Fiorenza, S., Estraneo, A., & Trojano, L. (2020). Eye blink rate increases as a function of cognitive load during an auditory oddball paradigm. *Neuroscience Letters*, *736*, 135293.

Magosso, E., De Crescenzio, F., Ricci, G., Piastra, S., & Ursino, M. (2019). EEG Alpha Power Is Modulated by Attentional Changes during Cognitive Tasks and Virtual Reality Immersion. *Computational Intelligence and Neuroscience*, *2019*.

Magyaródi, T., Nagy, H., Soltész, P., Mózes, T., & Oláh, A. (2013). Psychometric properties of a newly established flow state questionnaire. *The Journal of Happiness & Well-Being*, *1*(2), 85-96.

Malik, M., & Camm, A. J. (1995). *Heart Rate Variability*. New York: Futura.

Mallapragada, C. (2018). *Classification of EEG signals of user states in gaming using machine learning* (Master's thesis). Missouri University of Science and Technology, United States.

Mallick, R., Slayback, D., Touryan, J., Ries, A. J., & Lance, B. J. (2016). The use of eye metrics to index cognitive workload in video games. In *2016 IEEE Second Workshop on Eye Tracking and Visualization (ETVIS)* (pp. 60-64). IEEE.

Malone, T. W. (1981). What makes computer games fun? *Byte*, *6*(12), 258-277.

Manero, B., Torrente, J., Freire, M., & Fernández-Manjón, B. (2016). An instrument to build a gamer clustering framework according to gaming preferences and habits. *Computers in Human Behavior*, *62*, 353-363.

Mangai, U. G., Samanta, S., Das, S., & Chowdhury, P. R. (2010). A survey of decision fusion and feature fusion strategies for pattern classification. *IETE Technical review*, *27*(4), 293-307.

Mania, K., & Chalmers, A. (2001). The effects of levels of immersion on memory and presence in virtual environments: A reality centered approach. *CyberPsychology & Behavior*, *4*(2), 247-264.

Mansfield, B. E., Oddson, B. E., Turcotte, J., & Couture, R. T. (2012). A possible physiological correlate for mental flow. *The Journal of Positive Psychology*, *7*(4), 327-333.

Mao, Y., Roberts, S., Pagliaro, S., Csikszentmihalyi, M., & Bonaiuto, M. (2016). Optimal experience and optimal identity: A multinational study of the associations between flow and social identity. *Frontiers in psychology*, *7*, 67.

Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of eye-related measures of drivers' mental workload. *Procedia Manufacturing*, *3*, 2854-2861.

Marr, A. J. (2001). In the zone: A biobehavioral theory of the flow experience. *Athletic Insight: Online Journal of Sport Psychology*, *3*(1).

Martel, E., Su, F., Gerroir, J., Hassan, A., Girouard, A., & Muldner, K. (2015). Diving Head-First into Virtual Reality: Evaluating HMD Control Schemes for VR Games. In *FDG*.

Martey, R. M., Kenski, K., Folkestad, J., Feldman, L., Gordis, E., Shaw, A., ... & Strzalkowski, T. (2014). Measuring game engagement: multiple methods and construct complexity. *Simulation & Gaming*, *45*(4-5), 528-547.

Martínez-Cagigal, V. (2020). Topographic EEG/MEG plot (https://www.mathworks.com/matlabcentral/fileexchange/72729-topographic-eeg-meg-plot), *MATLAB Central File Exchange*. Retrieved October 27, 2020.

Martins, R., & Carvalho, J. M. (2015). Eye blinking as an indicator of fatigue and mental load-a systematic review. *Occupational safety and hygiene III*, *10*.

Maslow, A. H. (1943). A theory of human motivation. Psychological Review, 50, 370–396. http://dx.doi.org/10.1037/h0054346

Maslow, A. H. (1974). The good life of self-actualizing person. *Readings in human development: A humanistic approach*, *46*.

Massimini, F., & Carli, M. (1986). La selezione psicologica umana tra biologia e cultura. In F. Massimini & P. Inghilleri (eds.), *L'esperienza quotidiana* (pp. 65–84). Milan: Franco Angeli.

Mathiak, K. A., Klasen, M., Weber, R., Ackermann, H., Shergill, S. S., & Mathiak, K. (2011). Reward system and temporal pole contributions to affective evaluation during a first person shooter video game. *BMC neuroscience*, *12*(1), 66.

Mathiak, K. A., Klasen, M., Zvyagintsev, M., Weber, R., & Mathiak, K. (2013). Neural networks underlying affective states in a multimodal virtual environment: contributions to boredom. *Frontiers in human neuroscience*, *7*, 820.

Mathwick, C., & Rigdon, E. (2004). Play, flow, and the online search experience. *Journal of consumer research*, *31*(2), 324-332.

Matthews, G., & Dorn, L. (1995). Cognitive and attentional processes in personality and intelligence. In *International handbook of personality and intelligence* (pp. 367-396). Springer, Boston, MA.

McCall, G. J. (2003). The me and the not-me. In *Advances in identity theory and research* (pp. 11-25). Springer, Boston, MA.

McClelland, D. C. (1985). *Human Motivation*. NY: Cambridge University Press.

McCorry, L. K. (2007). Physiology of the autonomic nervous system. *American journal of pharmaceutical education*, *71*(4), 78.

McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of personality and social psychology*, *52*(6), 1258.

McCrae, R. R., & Costa Jr, P. T. (1985). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and individual Differences*, *6*(5), 587-597.

McCrae, R. R., & Costa Jr, P. T. (1991). Adding Liebe und Arbeit: The full five-factor model and well-being. *Personality and social psychology bulletin*, *17*(2), 227-232.

McCrae, R. R., & Costa Jr, P. T. (1997). Conceptions and correlates of openness to experience. In *Handbook of personality psychology* (pp. 825-847). Academic Press.

McCraty, R., & Shaffer, F. (2015). Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. *Global advances in health and medicine*, *4*(1), 46-61.

McCreery, M. P., Schrader, P. G., Krach, S. K., & Boone, R. (2013). A sense of self: The role of presence in virtual environments. *Computers in Human Behavior*, *29*(4), 1635-1640.

McDermott, A. F., Bavelier, D., & Green, C. S. (2014). Memory abilities in action video game players. *Computers in Human Behavior*, *34*, 69-78.

McElroy, J. C., Hendrickson, A. R., Townsend, A. M., & DeMarie, S. M. (2007). Dispositional factors in internet use: personality versus cognitive style. *MIS quarterly*, *31*(4), 809-820.

McIntire, L. K., McKinley, R. A., Goodyear, C., & McIntire, J. P. (2014). Detection of vigilance performance using eye blinks. *Applied ergonomics*, *45*(2), 354-362.

McLaughlin, A., Gandy, M., Allaire, J., & Whitlock, L. (2012). Putting fun into video games for older adults. *Ergonomics in Design*, *20*(2), 13-22.

McMahan, A. (2003). Immersion, engagement and presence. *Video Game Theory Reader* 67-86.

McMahan, R. P., Ragan, E. D., Leal, A., Beaton, R. J., & Bowman, D. A. (2011). Considerations for the use of commercial video games in controlled experiments. *Entertainment Computing*, *2*(1), 3-9.

McMahan, T., Parberry, I., & Parsons, T. D. (2015a). Evaluating Electroencephalography Engagement Indices During Video Game Play. In *FDG*.

McMahan, T., Parberry, I., & Parsons, T. D. (2015b). Evaluating player task engagement and arousal using electroencephalography. *Procedia Manufacturing*, *3*, 2303-2310.

McMahan, T., Parberry, I., & Parsons, T. D. (2015c). Modality specific assessment of video game player's experience using the Emotiv. *Entertainment Computing*, *7*, 1-6.

McMahon, N., Wyeth, P., & Johnson, D. (2012). Personality and player types in Fallout New Vegas. In *Proceedings of the 4th International Conference on Fun and Games* (pp. 113-116). ACM.

McNiel, J. M., & Fleeson, W. (2006). The causal effects of extraversion on positive affect and neuroticism on negative affect: Manipulating state extraversion and state neuroticism in an experimental approach. *Journal of Research in Personality*, *40*(5), 529-550.

Medler, B. (2011). Player dossiers: Analyzing gameplay data as a reward. *Game Studies*, *11*(1).

Meftah, I. T., Le Thanh, N., & Amar, C. B. (2012). Emotion recognition using KNN classification for user modeling and sharing of affect states. In *International Conference on Neural Information Processing* (pp. 234-242). Springer, Berlin, Heidelberg.

Mehroof, M., & Griffiths, M. D. (2010). Online gaming addiction: the role of sensation seeking, self-control, neuroticism, aggression, state anxiety, and trait anxiety. *Cyberpsychology, behavior, and social networking*, *13*(3), 313-316.

Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Kuehnast, M., & Jacobsen, T. (2015). Towards a psychological construct of being moved. *PloS one*, *10*(6), e0128451.

Mercier, J., Chalfoun, P., Martin, M., Tato, A. A., & Rivas, D. (2019). Predicting Subjective Enjoyment of Aspects of a Videogame from Psychophysiological Measures of Arousal and Valence. In *International Conference on Intelligent Tutoring Systems* (pp. 174-179). Springer, Cham.

Mert, A., Kilic, N., & Akan, A. (2012). ECG signal classification using ensemble decision tree. *J Trends Dev Mach Assoc Technol*, *16*(1), 179-182.

Meshkati, N. (1988). Heart rate variability and mental workload assessment. In *Advances in psychology* (Vol. 52, pp. 101-115). North-Holland.

Metin, B., Goktepe, A. K., Kaya, B., Tas, C., Serin, E., Dolu, F., & Tarhan, N. (2017). EEG findings during flow state. *The Journal of Neurobehavioral Sciences*, *4*(2), 1–52. https://doi.org/10.5455/JNBS.1496152464

Meyer, D. (2019). Support Vector Machines – The Interface to libsvm in package e1071. FH Technikum Wien, Austria. Accessed from https://cran.r-project.org/web/packages/e1071/vignettes/svmdoc.pdf.

Michailidis, L., Balaguer-Ballester, E., & He, X. (2018). Flow and Immersion in Video Games: The Aftermath of a Conceptual Challenge. *Front. Psychol.* 9:1682. doi: 10.3389/fpsyg.2018.01682

Michel, C. M., & Brunet, D. (2019). EEG source imaging: a practical review of the analysis steps. *Frontiers in neurology*, *10*, 325.

Mikulas, W. L., & Vodanovich, S. J. (1993). The essence of boredom. *The Psychological Record*, *43*(1), 3.

Miller, A., & Hom, H. L. (1990). Influence of extrinsic and ego incentive value on persistence after failure and continuing motivation. *Journal of Educational Psychology*, *82*(3), 539.

Miller, M. K., & Mandryk, R. L. (2016). Differentiating in-game frustration from at-game frustration using touch pressure. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces* (pp. 225-234). ACM.

Mills, D. J., Milyavskaya, M., Mettler, J., & Heath, N. L. (2018). Exploring the pull and push underlying problem video game use: A self-determination theory approach. *Personality and Individual Differences*, *135*, 176-181.

Min, S. N., Kim, J. Y., Subramaniyam, M., Kim, D. J., Jung, J. H., & Lee, K. S. (2018). Evaluation of Flow Level through Pupil Size, Blink Rate and Blink Duration in Different Types of Smartphone Games. *Journal of the Ergonomics Society of Korea*, *37*(5), 563-573.

Ming, D., Xi, Y., Zhang, M., Qi, H., Cheng, L., Wan, B., & Li, L. (2009). Electroencephalograph (EEG) signal processing method of motor imaginary potential for attention level classification. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4347-4351). IEEE.

Minguillon, J., Lopez-Gordo, M. A., & Pelayo, F. (2017). Trends in EEG-BCI for daily-life: Requirements for artifact removal. *Biomedical Signal Processing and Control*, *31*, 407-418.

Minnix, J. A., & Kline, J. P. (2004). Neuroticism predicts resting frontal EEG asymmetry variability. *Personality and Individual Differences*, *36*(4), 823-832.

Mirlohi, M., Egbert, J., & Ghonsooly, B. (2011). Flow in translation: Exploring optimal experience for translation trainees. *Target. International Journal of Translation Studies*, *23*(2), 251-271.

Missura, O., & Gärtner, T. (2009). Player modeling for intelligent difficulty adjustment. In *International Conference on Discovery Science* (pp. 197-211). Springer, Berlin, Heidelberg.

Mitchell, T. R. (1982). Motivation: New directions for theory, research, and practice. *Academy of management review*, *7*(1), 80-88.

Mockus, J., Tiesis, V., & Zilinskas, A. (1978). The application of Bayesian methods for seeking the extremum. *Towards global optimization*, *2*(117-129), 2.

Mohamad, I. B., & Usman, D. (2013). Standardization and its effects on K-means clustering algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, *6*(17), 3299-3303.

Mohammadi, Z., Frounchi, J., & Amiri, M. (2017). Wavelet-based emotion recognition system using EEG signal. *Neural Computing and Applications*, *28*(8), 1985-1990.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques* (pp. 23-26). Beverly Hills, CA: Sage Publications.

Monteiro, D., Liang, H. N., Abel, A., Bahaei, N., & de Cassia Monteiro, R. (2018). Evaluating Engagement of Virtual Reality Games Based on First and Third Person Perspective Using EEG and Subjective Metrics. In *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)* (pp. 53-60). IEEE.

Moore, K., & McElroy, J. C. (2012). The influence of personality on Facebook usage, wall postings, and regret. *Computers in Human Behavior*, *28*(1), 267-274.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326.

Morgan, M. J. (1974). Resistance to satiation. *Animal Behaviour*, *22*(2), 449-466.

Moroşan, M. (2019). *Automating Game-design and Game-agent Balancing through Computational Intelligence* (Doctoral dissertation). University of Essex, United Kingdom. Available at http://repository.essex.ac.uk/24233/1/THESIS.pdf.

Mosing, M. A., Magnusson, P. K., Pedersen, N. L., Nakamura, J., Madison, G., & Ullén, F. (2012a). Heritability of proneness for psychological flow experiences. *Personality and Individual Differences*, *53*(5), 699-704.

Mosing, M. A., Pedersen, N. L., Cesarini, D., Johannesson, M., Magnusson, P. K., Nakamura, J., ... & Ullén, F. (2012b). Genetic and environmental influences on the relationship between flow proneness, locus of control and behavioral inhibition. *Plos one*, *7*(11), e47958.

Moss, D. (2004). Heart rate variability (HRV) biofeedback. *Psychophysiology today*, *1*(1), 4-11.

Motamedi-Fakhr, S., Moshrefi-Torbati, M., Hill, M., Hill, C. M., & White, P. R. (2014). Signal processing techniques applied to human sleep EEG signals—A review. *Biomedical Signal Processing and Control*, *10*, 21-33.

Mulder, G. (1979). Mental load, mental effort and attention. In *Mental Workload* (pp. 299-325). Springer, Boston, MA.

Müller, B. C., Marx, A. K., Paulus, M., & Meinhardt, J. (2018). Frontal electroencephalogram alpha asymmetry relates to implicit achievement motives: A pilot study. *Mind, Brain, and Education*, *12*(2), 82-89.

Murayama, K., Matsumoto, M., Izuma, K., & Matsumoto, K. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences*, *107*(49), 20911-20916.

Muthukumaraswamy, S. (2013). High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. *Frontiers in human neuroscience*, *7*, 138.

Nacke, L. E., Bateman, C., & Mandryk, R. L. (2014). BrainHex: A neurobiological gamer typology survey. *Entertainment computing*, *5*(1), 55-62.

Nacke, L. E., & Lindley, C. A. (2008). Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (pp. 81-88). ACM.

Nacke, L. E., & Lindley, C. A. (2010). Affective ludology, flow and immersion in a first-person shooter: Measurement of player experience. *arXiv preprint arXiv:1004.0248*.

Nacke, L. E., Stellmach, S., & Lindley, C. A. (2011). Electroencephalographic assessment of player experience: A pilot study in affective ludology. *Simulation & Gaming*, *42*(5), 632-655.

Nagle, A., Riener, R., & Wolf, P. (2018). Personality-based reward contingency selection: A player-centered approach to gameplay customization in a serious game for cognitive training. *Entertainment Computing*, *28*, 70-77.

Nagle, A., Wolf, P., & Riener, R. (2016). Towards a system of customized video game mechanics based on player personality: Relating the Big Five personality traits with difficulty adaptation in a first-person shooter game. *Entertainment computing*, *13*, 10-24.

Nah, F. F. H., Eschenbrenner, B., Zeng, Q., Telaprolu, V. R., & Sepehr, S. (2014). Flow in gaming: literature synthesis and framework development. *International Journal of Information Systems and Management*, *1*(1-2), 83-124.

Nah, F. F. H., Yelamanchili, T., and Siau, K. (2017). A review on neuropsychophysiological correlates of flow. In *Proceedings of the International Conference on HCI in Business, Government, and Organizations* (Cham: Springer), 364–372.

Nakamura, J., and Csikszentmihalyi, M. (2014). *The Concept of Flow in Flow and the Foundations of Positive Psychology.* Dordrecht: Springer, 239–263.

Nakayama, M., Takahashi, K., & Shimizu, Y. (2002). The act of task difficulty and eye-movement frequency for the'Oculo-motor indices'. In *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 37-42). ACM.

Nascimento, S., Mirkin, B., & Moura-Pires, F. (2000). A fuzzy clustering model of data and fuzzy c-means. In *Ninth IEEE International Conference on Fuzzy Systems. FUZZ-IEEE 2000 (Cat. No. 00CH37063)* (Vol. 1, pp. 302-307). IEEE.

Neelamegam, S., & Ramaraj, E. (2013). Classification algorithm in data mining: An overview. *International Journal of P2P Network Trends and Technology (IJPTT)*, *4*(8), 369-374.

Newson, J. J., & Thiagarajan, T. C. (2019). EEG frequency bands in psychiatric disorders: a review of resting state studies. *Frontiers in human neuroscience*, *12*, 521.

Newzoo (2019). Key Numbers. Retrieved from https://newzoo.com/key-numbers/; https://newzoo.com/wp-content/uploads/2016/03/Newzoo-2019-Global-Games-Market-per-Segment.png (November 12, 2019).

Ngxande, M., Tapamo, J. R., & Burke, M. (2017). Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques. In *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)* (pp. 156-161). IEEE.

Nielsen, R. K. L., & Grabarczyk, P. (2019). Are Loot Boxes Gambling? Random reward mechanisms in video games. *Transactions of the Digital Games Research Association*, *4*(3).

Nijholt, A. (2008). BCI for games: A 'state of the art' survey. In *International Conference on Entertainment Computing* (pp. 225-228). Springer, Berlin, Heidelberg.

Nijholt, A., Bos, D. P. O., & Reuderink, B. (2009). Turning shortcomings into challenges: Brain–computer interfaces for games. *Entertainment computing*, *1*(2), 85-94.

Nordin, A. I., Denisova, A., & Cairns, P. (2014). Too many questionnaires: measuring player experience whilst playing digital games. In *Seventh York Doctoral Symposium on Computer Science & Electronics* (Vol. 69).

Nourbakhsh, N., Wang, Y., & Chen, F. (2013). GSR and blink features for cognitive load classification. In *IFIP Conference on Human-Computer Interaction* (pp. 159-166). Springer, Berlin, Heidelberg.

Nuñez Castellar, E. P., Antons, J. N., Marinazzo, D., and van Looy, J. (2016). Being in the zone: using behavioral and EEG recordings for the indirect assessment of flow. *PeerJ* 4:e2482v1.

Nylund, A., & Landfors, O. (2015). *Frustration and its effect on immersion in games: A developer viewpoint on the good and bad aspects of frustration* (Master's Thesis). Umeå University, Sweden.

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American society for Information Science and Technology*, *59*(6), 938-955.

O'Donnell, S., & Epstein, L. H. (2019). Smartphones are more reinforcing than food for students. *Addictive behaviors*, *90*, 124-133.

O'Donovan, S. (2012). Gamification of the games course. *Acesso em*, *17*, 1-8.

O'Hanlon, J. F. (1981). Boredom: Practical consequences and a theory. *Acta Psychologica*, *49*(1), 53-82.

O'Neill, M., Vaziripour, E., Wu, J., & Zappala, D. (2016). Condensing steam: Distilling the diversity of gamer behavior. In *Proceedings of the 2016 internet measurement conference* (pp. 81-95). ACM.

Oh, H., Hatfield, B. D., Jaquess, K. J., Lo, L. C., Tan, Y. Y., Prevost, M. C., ... & Blanco, J. A. (2015). A composite cognitive workload assessment system in pilots under various task demands using ensemble learning. In *International Conference on Augmented Cognition* (pp. 91-100). Springer, Cham.

Oh, H., Yang, H., & Yi, K. (2015). Learning a strategy for adapting a program analysis via bayesian optimisation. In *ACM SIGPLAN Notices* (Vol. 50, No. 10, pp. 572-588). ACM.

Oh, Y., & Yang, S. (2010). Defining exergames & exergaming. *Proceedings of Meaningful Play*, 1-17.

Ohira, H. (1995). Analysis of eyeblink activity during self-referent information processing in mild depression. *Perceptual and motor skills*, *81*, 1219-1229.

Öhman, A., Esteves, F., Flykt, A., & Soares, J. J. (1993). Gateways to consciousness: Emotion, attention, and electrodermal activity. In *Progress in electrodermal research* (pp. 137-157). Springer, Boston, MA.

Oliveira, L., Cardoso, J. S., Lourenço, A., & Ahlström, C. (2018). Driver drowsiness detection: a comparison between intrusive and non-intrusive signal acquisition methods. In *2018 7th European Workshop on Visual Information Processing (EUVIP)* (pp. 1-6). IEEE.

Ojeda, M. Á. N., Romera, C. S. N., & Velasco, J. R. (2019). In the Warcraft Universe We Trust: An Analysis of Transmedia Advertising Strategies in the World of Warcraft Video Game Series ("Battle Chest 3.0,""Cataclysm," and "Mists of Pandaria"). *International Journal of Communication*, *13*, 19.

Oman, C. M. (1990). Motion sickness: a synthesis and evaluation of the sensory conflict theory. *Canadian journal of physiology and pharmacology*, *68*(2), 294-303.

Ozkara, B. Y., Ozmen, M., & Kim, J. W. (2017). Examining the effect of flow experience on online purchase: A novel approach to the flow theory based on hedonic and utilitarian value. *Journal of Retailing and Consumer Services*, *37*, 119-131.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional science*, *32*(1), 1-8.

Paiva, J. S., Cardoso, J., & Pereira, T. (2018). Supervised learning methods for pathological arterial pulse wave differentiation: A SVM and neural networks approach. *International journal of medical informatics*, *109*, 30-38.

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 141-144). ACM.

Pallavicini, F., Ferrari, A., Pepe, A., Garcea, G., Zanacchi, A., & Mantovani, F. (2018). Effectiveness of virtual reality survival horror games for the emotional elicitation: Preliminary insights using Resident Evil 7: Biohazard. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 87-101). Springer, Cham.

Pallavicini, F., Pepe, A., & Minissi, M. E. (2019). Gaming in Virtual Reality: What Changes in Terms of Usability, Emotional Response and Sense of Presence Compared to Non-Immersive Video Games?. *Simulation & Gaming*, *50*(2), 136-159.

Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng*, *32*(3), 230-236.

Panda, S., Sahu, S., Jena, P., & Chattopadhyay, S. (2012). Comparing fuzzy-C means and K-means clustering techniques: a comprehensive study. In *Advances in Computer Science, Engineering & Applications* (pp. 451-460). Springer, Berlin, Heidelberg.

Pander, T., Przybyla, T., & Czabanski, R. (2008). An application of detection function for the eye blinking detection. In *2008 Conference on Human System Interactions* (pp. 287-291). IEEE.

Parikh, P., & Micheli-Tzanakou, E. (2004). Detecting drowsiness while driving using wavelet transform. In *Proceedings of the IEEE 30th Annual Northeast Bioengineering Conference* (pp. 79-80). IEEE.

Park, J., Song, Y., & Teng, C. I. (2011). Exploring the links between personality traits and motivations to play online games. *Cyberpsychology, Behavior, and Social Networking*, *14*(12), 747-751.

Parsons, T. D., McMahan, T., & Parberry, I. (2015). Neurogaming-based classification of player experience using consumer-grade electroencephalography. *IEEE Transactions on Affective Computing*.

Parsons, T. D., McMahan, T., & Parberry, I. (2020). Classification of Video Game Player Experience Using Consumer-Grade Electroencephalography. *IEEE Transactions on Affective Computing*.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: a meta-analysis of research findings. *Psychological bulletin*, *134*(2), 270.

Patil, A. J., Patil, C. S., Karhe, R. R., & Aher, M. A. (2014). Comparative study of different clustering algorithms. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng*, *3*(7), 10490-10497.

Paus, T., Koski, L., Caramanos, Z., & Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. *Neuroreport*, *9*(9), R37-R47.

Pausch, R., Proffitt, D., & Williams, G. (1997). Quantifying immersion in virtual reality. In T. Whitted (Ed.), *Proceedings from the ACM SIGGRAPH Annual Conference, Computer Graphics* (pp. 13–18). New York, NY: ACM Press/Addison-Wesley.

Payne, B. R., Jackson, J. J., Noh, S. R., & Stine-Morrow, E. A. (2011). In the zone: Flow state and cognition in older adults. *Psychology and aging*, *26*(3), 738.

Pecchia, L., Castaldo, R., Montesinos, L., & Melillo, P. (2018). Are ultra-short heart rate variability features good surrogates of short-term ones? State-of-the-art review and recommendations. *Healthcare technology letters*, *5*(3), 94-100.

Peever, N., Johnson, D., & Gardner, J. (2012). Personality & video game genre preferences. In *Proceedings of the 8th australasian conference on interactive entertainment: Playing the system* (p. 20). ACM.

Peifer, C. (2012). Psychophysiological correlates of flow-experience. In *Advances in flow research* (pp. 139-164). Springer, New York, NY.

Peifer, C., Schulz, A., Schächinger, H., Baumann, N., & Antoni, C. H. (2014). The relation of flow-experience and physiological arousal under stress—can u shape it?. *Journal of Experimental Social Psychology*, *53*, 62-69.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, *27*(8), 1226-1238.

Peng, X., Huang, J., Li, L., Gao, C., Chen, H., Tian, F., & Wang, H. (2019). Beyond Horror and Fear: Exploring Player Experience Invoked by Emotional Challenge in VR Games. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (p. LBW1616). ACM.

Perron, G. (2005). A cognitive psychological approach to gameplay emotions. In *Proceedings of DiGRA 2005 Conference*. http://www.digra.org/wp-content/uploads/digital-library/06276.58345.pdf.

Perry Hobson, J. S., & Williams, A. P. (1995). Virtual reality: a new horizon for the tourism industry. *Journal of vacation marketing*, *1*(2), 124-135.

Pessoa, L. (2009). How do emotion and motivation direct executive control?. *Trends in cognitive sciences*, *13*(4), 160-166.

Peters, C. (2010). Animating gaze shifts for virtual characters based on head movement propensity. In *2010 Second International Conference on Games and Virtual Worlds for Serious Applications* (pp. 11-18). IEEE.

Peters, C. S., & Malesky Jr, L. A. (2008). Problematic usage among highly-engaged players of massively multiplayer online role playing games. *CyberPsychology & Behavior*, *11*(4), 481-484.

Peterson, C., Park, N., & Seligman, M. E. (2005). Orientations to happiness and life satisfaction: The full life versus the empty life. *Journal of happiness studies*, *6*(1), 25-41.

Peterson, C., Ruch, W., Beermann, U., Park, N., & Seligman, M. E. (2007). Strengths of character, orientations to happiness, and life satisfaction. *The Journal of Positive Psychology*, *2*(3), 149-156.

Pfleging, B., Fekety, D. K., Schmidt, A., & Kun, A. L. (2016). A model relating pupil diameter to mental workload and lighting conditions. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 5776-5788). ACM.

Phillips, C. (2018). *Video game reward types and the player experience* (Doctoral dissertation). Queensland University of Technology.

Phillips, C., Johnson, D., & Wyeth, P. (2013). Videogame reward types. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications* (pp. 103-106). ACM.

Pilke, E. M. (2004). Flow experiences in information technology use. *International journal of human-computer studies*, *61*(3), 347-357.

Pinelle, D., Wong, N., & Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1453-1462). ACM.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). The ICLabel dataset of electroencephalographic (EEG) independent component (IC) features. *Data in Brief*, 104101.

Shuman, S. (2019). "PlayStation VR: The next wave of games coming in spring and summer 2019". Retrieved online from https://blog.us.playstation.com/2019/03/25/playstation-vr-the-next-wave-of-games-coming-in-spring-and-summer-2019/.

Plante, C. N., Gentile, D. A., Groves, C. L., Modlin, A., & Blanco-Herrera, J. (2019). Video games as coping mechanisms in the etiology of video game addiction. *Psychology of Popular Media Culture*, *8*(4), 385.

Plotnikov, A., Stakheika, N., De Gloria, A., Schatten, C., Bellotti, F., Berta, R., ... & Ansovini, F. (2012). Exploiting real-time EEG analysis for assessing flow in games. In *2012 IEEE 12th International Conference on Advanced Learning Technologies* (pp. 688-689). IEEE.

Poels, K., De Kort, Y., & IJsselsteijn, W. (2012). Identification and categorization of digital game experiences: a qualitative study integrating. *Westminster Papers in Communication and Culture*, *9*(1), 107-129.

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.

Poston, B. (2009). Maslow's hierarchy of needs. *The Surgical Technologist*, *41*(8), 347-353.

Price, T. F., & Harmon-Jones, E. (2011). Approach motivational body postures lean toward left frontal brain activity. *Psychophysiology*, *48*(5), 718-722.

Procci, K., & Bowers, C. (2011). An examination of flow and immersion in games. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 55, No. 1, pp. 2183-2187). Sage CA: Los Angeles, CA: SAGE Publications.

Procci, K., Singer, A. R., Levy, K. R., & Bowers, C. (2012). Measuring the flow experience of gamers: An evaluation of the DFS-2. *Computers in Human Behavior*, *28*(6), 2306-2312.

Proctor, S. L., & McCord, D. M. (2009). Correlates of the openness to experience domain. *Individual Differences Research*, *7*(4), 222-227.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of general psychology*, *14*(2), 154-166.

Przybylski, A. K., Ryan, R. M., & Rigby, C. S. (2009a). The motivating role of violence in video games. *Personality and Social Psychology Bulletin*, *35*(2), 243-259.

Przybylski, A. K., Weinstein, N., Ryan, R. M., & Rigby, C. S. (2009b). Having to versus wanting to play: Background and consequences of harmonious versus obsessive engagement in video games. *CyberPsychology & Behavior*, *12*(5), 485-492.

Putman, P., van Peer, J., Maimari, I., & van der Werff, S. (2010). EEG theta/beta ratio in relation to fear-modulated response-inhibition, attentional control, and affective traits. *Biological psychology*, *83*(2), 73-78.

Putze, S., Alexandrovsky, D., Putze, F., Höffner, S., Smeddinck, J. D., & Malaka, R. (2020). Breaking the experience: effects of questionnaires in vr user studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

Qin, H., Patrick Rau, P. L., & Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *Intl. Journal of Human–Computer Interaction*, *25*(2), 107-133.

Quinlan, J. R. (1996). Bagging, boosting, and C4. 5. In *AAAI/IAAI, 1*, 725-730.

Quinn, R. W. (2005). Flow in knowledge work: High performance experience in the design of national security technology. *Administrative science quarterly*, *50*(4), 610-641.

Radüntz, T. (2018). Signal quality evaluation of emerging EEG devices. *Frontiers in physiology*, *9*, 98.

Raedeke, T. D., & Stein, G. L. (1994). Felt arousal, thoughts/feelings, and ski performance. *The Sport Psychologist*, *8*(4), 360-375.

Raghunathan, R., & Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational behavior and human decision processes*, *79*(1), 56-77.

Rahman, T., Ghosh, A. K., Shuvo, M. H., & Rahman, M. M. (2015). Mental stress recognition using K-nearest neighbor (KNN) classifier on EEG signals. In *Int. Conf. Materials, Electronics & Information Engineering (ICMEIE)* (pp. 1-4).

Rakshit, R., Reddy, V. R., & Deshpande, P. (2016). Emotion detection and recognition using HRV features derived from photoplethysmogram signals. In *Proceedings of the 2nd workshop on Emotion Representations and Modelling for Companion Systems* (p. 2). ACM.

Ralph, B. C., Onderwater, K., Thomson, D. R., & Smilek, D. (2017). Disrupting monotony while increasing demand: benefits of rest and intervening tasks on vigilance. *Psychological research*, *81*(2), 432-444.

Ralph, P., & Monu, K. (2015). Toward a unified theory of digital games. *The Computer Games Journal*, *4*(1-2), 81-100.

Ramachandran, B. R. N., Pinto, S. A. R., Born, J., Winkler, S., & Ratnam, R. (2017). Measuring neural, physiological and behavioral effects of frustration. In *The 16th International Conference on Biomedical Engineering* (pp. 43-46). Springer, Singapore.

Ramírez-Gallego, S., Lastra, I., Martínez-Rego, D., Bolón-Canedo, V., Benítez, J. M., Herrera, F., & Alonso-Betanzos, A. (2017). Fast-mRMR: Fast minimum redundancy maximum relevance algorithm for high-dimensional big data. *International Journal of Intelligent Systems*, *32*(2), 134-152.

Rangaswamy, M., Porjesz, B., Chorlian, D. B., Wang, K., Jones, K. A., Bauer, L. O., ... & Begleiter, H. (2002). Beta power in the EEG of alcoholics. *Biological psychiatry*, *52*(8), 831-842.

Rani, P., Sarkar, N., & Liu, C. (2005). Maintaining optimal challenge in computer games through real-time physiological feedback. In *Proceedings of the 11th international conference on human computer interaction* (Vol. 58, pp. 22-27).

Rashid, M., Sulaiman, N., Majeed, A. P. A., Musa, R. M., Nasir, A. F. A., Bari, B. S., & Khatun, S. (2020). Current Status, Challenges, and Possible Solutions of EEG-Based Brain-Computer Interface: A Comprehensive Review. *Frontiers in neurorobotics*, *14*.

Rau, P. L. P., Tseng, Y. C., Dong, X., Jiang, C., & Chen, C. (2017). The Effect of Personality on Online Game Flow Experience and the Eye Blink Rate as an Objective Indicator. *Advances in Human-Computer Interaction*, *2017*.

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

Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., & Keltikangas-Järvinen, L. (2008). The psychophysiology of James Bond: Phasic emotional responses to violent video game events. *Emotion*, *8*(1), 114.

Recarte, M. Á., Pérez, E., Conchillo, Á., & Nunes, L. M. (2008). Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish journal of psychology*, *11*(2), 374-385.

Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. *Motivation and Emotion*, *13*(2), 83-103.

Regal, G., Voigt-Antons, J. N., Schmidt, S., Schrammel, J., Kojić, T., Tscheligi, M., & Möller, S. (2019). Questionnaires embedded in virtual environments: reliability and positioning of rating scales in virtual environments. *Quality and User Experience*, *4*(1), 1-13.

Rehbein, F., Staudt, A., Hanslmaier, M., & Kliem, S. (2016). Video game playing in the general adult population of Germany: Can higher gaming time of males be explained by gender specific genre preferences?. *Computers in Human Behavior*, *55*, 729-735.

Reilly, R. B., & Lee, T. C. (2010). Electrograms (ECG, EEG, EMG, EOG). *Technology and health care: Official journal of the European Society for Engineering and Medicine*, *18*(6), 443–458. https://doi.org/10.3233/THC-2010-0604

Reimer, B. (2009). Impact of cognitive task complexity on drivers' visual tunneling. *Transportation Research Record*, *2138*(1), 13-19.

Reuderink, B., Mühl, C., & Poel, M. (2013). Valence, arousal and dominance in the EEG during game play. *International journal of autonomous and adaptive communications systems*, *6*(1), 45-62.

Reyes del Paso, G. A., Langewitz, W., Mulder, L. J., Van Roon, A., & Duschek, S. (2013). The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: a review with emphasis on a reanalysis of previous studies. *Psychophysiology*, *50*(5), 477-487.

Reznik, S. J., & Allen, J. J. (2018). Frontal asymmetry as a mediator and moderator of emotion: An updated review. *Psychophysiology*, *55*(1), e12965.

Rieger, D., Wulf, T., Kneer, J., Frischlich, L., & Bente, G. (2014). The winner takes it all: The effect of in-game success and need satisfaction on mood repair and enjoyment. *Computers in Human Behavior, 39*, 281-286.

Rißler, R., Nadj, M., Li, M. X., Knierim, M. T., & Maedche, A. (2018). Got Flow?: Using Machine Learning on Physiological Data to Classify Flow. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (p. LBW612). ACM.

Ritterfeld, U., Cody, M., & Vorderer, P. (Eds.). (2009). *Serious games: Mechanisms and effects*. New York, NY: Taylor & Francis.

Rizzo, A., & Koenig, S. T. (2017). Is clinical virtual reality ready for primetime?. *Neuropsychology*, *31*(8), 877.

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). *Conscientiousness.* In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (p. 369–381). The Guilford Press.

Robinson, M. A. (2018). Using multi-item psychometric scales for research and practice in human resource management. *Human Resource Management*, *57*(3), 739-750.

Robinson, R., Murray, J., & Isbister, K. (2018). " You're Giving Me Mixed Signals!" A Comparative Analysis of Methods that Capture Players' Emotional Response to Games. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-6).

Rocha, J. B., Mascarenhas, S., & Prada, R. (2008). Game mechanics for cooperative games. *ZON Digital Games 2008*, 72-80.

Rodríguez-Sánchez, A., Salanova, M., Cifre, E., & Schaufeli, W. B. (2011). When good is good: A virtuous circle of self-efficacy and flow at work among teachers. *Revista de psicologia social*, *26*(3), 427-441.

Rodríguez - Sánchez, A. M., Schaufeli, W. B., Salanova, M., & Cifre, E. (2008). Flow experience among information and communication technology users. *Psychological Reports*, *102*(1), 29-39.

Roest, M., & Bakkes, S. (2015). Engaging Casual Games That Frustrate You. *FDG*, 2015.

Rogatko, T. P. (2009). The influence of flow on positive affect in college students. *Journal of happiness studies*, *10*(2), 133.

Rogers, J. M., Jensen, J., Valderrama, J. T., Johnstone, S. J., & Wilson, P. H. (2020). Single-channel EEG measurement of engagement in virtual rehabilitation: a validation study. *Virtual Reality*, 1-10.

Rogers, K., Ribeiro, G., Wehbe, R. R., Weber, M., & Nacke, L. E. (2018). Vanishing Importance: Studying Immersive Effects of Game Audio Perception on Player Experiences in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 328). ACM.

Ronimus, M., Kujala, J., Tolvanen, A., & Lyytinen, H. (2014). Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. *Computers & Education*, *71*, 237-246.

Roos, C., Van der Pal, J., & Sewnath, G. (2017). Maintaining optimal training value using Brain Computer Interfacing. *ITEC*, 2017.

Roque, N. A., & Boot, W. R. (2018). Action Video Games DO NOT Promote Visual Attention. In *Video Game Influences on Aggression, Cognition, and Attention* (pp. 105-118). Springer, Cham.

Rosenfield, M., Jahan, S., Nunez, K., & Chan, K. (2015). Cognitive demand, digital screens and blink rate. *Computers in Human Behavior*, *51*, 403-406.

Ross, S. R., & Keiser, H. N. (2014). Autotelic personality through a five-factor lens: Individual differences in flow-propensity. *Personality and individual differences*, *59*, 3-8.

Roth, C., Vorderer, P., Klimmt, C., & Vermeulen, I. (2010). Measuring the user experience in narrative-rich games: towards a concept-based assessment for interactive stories. *Entertainment Interfaces*.

Rousson, V., & Gasser, T. (2004). Simple component analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *53*(4), 539-555.

Ruckenstein, M. (1991). Homo ludens: a study of the play element in culture. *Leisure and Ethics*, *237*.

Russell, J. A. (2003a). Core affect and the psychological construction of emotion. *Psychological review*, *110*(1), 145.

Russell, J. A. (2003b). Introduction: The return of pleasure. *Cognition and Emotion*, *17*(2), 161-165.

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54-67.

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, *30*(4), 344-360.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of personality and social psychology*, *57*(6), 1069.

Sá, C., Gomes, P. V., Marques, A., & Correia, A. (2020). The Use of Portable EEG Devices in Development of Immersive Virtual Reality Environments for Converting Emotional States into Specific Commands. In *Multidisciplinary Digital Publishing Institute Proceedings* (Vol. 54, No. 1, p. 43).

Saadé, R. G., & Kira, D. (2009). Computer anxiety in e-learning: The effect of computer self-efficacy. *Journal of Information Technology Education: Research*, *8*(1), 177-191.

Sabet, S. S., Griwodz, C., & Möller, S. (2019). Influence of primacy, recency and peak effects on the game experience questionnaire. In *Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems* (pp. 22-27). ACM.

Saboul, D., Balducci, P., Millet, G., Pialoux, V., & Hautier, C. (2016). A pilot study on quantification of training load: The use of HRV in training practice. *European journal of sport science*, *16*(2), 172-181.

Sacco, D. F., & Ismail, M. M. (2014). Social belongingness satisfaction as a function of interaction medium: Face-to-face interactions facilitate greater social belonging and interaction enjoyment compared to instant messaging. *Computers in Human Behavior*, *36*, 359-364.

Sackett, A. M., Meyvis, T., Nelson, L. D., Converse, B. A., and Sackett, A. L. (2010). You're having fun when time flies: the hedonic consequences of subjective time progression. *Psychol. Sci.* 21, 111–117. doi: 10.1177/0956797609354832

Sahoo, F. M., & Sahu, R. (2009). The role of flow experience in human happiness. *Journal of the Indian Academy of Applied Psychology*, *35*, 40-47.

Sala, G., Tatlidil, K. S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological bulletin*, *144*(2), 111.

Salminen, M., Kivikangas, J. M., Ravaja, N., & Kallinen, K. (2009). Frontal EEG asymmetry in the study of player experiences during competitive and cooperative dual play. In *IADIS International Conference Game and Entertainment Technologies* (pp. 44-50).

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

Salamone, J. D., & Correa, M. (2016). Neurobiology of effort and the role of mesolimbic dopamine. In *Recent Developments in Neuroscience Research on Human Motivation* (pp. 229-256). Emerald Group Publishing Limited.

Sanchez, V. L. A., Cruz, O. I. I., Solorza, E. A. A., Monroy, I. A. E., Caro, K., & Castro, L. A. (2017). BeeSmart: A Gesture-Based Videogame to Support Literacy and Eye-Hand Coordination of Children with Down Syndrome. In *International Conference on Games and Learning Alliance* (pp. 43-53). Springer, Cham.

Sánchez-Maroño, N., Alonso-Betanzos, A., & Tombilla-Sanromán, M. (2007). Filter methods for feature selection–a comparative study. In *International Conference on Intelligent Data Engineering and Automated Learning* (pp. 178-187). Springer, Berlin, Heidelberg.

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332.

Sanders, T., & Cairns, P. (2010). Time perception, immersion and music in videogames. In *Proceedings of the 24th BCS interaction specialist group conference* (pp. 160-167). British Computer Society.

Saputra, R., & Iqbal, B. M. (2017). Stress emotion evaluation in multiplayer online battle arena (MOBA) video game related to gaming rules using electroencephalogram (EEG). In *Proceedings of the 2017 4th International Conference on Biomedical and Bioinformatics Engineering* (pp. 74-77).

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: reactions to tests. *Journal of personality and social psychology*, *46*(4), 929.

Sarkar, S. (2014). “Blizzard reaches 100M lifetime World of Warcraft accounts”. Retrieved from https://www.polygon.com/2014/1/28/5354856/world-of-warcraft-100m-accounts-lifetime (November 12, 2019).

Saunders, C., Rutkowski, A. F., Genuchten van, M., Vogel, D., & Orrego, J. M. (2011). Virtual space and place: Theory and test. *MIS Quarterly*, 1079-1098.

Sauseng, P., Hoppe, J., Klimesch, W., Gerloff, C., & Hummel, F. C. (2007). Dissociation of sustained attention from central executive functions: local activity and interregional connectivity in the theta range. *European Journal of Neuroscience*, *25*(2), 587-593.

Sawangjai, P., Hompoonsup, S., Leelaarporn, P., Kongwudhikunakorn, S., & Wilaiprasitporn, T. (2019). Consumer grade EEG measuring sensors as research tools: A review. *IEEE Sensors Journal*, *20*(8), 3996-4024.

Scharkow, M., Festl, R., Vogelgesang, J., & Quandt, T. (2015). Beyond the “core-gamer”: Genre preferences and gratifications in computer games. *Computers in Human Behavior*, *44*, 293-298.

Scheeringa, R., Bastiaansen, M. C., Petersson, K. M., Oostenveld, R., Norris, D. G., & Hagoort, P. (2008). Frontal theta EEG activity correlates negatively with the default mode network in resting state. *International journal of psychophysiology*, *67*(3), 242-251.

Schmid, S. E. (2011). Beyond autotelic play. *Journal of the Philosophy of Sport*, *38*(2), 149-166.

Schmidt, K. L., Ambadar, Z., Cohn, J. F., & Reed, L. I. (2006). Movement differences between deliberate and spontaneous facial expressions: Zygomaticus major action in smiling. *Journal of Nonverbal Behavior*, *30*(1), 37-52.

Schmidtke, J. I., & Heller, W. (2004). Personality, affect and EEG: predicting patterns of regional brain activity related to extraversion and neuroticism. *Personality and Individual Differences*, *36*(3), 717-732.

Schneider, E. F., Lang, A., Shin, M., & Bradley, S. D. (2004). Death with a story: How story impacts emotional, motivational, and physiological responses to first-person shooter video games. *Human communication research*, *30*(3), 361-375.

Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in cognitive sciences*, *6*(8), 339-344.

Schooler, J. W., & Mauss, I. B. (2010). To be happy and to know it: The experience and meta-awareness of pleasure. *Pleasures of the brain*, 244-254.

Schüler, J. (2007). Arousal of flow experience in a learning setting and its effects on exam performance and affect. *Zeitschrift für Pädagogische Psychologie*, *21*(3/4), 217-227.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*(2), 241-263.

Schutte, I., Kenemans, J. L., & Schutter, D. J. (2017). Resting-state theta/beta EEG ratio is associated with reward-and punishment-related reversal learning. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(4), 754-763.

Secerbegovic, A., Ibric, S., Nisic, J., Suljanovic, N., & Mujcic, A. (2017). Mental workload vs. stress differentiation using single-channel EEG. In *CMBEBIH 2017* (pp. 511-515). Springer, Singapore.

Sedeeq, A., & Romano, D. (2014). Brain controlled games: Concepts and classifiers. In *Proceedings of the 50th Annual Convention of the Artificial Intelligence and Simulation Behaviour* (pp. 1-5).

Sedghamiz, H. (2014). Matlab implementation of Pan Tompkins ECG QRS detector. Retrieved from https://uk.mathworks.com/matlabcentral/fileexchange/45840-complete-pan-tompkins-implementation-ecg-qrs-detector.

Seib, H. M., & Vodanovich, S. J. (1998). Cognitive correlates of boredom proneness: The role of private self-consciousness and absorption. *The Journal of psychology*, *132*(6), 642-652.

Seibert, J. (2014). *An exploratory study on virtual reality head mounted displays and their impact on player presence* (Doctoral dissertation). Baylor University, Waco, TX, United States.

Selye, H. (1975). *The stress of life* (Revised edition). New York, NY: McGraw Hill Book Co.

Seo, J., Laine, T. H., & Sohn, K. A. (2019). Machine learning approaches for boredom classification using EEG. *Journal of Ambient Intelligence and Humanized Computing*, 1-16.

Seok, S., & DaCosta, B. (2015). Predicting video game behavior: An investigation of the relationship between personality and mobile game play. *Games and Culture*, *10*(5), 481-501.

Setzer, V. W., & Duckett, G. E. (1994). The risks to children using electronic games. *Retrieved April*, *8*, 2003.

Shafer, D. M., Carbonara, C. P., & Korpi, M. F. (2019). Factors affecting enjoyment of virtual reality games: a comparison involving consumer-grade virtual reality technology. *Games for health journal*, *8*(1), 15-23.

Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays*, *29*(2), 58-69.

Shearer, P. B. (2016). *Physiological detection of flow* (Doctoral dissertation). University of South Dakota: United States.

Sheikholeslami, C., Yuan, H., He, E. J., Bai, X., Yang, L., & He, B. (2007). A high resolution EEG study of dynamic brain activity during video game play. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE* (pp. 2489-2491). IEEE.

Shelley, B. (2001). Guidelines for developing successful games. *Gamasutra*, Available from: https://www.gamasutra.com/view/feature/131450/guidelines_for_developing_.php.

Shelstad, W. J., Smith, D. C., & Chaparro, B. S. (2017). Gaming on the rift: How virtual reality affects game user satisfaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 2072-2076). Sage CA: Los Angeles, CA: SAGE Publications.

Sherry, J. L. (2004). Flow and media enjoyment. *Communication theory*, *14*(4), 328-347.

Shin, M., Heard, R., Suo, C., & Chow, C. M. (2012). Positive emotions associated with "Counter-Strike" game playing. *Games for Health: Research, Development, and Clinical Applications*, *1*(5), 342-347.

Shin, N. (2006). Online learner's 'flow'experience: an empirical study. *British Journal of Educational Technology*, *37*(5), 705-720.

Shirzad, N., & Van der Loos, H. M. (2016). Evaluating the user experience of exercising reaching motions with a robot that predicts desired movement difficulty. *Journal of motor behavior*, *48*(1), 31-46.

Shute, V. J., Ventura, M., Bauer, M., & Zapata-Rivera, D. (2009). Melding the power of serious games and embedded assessment to monitor and foster learning. *Serious games: Mechanisms and effects*, *2*, 295-321.

Sillaots, M., & Jesmin, T. (2016). Multiple Regression Analysis: Refinement of the Model of Flow. In *European Conference on Games Based Learning* (p. 609). Academic Conferences International Limited.

Sillaots, M., Jesmin, T., Fiadotau, M., & Khulbe, M. (2020). Gamifying Classroom Presentations: Evaluating the Effects on Engagement Across Demographic Factors. In *European Conference on Games Based Learning* (pp. 537-XIX). Academic Conferences International Limited.

Silpasuwanchai, C., and Ren, X. (2018). A Quick Look at Game Engagement Theories. In K. L. Norman, & J. Kirakowski (Eds.), *The Wiley Handbook of Human Computer Interaction* (pp. 657–679). Hoboken, NJ: Wiley.

Silvern, S. B., & Williamson, P. A. (1987). The effects of video game play on young children's aggression, fantasy, and prosocial behavior. *Journal of Applied Developmental Psychology*, *8*(4), 453-462.

Singer, R. N. (2002). Preperformance state, routines, and automaticity: what does it take to realize expertise in self-paced events?. *Journal of Sport and Exercise Psychology*, *24*(4), 359-375.

Singh, N. (2013). *Arousal Level Determination in Video Game Playing Using Galvanic Skin Response* (Doctoral dissertation). Thapar University, Patiala, India.

Singh, S., Barto, A. G., & Chentanez, N. (2005). Intrinsically motivated reinforcement learning. In *Advances in neural information processing systems* (pp. 1281-1288).

Sitompul, B. J. D., Sitompul, O. S., & Sihombing, P. (2019). Enhancement Clustering Evaluation Result of Davies-Bouldin Index with Determining Initial Centroid of K-Means Algorithm. In *Journal of Physics: Conference Series* (Vol. 1235, No. 1, p. 012015). IOP Publishing.

Sjöblom, M., & Hamari, J. (2017). Why do people watch others play video games? An empirical study on the motivations of Twitch users. *Computers in Human Behavior*, *75*, 985-996.

Skadberg, Y. X., and Kimmel, J. R. (2004). Visitors' flow experience while browsing a Web site: its measurement, contributing factors and consequences. *Comp. Hum. Behav.* 20, 403–422. doi: 10.1016/S0747-5632(03)00050-5

Skalski, P., Tamborini, R., Shelton, A., Buncher, M., & Lindmark, P. (2011). Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. *New Media & Society*, *13*(2), 224-242.

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, *3*, 74.

Smith, L. H., Hargrove, L. J., Lock, B. A., & Kuiken, T. A. (2010). Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *19*(2), 186-192.

Smith, R. P. (1981). Boredom: A review. *Human factors*, *23*(3), 329-340.

Smits, C. H., Deeg, D. J., & Bosscher, R. J. (1995). Well-being and control in older persons: The prediction of well-being from control measures. *The International Journal of Aging and Human Development*, *40*(3), 237-251.

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems* (pp. 2951-2959).

Snyder, E.E., & Spreitzer, E. (1974). Involvement in sports and psychological well being. *International Journal of Sport Psychology*, 5, 28-39.

So, R. H., Lo, W. T., & Ho, A. T. (2001). Effects of navigation speed on motion sickness caused by an immersive virtual environment. *Human factors*, *43*(3), 452-461.

Soetaert, R., Bourgonjon, J., & Rutten, K. (2011). Video games as equipment for living. *CLCWeb-Comparative Literature and Culture*, *13*(3), 1-9.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, *45*(4), 427-437.

Song, M. K., Lin, F. C., Ward, S. E., & Fine, J. P. (2013). Composite variables: when and how. *Nursing research*, *62*(1), 45.

Sony Interactive Entertainment (2019). Cumulative Worldwide Hardware Unit Sales (Sell-in). Retrieved from https://www.sie.com/en/corporate/data.html. Accessed on 27/12/2019.

Soutter, A. R. B., & Hitchens, M. (2016). The relationship between character identification and flow state within video games. *Computers in Human Behavior*, *55*, 1030-1038.

Speelman, C., & Maybery, M. (2013). Automaticity and skill acquisition. In *Implicit and explicit mental processes* (pp. 89-108). Psychology Press.

Spence, I., and Feng, J. (2010). Video games and spatial cognition. *Rev. Gen. Psychol.* 14:92. doi: 10.1037/a0019491

Spielberg, J. M., Stewart, J. L., Levin, R. L., Miller, G. A., & Heller, W. (2008). Prefrontal cortex, emotion, and approach/withdrawal motivation. *Social and personality psychology compass*, *2*(1), 135-153.

Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F. S., & Natalicio, D. S. (1971). Development of the Spanish edition of the state-trait anxiety inventory. *Interamerican Journal of Psychology*, *5*(3-4), 145-158.

Spring, D. (2015). Gaming history: computer and video games as historical scholarship. *Rethinking History*, *19*(2), 207-221.

Srinivasan, R. (1999). Methods to improve the spatial resolution of EEG. *International Journal of Bioelectromagnetism*, *1*(1), 102-111.

Stavrou, N. A. (2008). Intrinsic motivation, extrinsic motivation and amotivation: Examining self-determination theory from flow theory perspective. *New developments in the psychology of motivation*, 1-24.

Stavrou, N. A., Jackson, S. A., Zervas, Y., & Karteroliotis, K. (2007). Flow experience and athletes' performance with reference to the orthogonal model of flow. *The Sport Psychologist*, *21*(4), 438-457.

Steinkuehler, C., & Duncan, S. (2008). Scientific habits of mind in virtual worlds. *Journal of Science Education and Technology*, *17*(6), 530-543.

Stemmler, G., & Wacker, J. (2010). Personality, emotion, and individual differences in physiological responses. *Biological psychology*, *84*(3), 541-551.

Stenberg, G. (1992). Personality and the EEG: Arousal and emotional arousability. *Personality and individual differences*, *13*(10), 1097-1113.

Steptoe, A. S., & Butler, N. (1996). Sports participation and emotional wellbeing in adolescents. *The Lancet*, *347*(9018), 1789-1792.

Stern, J. A., & Skelly, J. J. (1984). The eye blink and workload considerations. In *Proceedings of the Human Factors Society Annual Meeting* (Vol. 28, No. 11, pp. 942-944). Sage CA: Los Angeles, CA: SAGE Publications.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: a possible measure of fatigue. *Human factors*, *36*(2), 285-297.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, *42*(4), 73-93.

Stoll, O. (2018). Peak Performance, the Runner's High and Flow. In *Handbook of sports and exercise psychology*. American Psychology Association Washington, DC.

Stolz, C., Endres, D., & Mueller, E. M. (2019). Threat-conditioned contexts modulate the late positive potential to faces—A mobile EEG/virtual reality study. *Psychophysiology*, *56*(4), e13308.

Strenge, H., Kropp, P., Hoffmeister, J., Verri, A., Galli, C., & Gerber, W. D. (1999). Contingent negative variation indicates phasic arousal for externally cued unilateral eye blink in human subjects. *Neuroscience letters*, *266*(1), 57-60.

Stryker, S., & Burke, P. J. (2000). The past, present, and future of an identity theory. *Social psychology quarterly*, 284-297.

Subahni, A. R., Xia, L., & Malik, A. S. (2012). Association of mental stress with video games. In *2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012)* (Vol. 1, pp. 82-85). IEEE.

Sulea, C., Van Beek, I., Sarbescu, P., Virga, D., & Schaufeli, W. B. (2015). Engagement, boredom, and burnout among students: Basic need satisfaction matters more than personality traits. *Learning and Individual Differences*, *42*, 132-138.

Sun, S., Zhang, C., & Zhang, D. (2007). An experimental evaluation of ensemble methods for EEG signal classification. *Pattern Recognition Letters*, *28*(15), 2157-2163.

Suresh, K. P. (2011). An overview of randomization techniques: an unbiased assessment of outcome in clinical research. *Journal of human reproductive sciences*, *4*(1), 8.

Sutcliffe, A. G., & Kaur, K. D. (2000). Evaluating the usability of virtual reality user interfaces. *Behaviour & Information Technology*, *19*(6), 415-426.

Suter, S. E., Huggenberger, H. J., Richter, S., Blumenthal, T. D., & Schachinger, H. (2009). Left side cradling of an appetitive doll is associated with higher heart rate variability and attenuated startle in nulliparous females. *International Journal of Psychophysiology*, *74*(1), 53-57.

Sutoyo, R., Winata, D., Oliviani, K., & Supriyadi, D. M. (2015). Dynamic difficulty adjustment in tower defence. *Procedia Computer Science*, *59*, 435-444.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological science*, *8*(3), 204-210.

Swann, C., Keegan, R. J., Piggott, D., & Crust, L. (2012). A systematic review of the experience, occurrence, and controllability of flow states in elite sport. *Psychology of Sport and Exercise*, *13*(6), 807-819.

Swann, C., Piggott, D., Schweickle, M., & Vella, S. A. (2018). A review of scientific progress in flow in sport and exercise: normal science, crisis, and a progressive shift. *Journal of Applied Sport Psychology*, *30*(3), 249-271.

Sweetser, P., & Johnson, D. (2019). GameFlow and Player Experience Measures: An Initial Comparison of Conceptual Constructs. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction* (pp. 317-321).

Sweetser, P., Johnson, D. M., & Wyeth, P. (2012). Revisiting the GameFlow model with detailed heuristics. *Journal: Creative Technologies*, *2012*(3).

Sweetser, P., Johnson, D., Wyeth, P., Anwar, A., Meng, Y., & Ozdowska, A. (2017). GameFlow in different game genres and platforms. *Computers in Entertainment (CIE)*, *15*(3), 1.

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

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Berlin: Springer.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, *10*(3), 251-296.

Szablewicz, M. (2010). The ill effects of "opium for the spirit": a critical cultural analysis of China's Internet addiction moral panic. *Chinese Journal of Communication*, *3*(4), 453-470.

Taft, L. M., Evans, R. S., Shyu, C. R., Egger, M. J., Chawla, N., Mitchell, J. A., ... & Varner, M. (2009). Countering imbalanced datasets to improve adverse drug event predictive models in labor and delivery. *Journal of biomedical informatics*, *42*(2), 356-364.

Tajerian, M. (2012). Fight or flight: The neuroscience of survival horror. http://www.gamasutra.com/view/feature/172168/.

Takatalo, J., Häkkinen, J., Kaistinen, J., & Nyman, G. (2010). Presence, involvement, and flow in digital games. In *Evaluating user experience in games* (pp. 23-46). Springer, London.

Tamborini, R., Bowman, N. D., Eden, A., Grizzard, M., & Organ, A. (2010). Defining media enjoyment as the satisfaction of intrinsic needs. Journal of Communication, 60, 758–777. doi:10.1111/j.1460-2466.2010.01513.x.

Tan, C. T., Leong, T. W., & Shen, S. (2014). Combining think-aloud and physiological data to understand video game experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 381-390). ACM.

Tan, C. T., Leong, T. W., Shen, S., Dubravs, C., & Si, C. (2015). Exploring gameplay experiences on the Oculus Rift. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (pp. 253-263). ACM.

Tan, C. T., & Pisan, Y. (2012). Towards Automated Player Experience Detection With Computer Vision Techniques. In *CHI 2012 Workshop on Game User Experience*.

Tanaka, Y., & Yamaoka, K. (1993). Blink activity and task difficulty. *Perceptual and Motor Skills*, *77*(1), 55-66.

Tandle, A., Jog, N., Dharmadhikari, A., & Jaiswal, S. (2016). Estimation of valence of emotion from musically stimulated EEG using frontal theta asymmetry. In *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* (pp. 63-68). IEEE.

Tao, D., Zhang, X., Cai, J., Tan, H., Zhang, X., & Zhang, T. (2019). Physiological Measures of Mental Workload: Evidence from Empirical Studies. In *International Conference on Man-Machine-Environment System Engineering* (pp. 217-225). Springer, Singapore.

Tatalović Vorkapić, S., & Gović, A. (2016). Relationship between flow and personality traits among preschool teachers. *Metodički obzori*, *11*(23), 24-40.

Tauscher, J. P., Schottky, F. W., Grogorick, S., Bittner, P. M., Mustafa, M., & Magnor, M. (2019). Immersive EEG: Evaluating electroencephalography in virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 1794-1800). IEEE.

Tay, T. (2016). *Physiological Correlates of the "Flow" Experience: EEG Measures of Attentional State* (Doctoral dissertation). University of Nevada, Reno, United States.

Taylor, L. N. (2002). *Video Games: Perspective, Point-of-view, and Immersion* (Master Thesis). University of Florida, Gainesville, FL.

Taywade, S. A., & Raut, R. D. (2014). A review: EEG signal analysis with different methodologies. In *Proceedings of the National Conference on Innovative Paradigms in Engineering and Technology (NCIPET'12)* (pp. 29-31).

Tax, D. M., Van Breukelen, M., Duin, R. P., & Kittler, J. (2000). Combining multiple classifiers by averaging or by multiplying?. *Pattern recognition*, *33*(9), 1475-1485.

Teipen, C. (2016). The Implications of the Value Chain and Financial Institutions for Work and Employment: Insights from the Video Game Industry in Poland, Sweden and Germany. *British Journal of Industrial Relations*, *54*(2), 311-333.

Teng, C. I. (2008). Personality differences between online game players and non players in a student sample. *Cyberpsychology & Behavior*, *11*(2), 232–234.

Teng, C. I. (2011). Who are likely to experience flow? Impact of temperament and character on flow. *Personality and Individual Differences*, *50*(6), 863-868.

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement science review*, *2*(2), 1-11.

Thorbjørnsen, H., Pedersen, P. E., & Nysveen, H. (2007). "This is who I am": Identity expressiveness and the theory of planned behavior. *Psychology & Marketing*, *24*(9), 763-785.

Thurber, M. R., Bodenhamer-Davis, E., Johnson, M., Chesky, K., & Chandler, C. K. (2010). Effects of heart rate variability coherence biofeedback training and emotional management techniques to decrease music performance anxiety. *Biofeedback*, *38*(1), 28-40.

Tian, Y., Bian, Y., Han, P., Wang, P., Gao, F., & Chen, Y. (2017). Physiological signal analysis for evaluating flow during playing of computer games of varying difficulty. *Frontiers in psychology*, *8*, 1121.

Todd, D. A., McCullagh, P. J., Mulvenna, M. D., & Lightbody, G. (2012). Investigating the use of brain-computer interaction to facilitate creativity. In *Proceedings of the 3rd Augmented Human International Conference* (pp. 1-8).

Tognetti, S., Garbarino, M., Bonarini, A., & Matteucci, M. (2010). Modeling enjoyment preference from physiological responses in a car racing game. In *Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games* (pp. 321-328). IEEE.

Toner, J., Montero, B. G., & Moran, A. (2015). The perils of automaticity. *Review of General Psychology*, *19*(4), 431-442.

Tosca, S., & Klastrup, L. (2009). "Because it just looks cool!"-Fashion as character performance: The Case of WoW. *Journal For Virtual Worlds Research*, *1*(3).

Tosun, L. P., & Lajunen, T. (2009). Why do young adults develop a passion for Internet activities? The associations among personality, revealing "true self" on the Internet, and passion for the Internet. *CyberPsychology & Behavior*, *12*(4), 401-406.

Tozman, T., Magdas, E. S., MacDougall, H. G., & Vollmeyer, R. (2015). Understanding the psychophysiology of flow: A driving simulator experiment to investigate the relationship between flow and heart rate variability. *Computers in Human Behavior*, *52*, 408-418.

Tran, D., Mac, H., Tong, V., Tran, H. A., & Nguyen, L. G. (2018). A LSTM based framework for handling multiclass imbalance in DGA botnet detection. *Neurocomputing*, *275*, 2401-2413.

Trepte, S., & Reinecke, L. (2011). The pleasures of success: Game-related efficacy experiences as a mediator between player performance and game enjoyment. *Cyberpsychology, Behavior, and Social Networking*, *14*(9), 555-557.

Triberti, S. (2016). This Drives Me Nuts!: How Gaming Technologies Can Elicit Positive Experiences by Means of Negative Emotions. In *Psychology and Mental Health: Concepts, Methodologies, Tools, and Applications* (pp. 271-294). IGI Global.

Ullén, F., de Manzano, Ö., Almeida, R., Magnusson, P. K., Pedersen, N. L., Nakamura, J., ... & Madison, G. (2012). Proneness for psychological flow in everyday life: Associations with personality and intelligence. *Personality and Individual Differences*, *52*(2), 167-172.

Ullén, F., de Manzano, Ö., Theorell, T., & Harmat, L. (2010). The physiology of effortless attention. In B. Bruya (Ed.), *Effortless attention* (pp. 205–218). Cambridge, MA, USA: The MIT Press.

Ullén, F., Harmat, L., Theorell, T., & Madison, G. (2016). Flow and individual differences–a phenotypic analysis of data from more than 10,000 twin individuals. In *Flow Experience* (pp. 267-288). Springer International Publishing.

Ulrich, M., Keller, J., & Grön, G. (2016). Neural signatures of experimentally induced flow experiences identified in a typical fMRI block design with BOLD imaging. *Social cognitive and affective neuroscience*, *11*(3), 496-507.

Ulrich, M., Keller, J., Hoenig, K., Waller, C., and Grön, G. (2014). Neural correlates of experimentally induced flow experiences. *Neuroimage* 86, 194–202. doi: 10.1016/j.neuroimage.2013.08.019

Utz, S., Jonas, K. J., & Tonkens, E. (2012). Effects of passion for Massively Multiplayer Online Role-playing games on interpersonal relationships. *Journal of Media Psychology*, 24(2), 77–86.

Uusberg, H., Allik, J., & Hietanen, J. K. (2015). Eye contact reveals a relationship between Neuroticism and anterior EEG asymmetry. *Neuropsychologia*, *73*, 161-168.

Vallat, R., Eichenlaub, J. B., Nicolas, A., & Ruby, P. (2018). Dream recall frequency is associated with medial prefrontal cortex white-matter density. *Frontiers in psychology*, *9*, 1856.

Vallerand, R. J. (2008). On the psychology of passion: In search of what makes people's lives most worth living. *Canadian Psychology*, *49*, 1-13.

Vallerand, R. J., Blanchard, C., Mageau, G. A., Koestner, R., Ratelle, C., Léonard, M., ... & Marsolais, J. (2003). Les passions de l'ame: on obsessive and harmonious passion. *Journal of personality and social psychology*, *85*(4), 756.

Van den Hoogen, W. M., Poels, K., IJsselsteijn, W., & de Kort, Y. (2012). Between challenge and defeat: Repeated player-death and game enjoyment. *Media Psychology*, *15*(4), 443-459.

Van den Hout, J. J., Davis, O. C., & Weggeman, M. C. (2018). The conceptualization of team flow. *The Journal of psychology*, *152*(6), 388-423.

Van Hooft, E. A., & Van Hooff, M. L. (2018). The state of boredom: Frustrating or depressing?. *Motivation and emotion*, *42*(6), 931-946.

Van Kerrebroeck, H., Brengman, M., & Willems, K. (2017). When brands come to life: experimental research on the vividness effect of Virtual Reality in transformational marketing communications. *Virtual Reality*, *21*(4), 177-191.

Van Lankveld, G., Schreurs, S., Spronck, P., & Van Den Herik, J. (2010). Extraversion in games. In *International Conference on Computers and Games* (pp. 263-275). Springer, Berlin, Heidelberg.

Van Lankveld, G., Spronck, P., Van Den Herik, H. J., & Rauterberg, M. (2009). Incongruity-based adaptive game balancing. In *Advances in computer games* (pp. 208-220). Springer, Berlin, Heidelberg.

Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T.-P. (2001). Activity correlates of workload during a visuospatial memory task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *43*, 111–121.

Van Ravenzwaaij, D., Boekel, W., Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2014). Action video games do not improve the speed of information processing in simple perceptual tasks. *Journal of Experimental Psychology: General*, *143*(5), 1794.

Varvatsoulias, G. (2014). Discussion on the Claim That the Human Mind is a Computational Process Device. *Psychological Thought*, *7*(2), 127–133.

Veltman, J. A., & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological psychology*, *42*(3), 323-342.

Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, *41*(5), 656-669.

Vermeulen, L., & Van Looy, J. (2016). "I play so I am?" A gender study into stereotype perception and genre choice of digital game players. *Journal of Broadcasting & Electronic Media*, *60*(2), 286-304.

Vicencio-Moreira, R., Mandryk, R. L., Gutwin, C., & Bateman, S. (2014). The effectiveness (or lack thereof) of aim-assist techniques in first-person shooter games. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 937-946). ACM.

Vicente, J., Laguna, P., Bartra, A., & Bailón, R. (2016). Drowsiness detection using heart rate variability. *Medical & biological engineering & computing*, *54*(6), 927-937.

Vo, T. H. T., Dang, T. L. N., Nguyen, N. V. T., & Van Huynh, T. (2019). Classification Electroencephalography Using Machine Learning. In *2019 19th International Symposium on Communications and Information Technologies (ISCIT)* (pp. 14-19). IEEE.

Vodanovich, S. J. (2003). Psychometric measures of boredom: A review of the literature. *The Journal of psychology*, *137*(6), 569-595.

Vodanovich, S. J., Verner, K. M., & Gilbride, T. V. (1991). Boredom proneness: Its relationship to positive and negative affect. *Psychological reports*, *69*(3), 1139-1146.

Vogeley, K., and Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends Cogn. Sci.* 7, 38–42. doi: 10.1016/S1364-6613(02)00003-7

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., and Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *J. Cogn. Neurosci.* 16, 817–827. doi: 10.1162/089892904970799

Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience*, *6*(7), 533.

Voiskounsky, A., & Wang, S. (2014). Flow experience while computer gaming: Empirical study. *Open Journal of Social Sciences*, *2*(8), 1.

Von Gemmingen, M. J., Sullivan, B. F., & Pomerantz, A. M. (2003). Investigating the relationships between boredom proneness, paranoia, and self-consciousness. *Personality and Individual Differences*, *34*(6), 907-919.

Von Rosenberg, W., Chanwimalueang, T., Goverdovsky, V., Looney, D., Sharp, D., & Mandic, D. P. (2016). Smart helmet: Wearable multichannel ECG and EEG. *IEEE journal of translational engineering in health and medicine*, *4*.

Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication theory*, *14*(4), 388-408.

Vriend, S. (2017). Can your game have too many choices?. *Gamasutra*, Available at: https://www.gamasutra.com/blogs/SitaVriend/20170818/303852/Can_your_game_have_too_many_choices.php.

Wacker, J. (2018). Effects of positive emotion, extraversion, and dopamine on cognitive stability-flexibility and frontal EEG asymmetry. *Psychophysiology*, *55*(1), e12727.

Wacker, J., Chavanon, M. L., & Stemmler, G. (2010). Resting EEG signatures of agentic extraversion: New results and meta-analytic integration. *Journal of Research in Personality*, *44*(2), 167-179.

Walker, C. J. (2010). Experiencing flow: Is doing it together better than doing it alone?. *The Journal of Positive Psychology*, *5*(1), 3-11.

Walker, J. E., Kozlowski, G. P., & Lawson, R. (2007). A modular activation/coherence approach to evaluating clinical/QEEG correlations and for guiding neurofeedback training: modular insufficiencies, modular excesses, disconnections, and hyperconnections. *Journal of Neurotherapy*, *11*(1), 25-44.

Wallace, P. M. (1999). *The psychology of the Internet*. New York, NY: Cambridge University Press.

Wan, C. S., & Chiou, W. B. (2006). Psychological motives and online games addiction: A test of flow theory and humanistic needs theory for Taiwanese adolescents. *CyberPsychology & Behavior*, *9*(3), 317-324.

Wang, C. A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., & Munoz, D. P. (2018a). Arousal Effects on Pupil Size, Heart Rate, and Skin Conductance in an Emotional Face Task. *Frontiers in neurology*, *9*.

Wang, Y. K., Jung, T. P., & Lin, C. T. (2018b). Theta and alpha oscillations in attentional interaction during distracted driving. *Frontiers in behavioral neuroscience*, *12*, 3.

Wang, C. K. J., Khoo, A., Liu, W. C., & Divaharan, S. (2008). Passion and intrinsic motivation in digital gaming. *CyberPsychology & Behavior*, *11*(1), 39-45.

Wang, C. K. J., Liu, W. C., & Khoo, A. (2009). The psychometric properties of dispositional flow scale-2 in internet gaming. *Current Psychology*, *28*(3), 194-201.

Wang, H., & Liu, S. (2016). An effective feature selection approach using the hybrid filter wrapper. *International Journal of Hybrid Information Technology*, *9*(1), 119-128.

Wang, H., & Sun, C. T. (2011). Game reward systems: Gaming experiences and social meanings. In *DiGRA Conference* (pp. 1-15).

Wang, X. W., Nie, D., & Lu, B. L. (2011). EEG-based emotion recognition using frequency domain features and support vector machines. In *International conference on neural information processing* (pp. 734-743). Springer, Berlin, Heidelberg.

Wanyan, X., Zhuang, D., Lin, Y., Xiao, X., & Song, J. W. (2018). Influence of mental workload on detecting information varieties revealed by mismatch negativity during flight simulation. *International Journal of Industrial Ergonomics*, *64*, 1-7.

Warmelink, H., Harteveld, C., & Mayer, I. (2009). Press Enter or Escape to Play-Deconstructing Escapism in Multiplayer Gaming. In B. Atkins, H. Kennedy, & T. Krzywinska (Eds.), *Breaking new ground: Innovation in games, play, practice and theory*: P*roceedings of DiGRA 2009*, September 2009, London, UK.

Warr, P., Bindl, U. K., Parker, S. K., & Inceoglu, I. (2014). Four-quadrant investigation of job-related affects and behaviours. *European Journal of Work and Organizational Psychology*, *23*(3), 342-363.

Waterman, A. S. (2005). When effort is enjoyed: Two studies of intrinsic motivation for personally salient activities. *Motivation and emotion*, *29*(3), 165-188.

Watson, D., & Clark, L. A. (1997). Extraversion and its positive emotional core. In *Handbook of personality psychology* (pp. 767-793). Academic Press.

Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structures of personality and their relevance to psychopathology. *Journal of abnormal psychology*, *103*(1), 18.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, *54*(6), 1063.

Watt, J. D., & Hargis, M. B. (2010). Boredom proneness: Its relationship with subjective underemployment, perceived organizational support, and job performance. *Journal of business and psychology*, *25*(1), 163-174.

Wauck, H., Xiao, Z., Chiu, P. T., & Fu, W. T. (2017). Untangling the Relationship Between Spatial Skills, Game Features, and Gender in a Video Game. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (pp. 125-136). ACM.

Weber, R., Tamborini, R., Westcott-Baker, A., and Kantor, B. (2009). Theorizing flow and media enjoyment as cognitive synchronization of attentional and reward networks. *Commun. Theory* 19, 397–422. doi: 10.1111/j.1468-2885.2009.01352.x

Webster, J., Trevino, L. K., & Ryan, L. (1993). The dimensionality and correlates of flow in human-computer interactions. *Computers in human behavior*, *9*(4), 411-426.

Weibel, D., & Wissmath, B. (2011). Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology*, *2011*, 6.

Weiner, E. A., & Concepcion, P. (1975). Effects of affective stimuli mode on eye-blink rate and anxiety. *Journal of clinical psychology*.

Wetzel, B., Anderson, K., Koutstaal, W., & Gini, M. (2012). If not now, where? time and space equivalency in strategy games. In *Eighth Artificial Intelligence and Interactive Digital Entertainment Conference*.

White, R. W. (1959). Motivation reconsidered: The Concept of competence. *Psychological Review*, *66*, 297–333.

Whitlock, L. A., McLaughlin, A. C., Leidheiser, W., Gandy, M., & Allaire, J. C. (2014). Know before you go: feelings of flow for older players depends on game and player characteristics. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play* (pp. 277-286). ACM.

Whitson, J., Eaket, C., Greenspan, B., Tran, M. Q., & King, N. (2008). Neo-immersion: Awareness and engagement in gameplay. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (pp. 220-223). ACM.

Whittinghill, D. M., Ziegler, B., Case, T., & Moore, B. (2015). Nasum virtualis: A simple technique for reducing simulator sickness. In *Games Developers Conference (GDC)* (Vol. 74).

Widiger, T. A. (2009). Neuroticism. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 129-146). New York, NY, US: The Guilford Press.

Widmann, A. (2006). *Firfilt EEGLAB Plugin*, Version 1.5.1. Leipzig: University of Leipzig.

Wiedemann, A. U., Gardner, B., Knoll, N., & Burkert, S. (2014). Intrinsic rewards, fruit and vegetable consumption, and habit strength: A three-wave study testing the associative-cybernetic model. *Applied Psychology: Health and Well-Being*, *6*(1), 119-134.

Williams, D. (2003). The video game lightning rod. *Information Communication & Society*, *6*(4), 523-550.

Williams, D., Yee, N., & Caplan, S. E. (2008). Who plays, how much, and why? Debunking the stereotypical gamer profile. *Journal of computer-mediated communication*, *13*(4), 993-1018.

Williams, P., & Aaker, J. L. (2002). Can mixed emotions peacefully coexist?. *Journal of Consumer Research*, *28*(4), 636-649.

Wilt, J., & Revelle, W. (2017). Extraversion. In T. A. Widiger (Ed.), *Oxford library of psychology. The Oxford handbook of the Five Factor Model* (pp. 57-81). New York, NY, US: Oxford University Press.

Witt, L. A., Burke, L. A., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of applied psychology*, *87*(1), 164.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, Massachusetts.

Wohn, D. Y., Ratan, R., & Cherchiglia, L. (2020). Gender and Genre Differences in Multiplayer Gaming Motivations. In *International Conference on Human-Computer Interaction* (pp. 233-248). Springer, Cham.

Wolf, M. J. (2001). Genre and the video game. *The medium of the video game*, 113-134.

Wong, T. T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, *48*(9), 2839-2846.

Wood, R. T., & Griffiths, M. D. (2007). Time loss whilst playing video games: is there a relationship to addictive behaviours?. *International Journal of Mental Health and Addiction*, *5*(2), 141-149.

Wood, R. T., Griffiths, M. D., & Parke, A. (2007). Experiences of time loss among videogame players: An empirical study. *Cyberpsychology & behavior*, *10*(1), 38-44.

Wu, M., & Wang, Y. (2015). A feature selection algorithm of music genre classification based on ReliefF and SFS. In *2015 IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS)* (pp. 539-544). IEEE.

Wu, T. C., Scott, D., & Yang, C. C. (2013). Advanced or addicted? Exploring the relationship of recreation specialization to flow experiences and online game addiction. *Leisure Sciences*, *35*(3), 203-217. doi:10.1080/01490400.2013.780497.

Wu, Z., Begley, C. G., Situ, P., Simpson, T., & Liu, H. (2014). The effects of mild ocular surface stimulation and concentration on spontaneous blink parameters. *Current eye research*, *39*(1), 9-20.

Wu, Z., Begley, C. G., Zhang, J., Adebisi, L. A., & Simpson, T. (2011). Does ocular surface stimulation oppose internal controls over blinking?. *Investigative Ophthalmology & Visual Science*, *52*(14), 3846-3846.

Wulf, T., Schneider, F. M., & Beckert, S. (2018). Watching Players: An Exploration of Media Enjoyment on Twitch. *Games and Culture*. https://doi.org/10.1177/1555412018788161.

Yamada, F. (1998). Frontal midline theta rhythm and eyeblinking activity during a VDT task and a video game: useful tools for psychophysiology in ergonomics. *Ergonomics* 41, 678–688. doi: 10.1080/001401398186847

Yan, S., & El-Nasr, M. S. (2006). Visual attention in 3D video games. In *Proceedings of the 2006 symposium on Eye tracking research & applications* (pp. 42-42). ACM.

Yang, S., Smith, B., & Graham, G. (2008). Healthy video gaming: Oxymoron or possibility?. *Innovate: Journal of Online Education*, *4*(4).

Yannakakis, G. N., Spronck, P., Loiacono, D., & Andre, E. (2013). Player modeling. *Dagstuhl Follow-Ups*, 6.

Yates, S. J., & Littleton, K. (1999). Understanding computer game cultures: A situated approach. *Information, Communication and Society*, *2*(4), 566–583.

Yazdani, A., Ebrahimi, T., & Hoffmann, U. (2009). Classification of EEG signals using Dempster Shafer theory and a k-nearest neighbor classifier. In *2009 4th International IEEE/EMBS Conference on Neural Engineering* (pp. 327-330). IEEE.

Ye, X., Ning, H., Backlund, P., & Ding, J. (2020). Flow Experience Detection and Analysis for Game Users by Wearable Devices-based Physiological Responses Capture. *IEEE Internet of Things Journal*.

Yee, N. (2001). The virtual Skinner Box. Available at http://www.nickyee.com/eqt/skinner.html.

Yee, N. (2006). The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and virtual environments*, *15*(3), 309-329.

Yelamanchili, T. (2018). *Neural correlates of flow, boredom, and anxiety in gaming: An electroencephalogram study* (Unpublished Master's thesis). Missouri University of Science and Technology, United States.

Yeo, G. B., & Neal, A. (2004). A multilevel analysis of effort, practice, and performance: effects; of ability, conscientiousness, and goal orientation. *Journal of applied psychology*, *89*(2), 231.

Yildirim, I. G. (2016). Time pressure as video game design element and basic need satisfaction. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2005-2011). ACM.

Yin, Z., Zhang, J., & Wang, R. (2016). Neurophysiological feature-based detection of mental workload by ensemble support vector machines. In *Advances in Cognitive Neurodynamics (V)* (pp. 469-475). Springer, Singapore.

Yoo, S., & Kay, J. (2016). VRun: running-in-place virtual reality exergame. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (pp. 562-566).

Yoshida, K., Sawamura, D., Inagaki, Y., Ogawa, K., Ikoma, K., and Sakai, S. (2014). Brain activity during the flow experience: a functional near-infrared spectroscopy study. *Neurosci. Lett.* 573, 30–34. doi: 10.1016/j.neulet.2014.05.011

Young, C. J. (2018). *Game changers: Everyday gamemakers and the development of the video game industry* (Doctoral dissertation). University of Toronto, Toronto, Ontario, Canada.

Youssef, A., & Cossell, S. (2009). Thoughts on adjusting perceived difficulty in games. In *Proceedings of the Sixth Australasian Conference on Interactive Entertainment* (p. 14). ACM.

Yuan, B., Folmer, E., & Harris, F. C. (2011). Game accessibility: a survey. *Universal Access in the Information Society*, *10*(1), 81-100.

Yun, K., Doh, S., Carrus, E., Wu, D.-A., and Shimojo, S. (2017). Being in the zone: flow state and the underlying neural dynamics in video game playing. *arXiv* [Preprint]. arXiv:1711.06967.

Xiao, Y., & Yu, J. (2012). Partitive clustering (K-means family). *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(3), 209-225.

Xiong, T., & Cherkassky, V. (2005). A combined SVM and LDA approach for classification. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* (Vol. 3, pp. 1455-1459). IEEE.

Zammitto, V. (2008). Visualization Techniques in Video Games. *Electronic Information, the Visual Arts and Beyond*, 267–276.

Zeigler-Hill, V., & Monica, S. (2015). The HEXACO model of personality and video game preferences. *Entertainment Computing*, *11*, 21-26.

Zhang, J., & Fu, X. (2015). The influence of background music of video games on immersion. *Journal of Psychology & Psychotherapy*, *5*(4), 1.

Zhang, S., Li, X., Zong, M., Zhu, X., & Wang, R. (2017). Efficient knn classification with different numbers of nearest neighbors. *IEEE transactions on neural networks and learning systems*, *29*(5), 1774-1785.

Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O. N. M., & Atkins, M. S. (2012). Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surgical endoscopy*, *26*(10), 2746-2750.

Zhou, X., Ji, X., & Chai, X. (2017). A Wrist-type Multi-physiological Signal Acquisition Device. In *2016 International Forum on Mechanical, Control and Automation (IFMCA 2016)* (pp. 150-153). Atlantis Press.

Zimmerman, B. J. (1995). Self-efficacy and educational development. *Self-efficacy in changing societies*, *1*, 202-231.

Zohaib, M. (2018). Dynamic difficulty adjustment (DDA) in computer games: A review. *Advances in Human-Computer Interaction*, *2018*.

9 Appendix

9.1 Informed consent – Participation sheet

ID	

Sony Interactive Entertainment

Centre for Digital Entertainment

Bournemouth University

I would like to invite you to take part in a research study. Before you make your decision, please, read carefully why this study is being done and what you will be asked to do. Bear in mind that agreeing to participate does not refrain you from withdrawing at any point.

This study is a collaboration of Center for Digital Entertainment (Bournemouth and Bath Universities) and Sony Interactive Entertainment and funded by The Engineering and Physical Sciences Research Council (EPSRC).

We wish to acquire a better understanding of how the body reacts during virtual reality video game play. You will be asked to play a video game in virtual reality (VR) and to give your opinion on an array of questionnaires. Your heart rate, eye movement and brain signals will be recorded during the experiment.

Please, consider the following:

- Although the focus of the study is to evaluate your experience with the game, you may choose not to respond to questions if you consider them intrusive.
- You may withdraw from the experiment at any time without providing any reason whatsoever, and there will be no consequences.
- Your personal details and experimental data will be destroyed after the completion of the study, and no third party will be given access to them.
- We respect the privacy of our participants, and we process everything with confidentiality.

- You understand that this study is not diagnostic and does not, in any way, aim to profile individuals based on their answers or performance in the game.
- In the case of discomfort, you can remove the headset and decide on the continuation of your participation.
- The results of this study may be published in a journal or used for teaching purposes. The results may also be presented at scientific meetings or in talks at academic institutions. Results will always be presented in such a way that data from individual volunteers cannot be identified.

If you agree with the above terms, please, fill in the boxes in the next page.

Age	
Gender	
Signature	
___________ / ___________ / 20	

Do you smoke?

Yes	No

If yes, how many years?

Do you drink coffee?

Yes	No

If yes, how many did you have today (in cups/glasses)?

Do you exercise?

Yes	No

If yes, how many days a week?

9.2 Player expertise questionnaire

This questionnaire will help us determine your expertise with video games. Please, read each question carefully and answer as accurately as possible.

1. Have you ever played video games (any console or platform)?

Yes	No

2. If yes, how long have you been playing?

Number	Days / Months / Years (Please, circle the corresponding option)

3. Have you abstained from playing video games for extensive periods? If yes, please mention an approximate, cumulative number of days, months or years. If not, write "0" in the "Number" field.

Number	Days / Months / Years (Please, circle or underline the corresponding option)

4. At what age did you start playing video games?

Age

5. How many hours per day do you dedicate in video game play (please, indicate an average number)?

Hours per day

6. Do you usually play games longer than you planned?

Never	Rarely	Sometimes	Frequently	Always

7. How often during the day, do you think about playing a video game?

Never	Rarely	Sometimes	Frequently	Always

8. Do you resort to video game play when you do not feel well?

Never	Rarely	Sometimes	Frequently	Always

9. How likely is it that you skip responsibilities or commitments in order to play video games?

Never	Unlikely	Sometimes	Likely	Always

10. When you play video games, do you lose track of time?

Never	Rarely	Sometimes	Frequently	Always

11. Does the time *feel* like it goes by more slowly than usual or more rapidly? (<u>Please answer this question only if you did not choose "Never" in question number 10</u>)

More slowly	More rapidly

12. Do you feel enjoyment when you play video games?

Never	Rarely	Sometimes	Frequently	Always

13. Have you ever intentionally avoided playing video games?

Yes	No

9.3 Time passing estimation (post-game)

1. How much time do you think passed since the beginning of the game?

Approximate number of minutes

2. How likely would you play this game again in your free time? Indicate your choice with a checkmark (✓) to the corresponding cell.

I would not replay this game	Not very likely	I would not care if I played it again or not	Likely	I would definitely replay the game