UNIVERSITY OF LIVERPOOL

Enjoyment in VR games: Factors, Challenges, and Simulator Sickness Mitigation Techniques

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

By

Diego Vilela Monteiro

January 2021

PGR Declaration of Academic Honesty

NAME (Print)	**DIEGO VILELA MONTEIRO**
STUDENT NUMBER	**201323875**
SCHOOL/INSTITUTE	**Department of Computing**
TITLE OF WORK	**Enjoyment in VR games: Factors, Challenges, and Simulator Sickness Mitigation Techniques**

This form should be completed by the student and appended to any piece of work that is submitted for examination. Submission by the student of the form by electronic means constitutes their confirmation of the terms of the declaration.

Students should familiarise themselves with Appendix 4 of the PGR Code of Practice: PGR Policy on Plagiarism and Dishonest Use of Data, which provides the definitions of academic malpractice and the policies and procedures that apply to the investigation of alleged incidents.

Students found to have committed academic malpractice will receive penalties in accordance with the Policy, which in the most severe cases might include termination of studies.

STUDENT DECLARATION

I confirm that:

- I have read and understood the University's PGR Policy on Plagiarism and Dishonest Use of Data.
- I have acted honestly, ethically and professionally in conduct leading to assessment for the programme of study.
- I have not copied material from another source nor committed plagiarism nor fabricated, falsified or embellished data when completing the attached material.
- I have not copied material from another source, nor colluded with any other student in the preparation and production of this material.
- If an allegation of suspected academic malpractice is made, I give permission to the University to use source-matching software to ensure that the submitted material is all my own work.

SIGNATURE 

DATE January 6, 2021.

Abstract

Although Virtual Reality (VR) has been developed for a while, the last decade has seen a surge in its popularity with the advent of commercial VR Head-Mounted Displays (HMDs), making the technology more accessible. One field that significantly benefits from VR is the entertainment industry, for example, games. Games can be challenging to design as they involve several components that are found in other types of applications as well, such as presentation, navigation, interaction with virtual agents, and in-game measurements. Despite recent advances, the optimal configurations for game applications in VR are still widely unexplored. In this thesis, we propose to fill this gap by a series of studies that analyse different components involved in making VR applications more enjoyable.

We propose studying three characteristics that are heavily influential in game enjoyment (1) the aesthetical realism and emotions of virtual agents; (2) viewing perspective (First-Person Perspective and Third-Person Perspective), its influence on subjective feelings and how to measure those feelings; and (3) how to reduce or eliminate VR Sickness without affecting the experience (or affecting it positively).

Our results showed that Virtual Agents' facial expressions are one of the most important aspects to be considered. On the second topic, we have observed that viewing perspective is influential on VR Sickness; however, other subjective feelings were challenging to measure in this context. On the last topic, we analysed existing tendencies in Simulator Sickness mitigation techniques that do not affect in-game mechanics and present a novel solution that has a good trade-off between mitigating VR Sickness and maintaining or enhancing immersion and performance. Finally, we propose some guidelines based on our results.

Acknowledgements

First, I would like to thank my supervisor Dr. Hai-Ning Liang for his support, encouragement, and remarkable patience. My time as a Ph.D. student required a lot of dedication but, through his guidance, I was able to achieve far more than I had initially imagined and still have a positive outlook on life. I would need a completely new thesis just to be able to thank him properly.

I also sincerely extend my gratitude to Dr. Andrew Abel for key insights in my development and the other member of my supervisory and advisory members (i.e., Mr. Phil Jimmieson, Dr. Airong Wang, Dr. Nilufar Baghaei, Dr. Jieming Ma, and Dr. Floriana Grasso) for their feedback and insightful comments on my work. Kok Hoe Wong for his guidance during my TA work. My examiners (Dr. Cheng-Hung Lo and Dr. Yiyu Cai) for the insightful comments that helped improve my thesis.

I would also like to thank everybody at the X-CHI lab for being great friends, colleagues, and collaborators. Special thanks to my Ph.D. colleague Wenge Xu: thank you for everything in the past three years. I hope our paths keep crossing in the future. See you in the UK. Thanks also to Xian Wang, who helped me a lot.

I cannot possibly thank my family enough. My mother Rita and father Fernando gave me so much support and love throughout my life, always making sure I did not give up. They have always stimulated my curiosity and sense of wonder. Without them, I would not be here. Renata for supporting my decision to study abroad. My grandparents and aunts and uncles, godfather, and godmother.

My appreciation goes to all my friends who helped me through my academic and life journey, Juliana “Perspicaz”, Nielsen, Nicolas, Luis Fernando, Enari, Ricardo, Iohans, Ina, Irwyn, Prateek, Mo, Samer, Rafael, Luiz, Fernando, Gabriel.

Honestly, my journey is such a collective effort that I should call it “our” journey. So, thank you, everyone!

Publications and Contributions

Materials from this dissertation have been previously published in the journals and conferences listed below. The chapter corresponding to each paper is noted in the parenthesis.

1. Diego Monteiro, et al., "Evaluating the Effects of a Cartoon-Like Character with Emotions on Users' Behaviour within Virtual Reality Environments," in Proceedings of the First IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR 2018), Taichung, Taiwan, 2018, pp. 229-236, DOI: 10.1109/AIVR.2018.00053. (Chapter 3)

2. Diego Monteiro, et al., "Evaluating the Need and Effect of an Audience in a Virtual Reality Presentation Training Tool," in Proceedings of the 33rd International Conference on Computer Animation and Social Agents (CASA 2020), Bournemouth, UK, 2020. (Chapter 4)

3. Diego Monteiro, et al., "Comparing the effects of a Familiar Audience with an Unfamiliar One in VR Using a Presentation Training Tool," in Frontiers in Virtual Reality, section Virtual Reality and Human Behaviour, under review. (Chapter 5)

4. Diego Monteiro, et al., "Evaluating enjoyment, presence, and emulator sickness in VR games based on first- and third- person viewing perspectives," in Computer Animation and Virtual Worlds, vol. 29, no. 3–4, p. e1830, 2018, DOI: 10.1002/cav.1830. (Chapter 6)

5. Diego Monteiro, et al., " Correlating gamers' Brainwaves to their subjective Feelings in Virtual Reality games under Different Viewing perspectives," in Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications in Industry (VRCAI 2018), Tokyo, Japan, 2018, pp. 1-4, DOI: 10.1145/3284398.3284430. (Chapter 7)

6. Diego Monteiro, et al., “Comparing Event-Related Arousal-Valence and Focus among Different Viewing perspectives in VR gaming," in Proceedings of the 9th International Conference on Brain-Inspired Cognitive System (BICS 2018), Xian, China, 2018, Lecture Notes in Computer Science, vol 10989. Springer, DOI: 10.1007/978-3-030-00563-4_75. (Chapter 7)

7. Diego Monteiro, et al. , “Evaluating Engagement of Virtual Reality games based on First- and Third-Person Viewing perspectives using EEG and subjective Metrics," in Proceedings of The First IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR 2018), Taichung, Taiwan, 2018, pp. 53-60, DOI: 10.1109/AIVR.2018.00015. (Chapter 7)

8. Diego Monteiro, et al., "A Review of visual Techniques for Simulator Sickness Mitigation in Virtual Reality Head-Mounted-Displays," in Springer Virtual Reality under-review. (Chapter 8)

9. Diego Monteiro, et al., "An In-depth Exploration of the Effect of 2D/3D views and controller types on first-person shooter games in Virtual Reality," in Proceedings of 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Beijing, China, 2020. (Chapter 9)

Table of Contents

List of Abbreviations

1PP	First-Person Perspective
3PP	Third-Person Perspective
ADS	Absolute Discomfort Score
AF	Frontal/Pre-Frontal Electroencephalogram Channel
AI	Artificial Intelligence
ANOVA	Analysis of Variance
AR	Augmented Reality
BM	Big Mask
BPM	Beats Per Minute
BS	Body Score
CAVE	Cave Automatic Virtual Environment
CD	Conventional Display
CEFR	Common European Frame of Reference
CV	Consumer Version
DK	Development Kit
DS	Discomfort Score
EA	Extremely Anxious
ECG	Electrocardiogram
EDS	Ending Discomfort Score
EEG	Electroencephalogram
EMG	Electromyogram
ERP	Event-Related Potential
FFT	Fast Fourier Transform
FMS	Fast Motion Sickness
FOV	Field-of-View
FP	Frontal Electroencephalogram Channel
FPS	First-person Shooter
GEQ	Game Experience Questionnaire
HA	Hypothesis
HMD	Head-Mounted Display
HR	Heart Rate

HS	Health Score
HUD	Heads-up Display
IEQ	Immersive Experience Questionnaire
IQR	Interquartile Range
MANOVA	Multivariate analysis of variance
MLP	Multilayer Perceptron
MS	Motion Sickness
NA	Not Anxious
ND	Not Disclosed
V/E	Voice and Emotion
V/NE	Voice and No Emotion
NV/E	No Voice and Emotion
NV/NE	No Voice and No Emotion
NPC	Non-Playable Character
NU	Not Utilized
PCs	Playable Characters
PC	Personal Computer
PF	PlaneFrame
P_{FoV}	Plane Field-of-View
PGQ	post-game questionnaire
PRCS	Personal Report of Confidence as a Speaker
PSD	Power Spectrum Density
RDS	Relative Discomfort Score
RLF	Real-life familiar audience
RLU	Real-life unfamiliar audience
RM-ANOVA	Repeated Measure Analyse of Variance
RMB	Chinese Yuan
RPG	Role Playing Game
RQ	Research Question
RVR	Regular HMD VR display without Adjustments
SA	Standard Anxious
SM	Small Mask
SS	Simulator Sickness

SSQ	Simulator Sickness Questionnaire
TV	Television
VA	Virtual Audience
VE	Virtual Environment
VR	Virtual Reality
VRE	Virtual Room Empty
VRF	Virtual Reality Familiar Audience
VRP	Virtual Room Practice
VRPF	Virtual Reality PlaneFrame
VRU	Virtual Reality Unfamiliar Audience

List of Equations

List of Figures

List of Tables

Chapter 1 Introduction

Over the last decade, Virtual Reality (VR), more specifically VR head-mounted displays (HMDs), have attracted ever-increasing interest from the public. According to a survey we performed[1] using over 280 people of diverse nationalities about the use of VR, 67% of respondents had used VR in some capacity. From that 67%, at least 49% of respondents felt some level of sickness after experiencing VR. Furthermore, 66% of all the participants declared they planned to buy a VR device within the next year, and over 70% imagined it would be fun to learn how to use it.

Even though these numbers suggest that VR is becoming increasingly popular, they also reveal a problem that can cause a great deal of discomfort for first-time and continuous HMD device users: VR Sickness. Thus, it is crucial to evaluate how different factors in VR affect participants' fun (i.e., enjoyment) and how we can maintain enjoyment while reducing the incidence of motion sickness.

Therefore, in this thesis, we will study the factors involved in the feeling of presence in VR, such as (1) virtual agents and how people react to them, and what kinds of developments in the area of virtual agents enhance the interactive experience, (2) viewing perspectives' influence on virtual environments, and whether VR must always be viewed in a first-person perspective, or whether there can be advantages in using a third-person perspective, and (3) how to better minimize VR Sickness without negatively affecting the user experience. We will also propose a novel technique to mitigate the VR Sickness.

[1] Article under preparation

Section 1.1 Virtual Agents

Interaction in VR is often mediated by virtual agents (hereafter referred to as *characters*), which can come in different formats, styles, and levels of realism. Their control and behaviour can be pre-programmed, can be based on a behavioural AI, or can be human-driven. However, one of the challenges when dealing with characters is how to make them express emotion, and how to make them easier to empathize with. Furthermore, it is unclear whether realism facilitates emotional expression or if it creates a barrier.

The advancement of 3D scanning technologies allows us to create ever more realistic characters in VR [1]; greater aesthetic realism could equate to more presence and other positive feedback when talking about the self [1], [2].

Most research studying how characters can present human emotions emphasizes creating realistic representations of humans [3]–[7], making them as lifelike and realistic as possible. However, cartoonish characters can be used to emphasize emotions [8]. Cartoon characters might have advantages like being less likely to fall into the “uncanny valley” because they are more conceptual and, thus, more flexible [9].

Other behavioural factors of the virtual character might mediate greater enjoyment and credibility in certain situations [10]; thus, it is essential to understand what kind of developments should be prioritized for improving character interaction. Is it really advantageous to 3D scan someone, or are other models enough?

Section 1.2 Viewing Perspective

One factor that promotes presence and is still underexplored in immersive VR is perspective [11], [12]. In this thesis, when talking about viewing perspective, we mean first-person perspective (1PP), that is, the virtual camera is positioned where the virtual characters head is, and third-person perspective (3PP), where the virtual camera is positioned so it can see the whole body of the player from a distance [13]. Even though other kinds of views exist, such as aerial view, in which the distance effectively forces the view to become nearly 2D, they are beyond the scope of this research. Viewing perspective can alter not only the levels of presence but also potentially reduce the levels of Simulator Sickness (SS) [14]. Thus, it is worth conducting analyses to understand better what factors can be the most beneficial to enjoyment in VR [15]. Such investigation can be done through novel psychophysiological measurement techniques such as Electroencephalogram (EEG) [16], which are now commercially available and could be more easily integrated into HMDs.

Section 1.3 Simulator Sickness Mitigation Techniques

Simulator Sickness affects a significant number of people who experience VR [17], at least for first-time users [18]. It has been studied for centuries in other forms [19]; however, no definitive conclusion has been reached yet to the root cause of Simulator Sickness [20]. Nevertheless, many empirical studies have sought to mitigate it [21]–[23], with varying levels of success and different trade-offs. When it comes to VR, these trade-offs might affect enjoyment and performance [24]. Hence, we propose investigating what has already been done to mitigate Simulator Sickness (SS) in VR and analyse a new technique regarding its trade-offs. Our investigation on mitigation techniques will focus on what can be done using only software.

Section 1.4 Thesis Statement

The goal of this thesis is to understand some of the factors influencing enjoyment in immersive VR games and to design a novel technique to mitigate SS in immersive VR games catering to players' performance and experience. In specific, this research focuses on four core factors of enjoyment in Virtual Reality games:

- Interaction with virtual agents.
- Viewing perspectives in VR games.
- Mitigating Simulator Sickness without disrupting immersion.
- Mitigating Simulator Sickness without hindering player performance.

Section 1.5 Contributions

This thesis studies the core factors affecting presence and enjoyment in VR HMDs. In this context, it makes the following main contributions:

- Showing empirically that cartoonish characters only need to express their emotions verbally or facially to affect the player, expressing both simultaneously did not increase affect.
- Identifying that characters are preferred to empty environments; however, their body language can be overlooked; participants did not notice it during interactive scenes but reported missing eye contact in empty environments.
- Determining that characters that are more realistic emotionally can increase players' competitiveness.
- Proposing an addition, called "Mood-Gravity", to a state-of-the-art mood controller algorithm; the addition brings the current emotion to neutral, expressing the short-term trait of emotions.
- Showing that some people present lower levels of Simulator Sickness when using the 3PP in VR racing games.
- Proving that current psychophysiological metrics do not correlate linearly with subjective metrics. However, they are potentially useful in detecting players' preferences.
- Presenting a study on the techniques used to mitigate sickness in commercial HMDs.
- Creating a technique that reduces Simulator Sickness, but that has little effect on immersion, and in-game performance.
- Designing a set of guidelines for VR games based on the current findings.

Section 1.6 Dissertation Organization

Chapter 2: Literature Review – This chapter provides a description of HMDs (i.e., history, user experience-related characteristics with current HMDs), and summarizes concepts related to games that are used in this thesis.

Chapter 3: Evaluating the Effects of a Cartoon-like character with Emotions on Users' Behaviour within Virtual Reality Environment – This part of the dissertation empirically investigates how, in an immersive Virtual Reality environment, a cartoonish character expressing emotions in different ways through a combination of facial and vocal expression affects a player's emotional state. The virtual character in this chapter has an emotional system based on the latest research on the topic, with improvements.

Chapter 4: Evaluating the Need and Effect of an Audience in a Virtual Reality Presentation – In this chapter, we evaluate how aesthetically realistic characters are perceived and whether they are preferred to no character at all. This evaluation is performed in an environment prone to cause social fear in participants: a public presentation. We accomplished this by using 3D scanning real people as the audience and adding them to the evaluation environment. Later we ask participants to train for a presentation either in front of the said audience or in an empty environment and to then present to a real-life audience. We then compared their feedback regarding how the audience made them feel. We observed a preference for training in front of an audience, which should be explored further.

Chapter 5: Comparing the Effects of a Familiar Audience with an Unfamiliar One in a VR Presentation Simulator – In this chapter, based on the feedback from Chapter 4, we explore how an audience with 3D-scanned faces compares to an audience using generic stock models according to the participants' perception. We evaluated the participant's level of nervousness based on their heart rate and their verbal feedback. The participants often reported not minding the differences of the audience's faces in a stressful environment of a presentation; however, they noticed a lack of facial feedback.

Chapter 6: Evaluating Enjoyment, Presence, and Emulator Sickness in VR Games Based on First- and Third-Person Viewing – This chapter presents an evaluation of the subjective feelings of players after experiencing three versions of the same racing game in Virtual Reality under different perspectives. This chapter aims to identify some trends for a subsequent deeper investigation using psychophysiological metrics in Chapter 7. We observe that third-person perspective induces Simulator Sickness slightly less often than first-person perspective and is slightly less immersive.

Chapter 7: Analysing Subjective Feelings in Virtual Reality Games Based on First- and Third-Person Perspective using EEG and Subjective Metrics – In this chapter, we present a follow-up to Chapter 6, in which we use not only subjective metric to evaluate how players feel about the same racing game under different perspectives while also using psychophysiological metrics, namely EEG, to identify correlations and evaluate if the use of commercial EEGs is feasible to identify patterns during VR gaming. Our data do not reveal clear differences between first- and third-person perspectives in VR. However, we observe some patterns when analysing other parameters.

Chapter 8: An In-depth Exploration of the Effect of 2D/3D Views and Controller Types on First-person Shooter games in – After the feedback from many participants from Chapter 6 indicating Simulator Sickness to be a problem in VR, this chapter investigates through a systematic literature review existing mitigation techniques that do not require extra hardware and do not affect how existing games behave. These techniques only alter how the VR environment is presented without altering locomotion behaviours, which would be disruptive in games that require continuous movement, such as racing games. Later in this chapter, we investigate the effects the interaction between controller and 2D and 3D views have in-game performance, immersion, and Simulator Sickness in VR. Further, to conduct such an investigation, a new technique to mitigate Simulator Sickness is proposed. This method is investigated and presents a good trade-off between reducing Simulator Sickness and being immersive and allowing players to keep a good in-game performance. Beyond that, we observe that controllers can be extremely influential in causing Simulator Sickness.

Chapter 9: Discussion, Conclusion, and Future Work – This chapter first presents a summary of our investigation and partial conclusions for each aspect investigated and proposes some guidelines for the creation of new VR games depending on the goals of the game. Then, we discuss the future challenges in creating enjoyable games in VR. Finally, Chapter 9 concludes the dissertation and provides recommendations for future work based on this dissertation.

Chapter 2 Literature Review

This chapter reviews the existing literature that is essential to the investigations within this thesis, VR HMD technologies and how they are affected by their interactions. Furthermore, we present the concept of games, a brief taxonomy, and explain the effects games have and their measurements.

Section 2.1 Virtual Reality Head-Mounted Displays

Depending on how one defines VR, VR can be traced to super-wide paintings designed to give the impression that one is inside them [25]. However, such paintings are not dynamic and resemble more the current CAVE systems than modern VR HMDs. Only in the 1960s was the first computer-supported VR HMD developed by Sutherland [26]. It was a bulky system that had to be supported from the ceiling, and the quality of the graphics was akin to the time.

Until the 2010s, VR HMDs were used mainly for scientific applications; they struggled to reach consumers. There have been many reasons for the failed attempts of consumer VR HMDs like the ones produced by Nintendo or Sega [25]. Poor technical quality (e.g., resolution, field of view, comfort), graphical quality, and prohibitive cost were likely important factors [27]. In 2009, Palmer Luckey, frustrated by the limitations from the previous generations of VR HMDs, started to develop his own device. He campaigned to have it developed and raised over US$2 million in 2012 for his VR device, which became known as the Oculus Rift [28]. The Rift was the first commercially successful VR HMD and was bought for US $2 billion by Facebook [25], [29]. Since then, there has been an increase in the amount of research and number of applications for VR, including games [29]. A 2016 report from Goldman Sachs predicted that the VR market could reach US $80 billion in 2025 [30]. This was before the Covid-19 pandemic, which has increased the sales of VR devices [31].

Because of how novel this technology is, many aspects of interaction within the VR are still being investigated [29]; it is necessary to investigate how to move within the virtual world [32], control one's avatar [33], and best present the virtual world itself [15]. In this section, we will present research that addresses some of these issues.

Section 2.1.1 Locomotion

Moving within VR is a fundamental question of its interaction. Walking within the virtual world often cannot be translated one-to-one with walking in the real world due to various limitations [34]. For one, the virtual world can be made infinite, whereas the real world is, in practice, limited in its dimensions. Moreover, objects in the real world can interfere with the movement and serve as barriers. Several techniques propose solutions to these limitations [34].

When using a gamepad control, users can move indefinitely, either continuously or discretely. The continuous movement does not cut the movement scenes and translates the joystick's input to the on-screen character actions in a fixed ratio fashion [35]. However, this kind of movement is associated with higher levels of Simulator Sickness [36]. One solution to this problem is to use discrete movement (a.k.a., teleportation). Teleport removes translational movement from VR and is associated with reducing nausea [35], [36]. However, it can lead to lower performance in games and cause greater dizziness [24], [37]. It is not ideal for social scenarios either, since one interlocutor can disappear from the other's field of view, causing confusion [38]. There are intermediary solutions [35] that involve hiding the movement from one interlocutor while presenting their movement to another [38]. Other solutions for locomotion when using a controller are changes in the scale of the movement [34] or pre-planned paths that reduce nausea-inducing movements [39].

Another solution to movement is to use walking metaphors such as walking-in-place and scaled walking, or the Human Joystick technique [34]. Walking-in-place consists of the user making movements that simulate walking without changing positions; however, it is still not a substitute to real walking as it elicits lower levels of presence and navigational ability. The Human Joystick technique involves treating the human body as a joystick and moving the character in the direction the user is tilting. This technique generates high levels of Simulator Sickness. Scaled walking performs better than the other two; however, as with normal walking, it is limited by the existing space [34].

Alternatively, walking platforms are also possible [40]–[42]; however, they require a more considerable investment in terms of money and space [40]. Their size makes them difficult to be installed in many homes. Furthermore, they require time to change movement speed, which can affect the sense of presence and cause sickness as well [42].

Considering that the alternatives also present limitations in the levels of presence and Simulator Sickness experienced by the user, in this study, we will consider locomotion done through the controllers as the default for VR applications.

Section 2.1.2 Control

How to control characters in virtual environments is still a topic of discussion. The virtual environment can be controlled, for example, by sphere-like devices [43], by standard gamepads [44], by keyboards, or by the default VR hybrid controllers that track the user's position and serve as a gamepad [44].

Furthermore, controlling VR avatars goes beyond locomotion; VR avatars move their arms and heads independently from the body [42]. Thus avatar control through body tracking can be an important tool in VR [1]. The way VR is controlled can increase enjoyment and make simulations more realistic [1].

Studies on the use of controllers revealed that playing a game using a WiiMote[2] activates different brain areas than when using the standard controllers [45].

[2] https://www.nintendo.co.uk/Wii/Wii-94559.html

Furthermore, the choice of the controller can affect performance [46]. Thus, the type of controller used is an important topic to be investigated regarding enjoyment in games.

Section 2.1.3 View and Perspective

3D First Person Perspective (1PP) is a view that lends itself naturally to VR games [42]. 1PP is when the camera is positioned where the character's eyes are [13]. In the case of VR games, 1PP is also when the camera is positioned where the eyes of the player are. However, a 3PP in which the camera is positioned right-behind the player [13] can have advantages in navigation tasks and can reduce Simulator Sickness under certain circumstances [14]. However, not much work has been done on the effects it has on enjoyment in VR games. Moreover, 2D views like that exhibited by a regular screen are also not well-explored in VR. 2D and 3D views are processed differently in the brain [47], and their differential effects on feelings of immersion in-game environments are yet to be investigated.

Section 2.2 Games

Game playing is comomly-observed phenomenon across the animal kingdom. Many young animals engage in playful behaviour [48]. However, humans retain this behaviour into adulthood. Games take a variety of forms and are present across widely different cultures. Hence, defining *games* is a challenging task.

One possible definition of games is “a closed formal system that subjectively represents a subset of reality”. That is, a game is an activity simulating some aspects of reality that has a set of rules that do not need external knowledge to be interpreted [49]. However, this definition misses one crucial component of games: enjoyment [50]. Thus, in this work, we will work with a mixed definition of games: subjective activities with a subset of rules of reality that generates enjoyment to the player. Moreover, in this thesis, we are interested in a subset of games: video games [49]. Video games (hereafter simply games) are in themselves divided into different genres

[13], [49]. In this section, we will (1) present an overview of some game genres which will be referenced in this thesis and (2) present and define some of the feelings generated in-game and how to evaluate them.

Section 2.2.1 Game Genres

According to Scott Rogers [13], there are over 50 types of game genres and subgenres. In this work, we will focus on a few that are big markets in VR [51]: (1) First-person shooters (FPS), (2) Horror, (3) Casinos, (4) Puzzle, (5) Simulation, (6) Sandbox, and (7) Race games.

First-person shooters are games in which the goal is to fire some sort of projectiles (e.g., bullets and lasers) at enemies. They are fast-paced and are seen from the player's perspective. This style of camera can be considered personal. Some examples of this kind of game are *Quake* and *Medal of Honor*.

The horror genre has great success in VR with games, such as *Resident Evil 7*. This style can be especially scary in VR, in which the player is not an observer but is instead the victim. These games often constitute players attempting to survive a horror or survival scenario with limited resources. They often include a dark atmosphere with many parts of the scenario hidden.

Casino games are games of luck, often involving cards, dices, and betting on sports. They involve betting, even if the currency used is not real money. Even though some casino games require skills to play, such as Poker, the luck and bet components are essential in this kind of game. In this regard, casino games directly contrast with puzzle games, which are based on logic, observation, and pattern completion. Puzzle games can utilize slow, methodical piece placement or use quick hand-eye coordination. Even though they might rely on luck sometimes (e.g., Tetris), luck is by no means the main component of puzzle games.

Because simulation and social games do not have clear victory goals, they can sometimes be called “toys” by some scholars. Inside these games, other activities with clearer goals can be created and explored. They often involve collaboration between players and can be used to simulate social interactions. They can also involve building locations that can be visited by other players.

Sandbox (a.k.a. “open world”) games are games with non-linear gameplay that take place in explorable worlds. Sandbox games are often exceptionally large games. Similar to some social games, sandbox games offer a variety of activities, from everyday life activities such as driving to cathartic experiences like shooting. However, unlike social games, they are often for a single player.

In race games, players simulate piloting/driving a vehicle, arranging from sports cars to spaceships. Emphasis is placed on making the experience as “real” as possible or on creating an action-based, arcade-like experience. In arcade-like racing games, cartoony drivers and environments, crazy-looking karts, and power-ups can be used to improve game performance and attack other players.

Section 2.2.2 In-game Feelings and Measurements

Games can generate a variety of feelings, like enjoyment [50], [52], [53], frustration [53], [54], and flow [53]. Flow is characterized by a feeling of being lost in the game when the challenge and skill level are in perfect balance. Games that are too easy can become boring, offering no challenge, whereas games that are too hard become frustrating [53], [55]. Matching the game’s difficulty to the players' skill level is a demanding task [56], [57]. Thus, determining how the player feels in each part of the game is extremely valuable so that the game can be adapted either through some dynamic difficulty adaptive techniques [58] or through a redesign.

In addition to enjoyment, frustration, and flow, there are other types of feelings that can be experienced by game players. For example, the aforementioned Simulator Sickness is the feeling of discomfort that users have after experiencing the virtual world [59]. This can be improved and sometimes prevented by removing movement from VR [35] or mitigated by changing how the virtual world is presented [59], for example, by adding a crosshair [60] or using ramps instead of stairs in the world design [61].

An additional feeling that is of great interest in the development of VR games is presence. The feeling of presence is the feeling of "being there" [62], [63], and is also critical in VR games. This feeling can be affected by different factors, such as embodiment, when you believe the virtual body is your own, and realism, in which the virtual world resembles the real one [63].

However, measuring in-game experience is a challenging task. It can be done through the use of questionnaires such as the Game Experience Questionnaire [64], or inferred by using psychophysiological signals (e.g., heart rate (HR) [65], brainwaves [66], skin conductance [53], [67]). Questionnaires can be used to identify several subjective factors involved in the gameplay. Nevertheless, to avoid interfering with the game-play experience, they can only be applied after the experience is over, making it harder to identify the specific elements that evoked each subjective feeling.

Psychophysiological signals can identify changes in the players' emotional condition in real-time; however, they are not as easily interpretable as questionnaires [65]. For example, even though it generally signals an increase in excitement, an increase in HR does not reveal if this increase represents a "positive" or "negative" feeling [65]. Thus, psychophysiological signals require special conversations and some complementary methods to be useful.

For example, in the work of Abujelala et al. [66], we can see an example of an EEG device used to measure game experience. They used a portable EEG measuring device named MUSE. They were able to infer engagement through a formula they devised. Similar methods have been implemented by Shin et al. [68]. In the work of Tauscher et al. [67], skin conductance was used as one of the factors to detect Simulator Sickness.

Thus, we conclude that investigating in-game experience is an open field to explore, especially through the use of EEG, which can detect engagement and possibly immersion. Thus, we judge it essential to investigate if we can create guidelines to improve the VR gaming experience using EEG.

Chapter 3 Evaluating the Effects of a Cartoon-like character with Emotions on Users' Behaviour within Virtual Reality Environment

Section 3.1 Introduction

Virtual Reality is a medium that thrives on the feeling of being there, being present [69]. It is built on the feeling that the virtual environment is tangible and real, even when our conscious minds know it is not [70]. Much research has investigated what brings a greater sense of presence to virtual environments [12], [69]–[73]. Non-playable characters (NPCs) are virtual entities that are often meant to act like humans and are thought to increase the feeling of presence and realism within virtual environments [3].

Humans are social beings [74] and, as such, we empathize with other people who share the same feelings as us[75]. It is natural for a person to sob and cry after watching a sad movie or to feel excited by watching an adventurous character do extraordinary deeds [76]. In addition, humans also tend to anthropomorphize non-humans (i.e., projecting humans' feelings and emotions onto animals and inanimate objects) [77]. It is then also natural for humans to display an emotional response towards virtual characters and virtual agents that can express believable emotions [78].

Most research studying how virtual characters can present human emotions emphasize creating realistic representation of humans [3]–[7], making them as life-like and realistic as possible, sometimes even risking falling into the uncanny valley [9]— an idea from aesthetics which holds that when features look and move almost, but not exactly, like natural beings, it causes a negative response among some observers. However, just like not all virtual characters that behave in human-like ways bring positive feelings, not all virtual environments require lifelike, very realistic representations. Games are environments that do not have to be exactly like the real world.

Games are tools to entertain ourselves and for enjoying our leisure time [50], [79]–[81] and sometimes for teaching and training purposes. They abstract information and concepts so we can practice skills that are sometimes important in our lives. For instance, games such as hide-and-seek and catch train hunting and hiding skills in a safe environment. Games such as chess or *Civilization*[3] can also be used to train more advanced strategic thinking in imaginary or virtual environments [50], [82], [83]. Because of their abstract characteristics, games do not need to present characters that are exact copies of what they are supposed to represent [84]. In some cases, simpler characters might even be more appropriate and appreciated by gamers [85].

Intuition, supported by past research, suggests that virtual characters would emotionally affect players more than the same character emotionless. However, it is important to empirically confirm such intuition. The goal of this research is to examine if basic interactions with a cartoon-like character that is programmed to express simple emotions imbue players with greater feelings than interacting with the same character as it does not express any type of emotion.

Section 3.2 Theoretical Foundations

Section 3.2.1 Emotion, Mood and Personality Models

Some behavioural models divide human emotional behaviour into three parts: (1) Personality, (2) Mood, and (3) Emotion [5]. These models are often referred to as hierarchical models of affect [86].

Personality is a set of stable mental characteristics (long-term affect) [86] (see Figure 3-1). One of the most-used personality models is called Big Five model of personality (see Table 3-1). This model divides human personality into five factors: (1) Openness to

[3] civilization.com

Experience, (2) Conscientiousness, (3) Extraversion, (4) Agreeableness, and (5) Neuroticism. According to this model, an individual's personality is characterized by how much he or she scores in each of these factors [87].

Whereas personality is a long-term affect, emotions are characterized as short-term affect because they usually decay fast and people do not focus on them for long [86]. Finally, moods are medium-term affect, and are considered emotional states that connect personality and emotions [5].

Table 3-1 Five Personality Dimensions (adapted from [5])

Factor	Description	Related Adjectives
Extraversion	Preference for and behaviour in social situations	Talkative, energetic, social
Agreeableness	Interactions with others	Trusting, friendly, cooperative
Conscientiousness	Organized, persistent in achieving goals, methodical	Methodical, well organized, dutiful
Neuroticism	Tendency to experience negative thoughts	Insecure, emotionally distressed
Openness	Open mindedness, interest in culture	Imaginative, creative

An especially influential model to describe emotions is by sorting emotions according to the dimensions of arousal and valence [88]. This model has been used in large bodies of works that deal with some form of emotional representation. It is used sometimes to understand the emotions elicited during a process [71], [89] or the context needed to recreate them [3], [7]. This model places emotions in a Cartesian system in which one axis is labelled Arousal and the other Valence; emotions are located on the plane (see Figure 3-2). Arousal refers to how exciting an emotion is, whereas Valence represents how "positive" the emotion is [71].

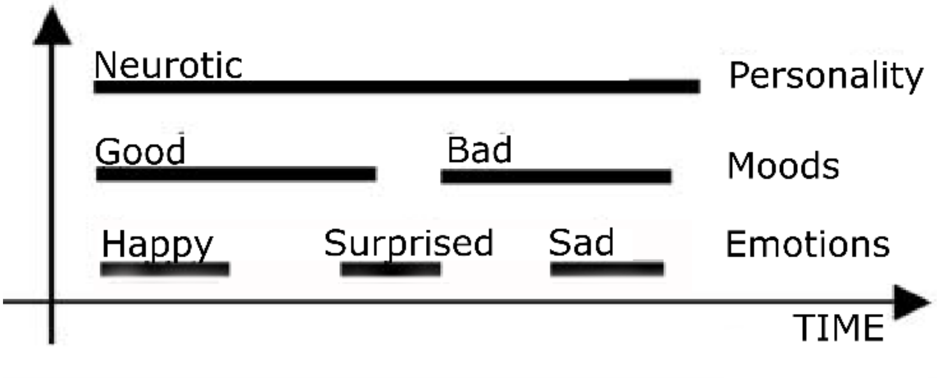

Figure 3-1 Hierarchical model of Affect (Image adapted from [5])

Section 3.2.2 Virtual Characters and Emotions

Virtual characters are present in almost all virtual environments. They can be Playable characters (PCs) and non-playable characters (NPCs) in games [84], [90]. They can be found as chatbots in healthcare and education applications [91], and they are sometimes answering calls in customer service [92]. Due to their importance and ubiquitous use, researchers for nearly three decades have been studying the use of virtual characters and their effect on users of virtual environments [93].

It is human nature to anthropomorphize inanimate things or non-human life beings [94] and, as such, it is only natural that virtual characters become channels for projecting human-like behaviours so that people who interact with them can also relate to them in a natural way [84], [85]. Therefore, users expect that computers (and robots) demonstrate similar behaviour to humans [95]. By giving them human-like attributes and behaviours, like a simple gaze, it is enough to make people feel they are being understood and that virtual beings understand what they are looking at [96].

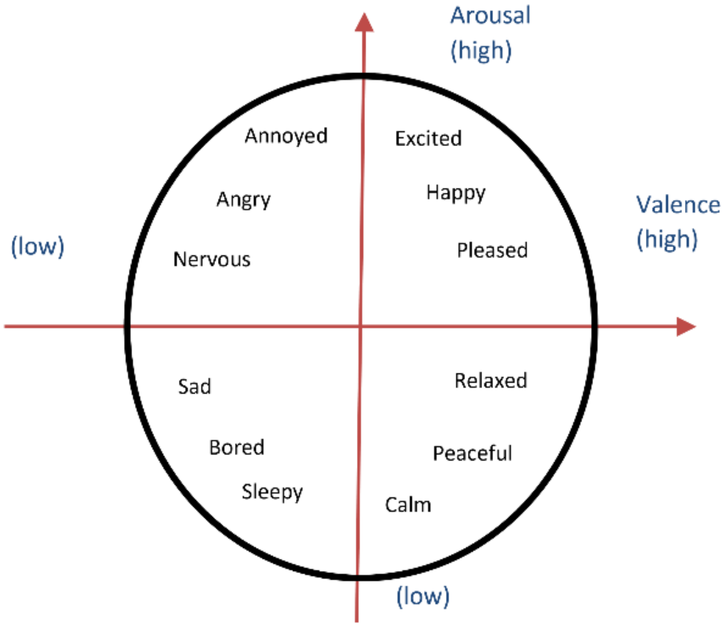

Figure 3-2 Emotional states and their positions in the arousal/valence plane (Image adapted from [89])

We can see in Figure 3-3 Funge's model that shows the different levels of requirements to create a high-level virtual character (i.e., cognitive character) [97]. Following this model, early works on virtual characters focused on allowing characters to move and speak via animations [93]. More recently, the focus is to give characters personality so they can somehow become autonomous [4].

As technology advances and we can move higher in Funge's hierarchy, it is interesting to see that, even though people expect virtual beings to exhibit human-like characteristics [98], virtual characters that present too many human traits are often classified as unsettling [99] and, at other times, are even deemed hard to relate to. Based on this review, a cartoonish character might oftentimes be more appropriate for a virtual environment [9].

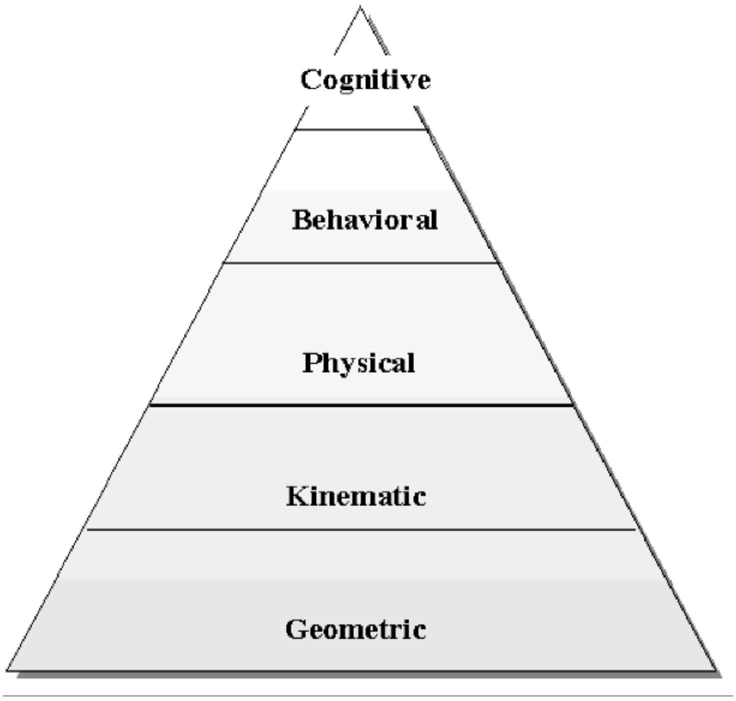

Figure 3-3 Funge's CG Modeling Hierarchy (Image adapted from [4])

Section 3.3 Related work

Section 3.3.1 Building Emotional Virtual Characters

One of the first works that have tried to create social behaviour and emotions for a computer agent is based on goals and a set of general personality traits, such as "curiosity" and "aggressiveness". Its main changes in behaviour are based on achieving or failing its goals [100]. This early system actually required the users to suspend their beliefs in order to "immerse in the world". A few years later, Rousseau and colleagues [101] have proposed, given time constrains, "a model providing a sufficiently rich structure based on the convention architecture of an intelligent agent, and available for computer actors". In their work, they use the personality model OCC [102]. However, this model is somewhat complex, and the authors state that, by the time the paper was published, the model remained still incomplete.

In [4], the authors use the Big-Five Model to develop their virtual character. They do not wish only to control the character's verbal response but also his movements. That is, they want to make the character's movements correspond to the personality attributed to it.

Even though quite interesting, this work did not inform whether players felt better when interacting with a virtual character whose movements corresponded to the personality given to it. Another work that uses the Big-Five model is [5]; it makes use of the hierarchical model of affect (see Figure 3-1) to develop a chatbot with a face associated to it. In later works, the same authors also used the OCC model [6], [103].

Some research has later proposed a layered model of affect (ALMA)). It uses the Big Five personality traits as parameters of a function that maps the emotions to a 3D version of the Arousal-Valence model. In this version, the model is named Pleasure-Arousal-Dominance. It also uses the OCC model [104]. However, in this research, no information is given about how well this model performs.

Using OCC and Pleasure-Arousal-Dominance models in [90], the authors propose an engine for emotions in games that implements similar ideas to the ones proposed in previous works in a black-box for use in development. In short, our literature review points to a trend in the evolution of the development of emotional systems. We believe that the use of Arousal-Valence and the Big-Five model of personality in character development appears to be promising, with recent studies shifting focus to it.

Section 3.3.2 Impact of Emotional Virtual Characters in Humans

The ALMA model proposed in [104] was later tested and the results are reported in [86]. The authors verify that the emotions of the in-game character were believable to users, but the moods were not very believable. They have also observed a difference in how believable the state of sadness was, depending on the age of the participants. What they have observed is that there is no "right way" to recreate an emotion.

One well-known game which tries to represent personality traits is The Sims, the life simulation game series. The second instalment has been used to test which kind of personality and appearance will get players more attached to the virtual characters. The findings suggest that the characters' appearance and perceived personality do matter [105].

Non-playable characters are especially common in role-playing games (RPGs). It has been observed in a virtual market that players become more engaged with the game if the NPCs present some sort of personality, even if the NPCs are not modelled to look like humans (but are still humanoid) [10]. It has been observed in a market setup experiment that people within a VR environment would interact more with human-like characters if those characters expressed some sort of emotional response [3].

A more recent study has demonstrated that realistic virtual avatars can have their emotions understood by both male and female participants in a VR environment. These avatars would tell a story, and the participants would later explain how involved they were with the avatars [76].

The results of the above studies are interesting because they suggest that a non-life-like (i.e., cartoonish) human character that presents some emotional responses might be more interesting for players in a VR game (although the interest and believability might vary according to different factors pertinent to the player).

Section 3.4 Environment and character Design Rationale

In this research, we wanted to explore whether a cartoon-like character in a VR environment would elicit different reactions from users if it could show simple, common emotional states or feelings based on users' input. Similar to [76], where the author used voice actors to help the user empathize with the character, we also wanted to see if sound would be an influencing factor, or if just being able to express simple emotions visually would be enough to elicit desired reactions from users.

We decided to use an anime-stylized avatar because we could modify its eyes to make them bigger so that they could be seen easily, which we believed would help in the identification of the emotions (See Figure 3-4). No body emotion was added. Since the virtual character is a female, hereafter we will refer to the character using the pronouns "she" and "her".

For her voice, we had a few sentences recorded to represent each feeling, following the directions of a performing arts professional so that the emotions would be credible. These sentences were written to be easily understandable and translatable. They were "I'm so happy", "I feel sad", "Hey! Stop it", "I'm bored", and "This is good isn't it?".

Figure 3-4 The female virtual character showing feelings of Anger (Left), sadness (Middle) and Happiness (Right)

The algorithms for showing emotional expressions were mainly based on the work of [3]. In this algorithm, the different emotions in the arousal plane are reached through interactions (events). Every interaction modifies the current position of the emotion and is mediated through the Big-Five personality values. In addition, we added a parameter that we called "Mood-Gravity", in which gravity would change the mood to a neutral state after some time to simulate the short-term trait of emotions. This "Mood-Gravity" parameter was considered by the system as a normal event from [3] that happens during every update.

In our VR environment, we devised a simple game of Rock-Paper-Scissors to be played against the virtual character. In this game, cheating was allowed, and victories would cheer up the character whereas defeats would sadden her (see Figure 3-5). Cheating was permitted to facilitate emotions to be achieved and to see if people who cheated would feel guilty more often if the character could express her feelings. The groups were not instructed differently to ensure comparability among results.

To be able to bring forth the other emotional states, the player could use balloons to tease the character by throwing them at the character until she got angry. We also placed a guitar in the environment and the users could play it to help the character go into a relaxed state.

Figure 3-5 Beginning of a Rock-Paper-Scissor match in the game and the virtual character is a neutral state (neither happy nor sad)

Section 3.5 Study Design

To evaluate the possible differences in feelings and perceptions of the participants between the Emotion and Non-Emotion conditions, and to also determine whether voice would have a significant effect, we devised the following study. Volunteers would interact with the game under the following conditions: Voice and Emotion (V/E), Voice and No Emotion (V/NE), No Voice and Emotion (NV/E) and, finally, a baseline No Voice and No Emotion (NV/NE) condition.

Participants wore an Oculus RIFT CV1 HMD while they played the four versions of the game. After each version, participants also complete a questionnaire. This questionnaire was constructed using two different subjective questionnaires, one measuring affection and realism (using values from 0 to 100) [3] and the other measuring Simulator Sickness [106]. It also contained a few other questions that we constructed (see Table 3-2 and Table 3-3).

Table 3-2 Questions in between versions

Q1.1	Did you cheat on the game?
Q1.2	If you cheated, did you feel guilty?
Q1.3	How did you feel when she was[happy, angry, sad, bored]

Table 3-3 Questions at the end of the study

Q2.1	Order the versions by preference
Q2.2	Do you think her expressions reflect the right feelings?
Q2.3	Which version made you feel the most competitive?
Q2.4	Which of her emotions made you feel the most competitive?

Section 3.5.1 Volunteers

We recruited, through social media, volunteers from a local university in China that followed a UK-based curriculum. They were from different study programs. We had a total of 12 volunteers (three females) with an average age of 21.67±2.1 years that ranged from 20 to 27 years of age. All volunteers were in good mental and physical health (that is, they were not aware that they had any neuropsychiatric disorders) and had normal or corrected-to-normal vision. All volunteers were briefed about the procedures and decided to participate in the study on a voluntary basis without any compensation.

Based on the data from the pre-experiment survey, 100% of the participants had experience with computer games, and 91.67% of them already had experience playing other games in VR.

Section 3.5.2 Apparatus

As stated earlier, we used an Oculus Rift CV1 as our HMD, as it was one of the most popular VR devices available. It was connected to a laptop with 32.0GB RAM, an Intel(R) Core (TM) i9-8950HK CPU @ 2.90GHz, and an NVIDIA GeForce GTX 1080Ti dedicated GPU.

Section 3.5.3 Procedure

On arrival, each participant was assigned a specific order of conditions that they would experience when they would play the game. The order had been formed through a Latin Square design so that carry-over effects would be mitigated. The participants then filled out a questionnaire to collect demographic data, such as age, and past gaming experience information, such as how often they played games and whether they had experience with VR.

A researcher would then explain to the participants how the experiment and the VR device worked. Then, the official information sheet was given to them to read. The volunteer would then sign a consent form to confirm their participation. Afterward, the researcher would position the HMD equipment on the participants' heads, checked their comfort level, and reposition the device if necessary.

In the next step, the participants were introduced to the controller and the game mechanics. They were instructed to experiment in the virtual environment, as the goal was to try to make the character feel as many emotions as they could. They could then interact with the character by playing the rock-paper-scissors game against, playing guitar to her, or throwing balls at her.

We required the participants to take 10-minute breaks between the different conditions. Within these 10 minutes, they would fill out questionnaire. To ensure comparability between participants and greater accuracy in the questionnaires, we requested the participants to rest for the whole 10 minutes regardless of how fast they answered the questionnaires.

On average, the experiment lasted around 80 minutes for each participant. Towards the end, they were offered refreshments and were asked to stay in the lab for a few minutes to make sure they were not impaired or uncomfortable due to possible Simulator Sickness. Finally, they were informed that they could leave the study lab at any moment if they so desired.

Section 3.6 Results

After the experiment, we had a total of 48 answered questionnaires, four for each participant and each of the four representing a different condition. Due to technical issues, the answers for Q2.3 and Q2.4 for two participants were not recorded. We used the subjective values given by the players. We then calculated the average of each of these indices by participant and condition following the instructions given in the respective references [3], [106]. We used IBM SPSS software (IBM Analytics) to calculate multiple two-way within-subjects ANOVAs. As our independent variables, we used each condition and, as our dependent variable, we used the scores of the quantitative questionnaire [3], [106]. The non-quantitative questions were used in an in-depth analysis. However, we did not have enough different responses in Q1.1 and Q1.2 to allow its use in further analysis.

According to the questionnaires, the participants overwhelmingly preferred the V/E condition, with 11 out of 12 declaring it as their favourite. The second and third choices were not as clear and were tied between V/NE and NV/E. The NV/E condition was chosen only three times as a second choice, but was chosen at least once as a first choice, whereas V/NE was chosen eight times as a second choice but was also put once as a last choice. Finally, the least-liked version was NV/NE, with 11 out of 12 participants ranking it as their least favourite. Simulator Sickness scores did not present any relevant information to report; that is, no participant reported feeling any level of Simulator Sickness in any of the versions. Thus, the simulator sickness data was not used in further analyses.

The mean values for Affect (out of 100) between player and virtual character were higher on the V/E version. However, the difference between having Voice and Emotion together and having only one of them was not as big as the difference between having only one of them and having neither. Other analysed data was omitted from Table 3-4 for readability reasons, since according to the ANOVAs, they were not significant.

The effects that only Voice and Emotion alone had on how the character affected the participants were significant $F(1,11)=4.6$, $p=.05$ and $F(1,11)=5.6$, $p=.04$, respectively. The effect of having them combined was not significant $F(1,11)= 2$, $p=.19$.

Table 3-4 Mean and standard deviation of selected analysed items

	Condition	Mean	Std. Deviation
Character Affect on Player	V/E	52.50	29.35
	NV/E	45.83	24.66
	V/NE	41.67	25.17
	NV/NE	26.67	18.26
Player Affect on character	V/E	63.75	31.99
	NV/E	57.50	26.67
	V/NE	55.83	26.10
	NV/NE	40.83	29.37
Overall realism	V/E	56.67	28.31
	NV/E	50.83	22.24
	V/NE	50.83	22.65
	NV/NE	35.42	23.30

All the participants understood her expression of Happiness; 10 out of the 12 understood her sadness; 9 understood anger; and 7 understood boredom. Players also tended to share the character's emotions more often when they were presented using the character's face (see Table 3-5). Most participants declared feeling happy when she was angry, and bored when she was sad.

Table 3-5 Participants sharing character's feeling based version

Version/Feeling	Happy	Sad	Bored	Angry
V/E	6	2	6	2
NV/E	5	1	5	1
V/NE	4	1	4	1
NV/NE	3	0	5	1

Most participants self-declared competitive in the V/E condition and most self-declared competitive when she was bored (see Table 3-6 and Table 3-7).

Table 3-6 Version in which participants declared to be most competitive while playing the game

Version	Number of Participants
V/E	7
NV/E	1
V/NE	1
NV/NE	1

Table 3-7 Character's emotion in V/E and number of Participants who self-declared to be competitive

Version/Feeling	Participants
happy	1
sad	0
bored	6
angry	3

Section 3.7 Discussion

Section 3.7.1 General Observations

These results show that a virtual avatar showing even simple emotions can bring value to characters in a VR environment and that non-life-like characters can also elicit feelings, some quite strong, in the players.

We found it interesting to see that any kind of emotional expression from the virtual character has the potential to make the player feel more affected by the game, even if the character is not very life-like looking. Furthermore, aggregating many emotional expressions (i.e., using both verbal expressions and facial expressions) might not bring extra benefits in terms of eliciting emotional affect in the game players.

Sadness and anger were feelings not shared often between players and the character. From the data, we can extrapolate that, since the player was teasing the character, they felt accomplished when she got angry since teasing would often be seen as a friendly behaviour [107]. We came to this conclusion because, when the character was sad, the participants did not feel any joy but felt bored instead.

We also observed the participants became more competitive when the character was bored because they wanted her to express some other emotions. This finding was supported by our other observations, which indicated that participants wanted to be more competitive when they were playing the version of the game in which the character tended to be the most realistic (V/E).

Section 3.7.2 Limitations

Since we were evaluating only whether simple, but common, emotions would be enough in a VR environment to affect a users' mood and behaviour, we kept the Big-Five personality trait values in a "neutral" configuration; all values were set to half of the maximum value, indicating moderate levels of each of the five personality traits. This does not change how the emotions are expressed in our study, only the speed at which the emotions change. Our results might have been more varied if we had used more personality configurations.

We only had a relatively small number of volunteers, volunteers were fairly homogeneous in terms of age and ethnicity, and virtually all had experience with games and VR. This makes it difficult to assert that our results are completely generalizable (for example, to other ethnicities and non-gamers). However, in line with findings from other researchers [86], it is extremely likely our findings can be applicable to a wider population because the character's emotions are basic and can be easily understood by most people across generations [108].

Finally, we have not compared our character with a life-like one. However, this is purposeful, since we wanted only to see if cartoon-like characters in VR environments could benefit from adding some simple emotions to increase users' interest and engagement.

Section 3.8 Summary and Remarks

In this research, we described a brief history of the creation of personality-based systems for virtual agents. Our review of the literature shows that non-life-like characters are not often studied, especially in virtual reality environments. This has inspired us to explore further the effects of adding simple expressions to cartoon-like avatar to alter users' behaviours within the gaming environment. To do this research, we have adapted a personality system from the literature and applied it to a female cartoon-like character so that it can express feelings of happiness, sadness, anger, boredom, and relaxation. The character can also talk or be silent (while showing emotions or not showing emotions).

We have then conducted a study with different versions of the character (with Voice and/or Emotion and without them). The findings of our experiment suggest that being able to express simple feelings that users can clearly distinguish, either by using facial expressions or by using vocal expressions, is enough to have a significant effect on their behaviour. Users can become more interested and engaged with the avatar and environment.

Their interaction also becomes focus on having an impact on the mood of the characters. The findings of this research are applicable to VR systems that use non-player characters that are aimed at increasing users' engagement and interest in the environment. These results are especially important for low-budget game developers who might want to make educated choices in design.

In the future, we plan to do further analyses, such as analysing other types of vocalized cues, different genders, and different levels of realism of the virtual characters. Such studies will help better understand our findings and allow us to better generalize them.

Chapter 4 Evaluating the Need and Effect of an Audience in a Virtual Reality Presentation Simulator

Section 4.1 Introduction

The fear of public speaking is the most mentioned in surveys on fears [109]. Training for presentations can be challenging. However, in VR environments, virtual characters can be used to train presenters to better cope with their fears. Studies on the usefulness of a virtual audience (VA) for training people to make public presentations are limited, focusing mostly on people seeking help for diagnosed phobias [110], [111]. However, in many situations, other people would also like to practice interactive skills but lack the proper training environment, as is the case in language learning [112], [113]. Finally, some people would just like to enjoy talking scenarios like those present in RPG and simulation games. There are several games in which conversations are a core mechanic, for example, *We should talk*[4] and *Secret Hitler*[5].

It has been shown in previous studies that people, regardless of their fear of public speaking, present different heart rates when exposed to VA [114]. However, because this study was done with older technology (i.e., low polygon models), it is unclear if all people will still be affected by a VA [115]. This is because even simple changes such as altering the viewing perspective can affect VR interaction [116], [117].

Furthermore, even though much research has been done to demonstrate that a VA can yield positive results in the treatment of phobic people [118]–[121], not as much has been done to determine the impact of a VA on the general population. To this end, in this research, we developed a VR simulation environment to enable players to simulate virtual

[4] www.weshouldtalkgame.com/
[5] www.secrethitler.com/

public speaking activities or presentations. We ran an experiment to gather feedback from people after the simulation to evaluate the acceptability of our system as a simulation scenario.

Section 4.2 Related Work

Virtual reality exposure therapies have shown potential, even before modern VR HMDs [110], [119], [122], [123]. For instance, they have been shown to work well for different kinds of anxieties [111], [124], including public speaking [125]. Beyond therapy, one of the most prominent uses of VR is for training [126]. Several studies have shown that VR is a disruptive and useful tool for the training of various skills [123], [127], [128], including technical [126], physical [16], [129], and sociological training [130]. Yet, realism is not always best [131]; even low-end VR can yield positive results [132].

Section 4.2.1 Simulation Games

Simulators as a form of gaming have existed since at least the 1980s, with city simulators like *Fortune Builder*[6] and *SimCity*[7], used to simulate urban planning. Many have started as a tool for teaching some specific set of skills, especially resource management [133], or to be able to experiment with different techniques and actions to predict an outcome. *The Sims*[8], for example, was proposed as a social simulator [134] and has been used and studied for its potential in language teaching [135].

More recently, VR simulation games have been proposed for a plethora of activities[136], from the most mundane to the highly specific. In this genre, we can mention games such

[6] coleco.com/
[7] www.ea.com/games/simcity
[8] www.ea.com/en-gb/games/the-sims

as *Job Simulator*[9], *Bus Driver Simulator*[10], *Thief Simulator*,[11] and *Surgeon Simulator*[12], games whose name are self-explanatory of the game's content.

As they can be used for mundane and specific activities and have been historically used to teach different skills, it is natural for this game genre to be used to teach in VR. For example, researchers have developed games to teach firefighting skills, with an undetermined degree of success [137], and teach medical procedures with a promising degree of success [138]. As we will explore in the next session, VR can be used to simulate presentations.

Section 4.2.2 VR as a Presentation Simulator

If conditions are adequate, VR can elicit similar presence levels compared to the real world in an interview setting [139]. It can even provoke fear [114]. Even though these earlier models were not as aesthetically realistic as what we can see in today's VR, people with social anxiety presented clear physical responses to them. However, these responses were not as prevalent in confident people. The reason for these responses may be because these participants thought the audience was not as realistic [114], or they ignored them [140]. Recent improvements in the aesthetical realism level might have changed this paradigm or caused socially anxious people to get so used to the technology that they do not get triggered anymore. Hence, this raises a question (RQ1): *What are people's opinions about VA made with models built with a high polygon count derived from 3D-scanned human faces*?

A newer study analysed if a realistic environment could increase feelings of presence in a virtual presentation environment [140]. The researchers in this study created a

[9] jobsimulatorgame.com/
[10] en.kishmish-games.com/bus-driver-simulator
[11] www.playway.com/component/content/article/10004-games/738-thief-simulator-vr?Itemid=131
[12] www.surgeonclassic.com/

classroom identical to the one available in their real-world environment. They then asked participants to present to a VA and participate in a semi-structured interview. They observed that participants, overall, appreciated the experience, and some felt as if it became more real as they got accustomed to it. However, some stated not perceiving the audience, even though they appreciated the experience. These responses raise our main research question (RQ2): *Is simply having an empty virtual classroom equivalent to having a VA for (some) people in a presentation simulator?* To find the answer, we created a presentation environment in which participants could present either to a realistic empty room or partially 3D-scanned VA.

Section 4.3 The VR Presentation Simulator

We developed a VR presentation simulation game to give players the chance to see how they would fare in a presentation. We utilized pre-made models for the VA's body and movements. However, their heads were customizable 3D models of actual individuals; in our case, the researchers (see Figure 4-1). Based on the literature, we hypothesized that this would improve the simulation and help the users feel more at ease before a real-life presentation [114], [140].

The VR simulator contained a slide presentation on the back and footnotes in the front, which allowed users to present looking at the audience. The footnotes were akin to the presentation component in Microsoft PowerPoint; it let participants have notes in front of them, like a prompter. The challenge within this simulator was to hold a good posture while presenting to the always attentive audience.

Section 4.4 Experiment Design

To evaluate whether the VR audience was indeed impactful to the players and to validate our VR presentation simulation game, the experiment had three conditions. In one condition, the participants practiced in front of a virtual audience (VRP). In another condition, the practice happened in an empty virtual room (VRE). The control participants also practiced in the room alone without VR (RLE).

We made a series of presentations in English with topics that the volunteers were not familiar with, according to their backgrounds. The participants prepared within a specific time to become familiar with the topic, and then presented the slides to a real-life audience. For all these later presentations, there were always the same audience members. Since the participants could not see their virtual face, the face model was not altered (see Figure 4-1). The participants' virtual hands matched their overall skin tone.

Figure 4-1 The faces of the audience and the prompter (left). The virtual reality presenter and environment (middle). One presentation in real-life (right).

Section 4.4.1 Environment Metrics

Each participant was asked to complete a simple questionnaire to collect demographic data. All participants were asked to self-assess their English level among the 6 levels of the CEFR [139]. They were presented with a list that explained what abilities each level represented and were asked to select the one closest to what they believe their level was: A1 represents a beginner and C2 fluency, B1 and B2 are intermediary levels.

Afterwards, they completed the Personal Report of Confidence as a Speaker (PRCS) [141], which is a true-or-false questionnaire with a maximum score of 12 points. The higher the volunteer scored, the more anxious the volunteer indicated feeling.

We used the 5-point Likert-scale Self-Statements During Public Speaking (SSPS) [142], which is referred to as a marker for short-term treatments. The higher the score, the more confident the presenter. Furthermore, we asked the participants to rate from 0 (worst) to 5 (best) how good they felt the conditions were, and their familiarity with the topic.

After going through the three conditions, we asked participants to rank each condition according to a list of criteria and to answer a few open-ended questions (see Table 4-1).

Table 4-1 Open questions presented at the end of the experiment.

ID	Question	Type
R1	Rank your preparation after each version	Rank
R2	Rank your nervousness after each version	Rank
R3	Rank your version preference	Rank
Q1	How do you feel about the virtual environment, not including the audience?	Open
Q2	How do you feel regarding the virtual audience?	Open
Q3	Is there anything you would change in the virtual Environment?	Open

Section 4.4.2 Apparatus

We used an Oculus Rift CV1 as our HMD, and two standard 27” 4K monitors. We used the Oculus Touch to control the presentation (even the real-life presentation). Participants did not need to learn different controls for each condition. We used monitors in real-life to simulate both the screen on the back and the text area facing the volunteer. All text sizes were set to be the same in VR and on the monitor.

Section 4.4.3 Participants

We recruited a total of 12 participants (seven females; four males; one non-binary) from a local university. They had an average age of 19.88 years (*SD* = 1.32), ranging between 19 and 23 years. No volunteer mentioned health issues, physical or otherwise. No participant was a native English speaker. Half had experience with VR systems before the experiment. The participants were not offered any reward to participate in the experiment.

Section 4.4.4 Procedures

Each participant was assigned a specific order of the three conditions in which to practice the presentation. The order was Latin Square counterbalanced to mitigate carry-over effects. Participants were briefed about the experiment.

Next, all participants were presented with a demonstration introducing the Oculus Rift and a static virtual environment to get them acquainted with the VR HMD and set up; they were presented to the virtual room and a mock slide presentation.

Then, each participant was left alone to complete the training without being observed. The participants had 15 minutes to train for the presentation on the current condition. After this time, the participants were requested to remove the HMD and answer the questions. They then presented the topic in real-life.

The participants were given 30 minutes to rest before starting the next condition. After all conditions were tested, the last questionnaire was given to the participants. The participants were told that the open-ended questions could be answered in their native language if they felt more comfortable doing so. At the end of the experiment, the participants were thanked for their time, and offered time to rest and some refreshments. On average, each volunteer took one and a half hours to finish the whole data collection process.

Section 4.5 Results

The data were analysed using both statistical inference methods and data visualizations. We conducted Mauchly's Test of Sphericity. We also employed Repeated Measures ANOVA (RM-ANOVA) using Bonferroni correction to detect significant main effects. If the assumption of sphericity was violated, we used the Greenhouse-Geisser correction to adjust the degrees of freedom. To control for both Type I and Type II errors we chose $p < 0.10$ for our analyses [143]. For the open-ended questions, we analysed the sentences individually.

Section 4.5.1 Pre-Questionnaire

We defined the bins for PRCS, normally distributed, as between 0 and 4 indicating feeling not so nervous, up to 8, which indicated moderate feelings of nervousness. Anything above 8 indicated extreme nervousness. We dubbed these bins as groups Not Anxious (NA), Standard Anxious (SA), and Extremely Anxious (EA), respectively. Fifty percent of the EA participants declared their level as B2. The other volunteers were mainly self-declared as B1.

Section 4.5.2 Questionnaires of Nervousness and Preference Order

The SSPS questionnaire revealed significant differences after having practiced in the virtual environment; the participants overall felt more positively to present to a real audience ($F_{2,\ 22}$ = 3.051, p = .068). The positive effect was more present on the NA participants (F2, 6 = 4.043, p = .077). Post-hoc analysis showed that VRP was especially more successful (see Figure 4-2) in making the participants feel less anxious (*p* = .085). The participants rated the VRP version the best simulator or training tool (see Figure 4-2). However, EA participants did not rate it more positively as a simulator than the other versions. The main difference in ratings came from the NA participants.

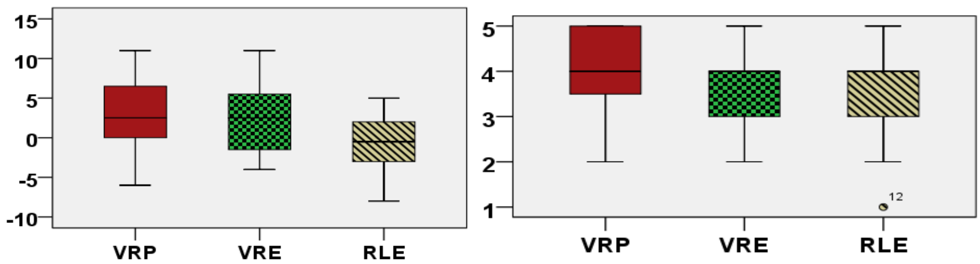

Figure 4-2 SSPS answers after training in each version (left); VR versions scored slightly better. Participants' ratings for each condition as a training tool (right).

The overall ranking disregarding PRCS can be seen in Figure 4-3; 75% of the EA participants chose one of the empty rooms as the version they felt most prepared in (R1). However, they all selected the VRP as their second option. All EA participants selected the VRP as the version in which they felt the most nervous (R2). Half chose it as their first choice, and 50% preferred it as their second version (R3). Only one EA participant chose VRE as their first choice. Overall, VRP was the condition that received the most positive overall feedback. Only one NA participant chose VRP as their first choice. The SA participants consistently chose it as their first choice.

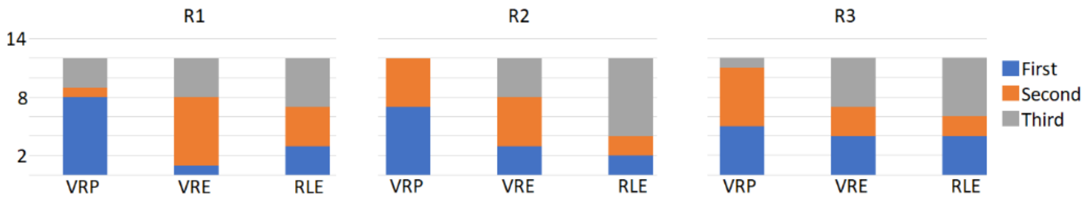

Figure 4-3 Ranking to questions from R1 to R3, in their respective order from left to right.

Section 4.5.3 Open-Ended Questions

The analysis of the interviews showed positive results and feedback. For Q1, one of the most common words reported by the volunteers is that the environment made them feel "good," and it was or felt "real".

According to Q2 and Q3, participants believed that a virtual audience made them nervous, but they thought it was beneficial. For example, P3 said "[I feel] a little bit nervous, especially when the audience's action changes." Some participants mentioned this as a positive point for making them better prepared to speak to the real audience later. P11 said that "There will be some tension, but it will then be easier in reality." P4 mentioned that the audience was good, even though they made him feel "serious".

The EA participants were the ones who most used words such as "nervous" and "scared". However, those were also the ones who commented most positively about the use of the application. As P3, who scored the maximum in the PRCS, puts it, "[It] made me not so relaxed but really helped me improve my presentation." In general, participants said that they did not want to change the audience.

Section 4.6 Discussion

The answers to the open questions and the ranking suggest a preference for VRP, which also indicates that the use of this simulation game can be a VR tool for training that has the potential to be well-accepted, which is aligned with previous research [114], [140]. Furthermore, the answers also show a positive aspect of having the faces of the real audience in the tool, which were described as having given them "good nervousness".

Hence, the participants' opinions about a VA made with a high polygon count derivate from 3D-scanned faces (*RQ1*) were positive.

The rankings and the open-ended questions led us to believe that the fear factor generated by the virtual audience is favourable to the participants who, in return, trained harder or started feeling more secure with themselves. This is expected based on the exposure therapy treatments found in the work of Slater et al. [114]. This might indicate that an even greater audience might bring even more beneficial results in response to the simulation.

Overall, VRP was verbally declared the most popular presentation simulation environment. Thus, in most cases, it is valuable to have a virtual audience interacting with the player to improve the scenario. However, some people are satisfied with the empty room simulation, which answers positively to the *RQ2*. These results indicate that the game used in this study can be used similarly to other simulation games, such as the *Bus Driver Simulator*.

Although most participants are content with the VRP version, we recommend adding a toggle button to satisfy players who might appreciate the VRE, or alternative game modes, as the requirements for EA people seem to be different than that of those who just need to practice presentation.

Section 4.7 Summary and Remarks

In this research, we explored whether there is a need (and an effect) of a virtual audience in a presentation simulator using VR. We did this analysis through a series of subjective metrics which indicate that even though most people do appreciate the virtual audience generated with the 3D-scanned faces, it makes them somewhat positively nervous. For some people, the current virtual audience is not necessary and they are just as satisfied with the game without it.

Furthermore, we observed that a simple audience that follows the users by looking at them is enough for creating a stimulating simulator for presentations in a VR environment. This is a game that aims to invoke feelings of anxiety and promote effective training. This VR environment was able to make users nervous and was rated 4 out of 5 as an effective training tool. We observed that the development of a public speaking simulator is the most adequate when it can have an audience. We also observed that players would prefer to have the option to turn it on and off, depending on their preferences. Overall, the results show that the game we created seems able to help students prepare for presentations.

Chapter 5 Comparing the Effects of a Familiar Audience with an Unfamiliar One in a VR Presentation Simulator

Section 5.1 Introduction

As seen in the previous chapter, the fear of public speaking has been on the top of the "human fear list" for about 30 years, even above "death" [144], and it is experienced by many, from students to presidents [109]. Moreover, it may preclude many people from opportunities for social contact, or even career advancement by making it difficult to create social networks [145]. Thus, getting used to speaking in front of people is a valuable skill.

Training or practicing public speaking alone is not ideal. The lack of feedback can be a challenge, and it is not uncommon to wrongly underestimate or overestimate one's own skills [146], [147]. On the other hand, training with real people can be either costly or involve asking favours to friends and family whose advice may be biased or lacking.

The idea of using VR to train and practice public speaking skills as a practical alternative to practicing in front of a mirror or enlisting reluctant friends has been widely researched for several years [112], [114], [118], [120], [148], [149]. A virtual environment (VE) with an audience rendered in high representational fidelity can affect the speaker's confidence [140], [149], [150], and has often been used to conduct exposure therapies to treat public speaking anxiety [114], [118], [149]. However, in some cases people ignore the audience in VEs, perhaps due to a lack of connection with the virtual audience [140]. From our previous analyses with our simulator, an audience made partially from 3D-scanned faces increases is preferred over an empty room. Yet, it remains unknown if increasing realism in the facial features of a virtual audience to mimic the faces of a future, real-life audience can affect how users interact with the environment and their anxiety levels compared to other kinds of audience.

This study investigates if a 3D scanned virtual audience of actual people influences the participants' anxiety and if this extra layer of realism is appreciated by people simulating giving speeches.

Section 5.2 Related Work

Section 5.2.1 Body tracking in Games

Physical gestures have long been thought to be the foundation for all human verbal communication, and using gesture interfaces allows one to create unique learning opportunities that take this into account [151]. With the advent of gesture-based computing and low-cost gesture recognition technologies, educational games will now experiment with new interactions [151]. That means feedback can be based on the players' physical actions.

Some games like *Fruit Ninja*[13] for the X-Box Kinect use simple positioning and changes in the player's arm position to "split" virtual fruits. However, other games make use of more complete models, such as *Dance Central*[14] and *Just Dance*[15]. These games base the player dance score on how a complete player model is positioned and moving compared to an existing ideal model. For the model creation, they make use of camera-based body tracking. This kind of tracking can be used for other in-game feedback, for example, to evaluate the body language in a presentation simulator.

Section 5.2.2 Presentation Training Feedback

Because nonverbal communication (e.g., body language) is a fundamental trait of human communication, it is an important skill that should not be overlooked [152]. It can be profoundly influential in how someone is perceived by others and the quality of their interactions [153]. Hence, it is essential to identify when ineffective and inadequate body

[13] www.halfbrick.com/games/fruit-ninja
[14] www.dancecentral.com/
[15] justdancenow.com/

language is being used [153], [154]. An effective method of recording body language is by using motion tracking.

One of the most accessible consumer-grade devices for motion tracking is the Kinect by Microsoft. Kinect is a versatile tool; it has been used to scan people to create 3D virtual replicas [155] as well as in a variety of training applications that require body movements to be recorded, such as driving simulations [156], and even in elderly rehabilitation [157]–[159]. Moreover, researchers have used it to help with assessing the quality of a presentation based on body language [154], [160].

Section 5.2.3 Audience in VR Presentation Simulators

Due to the benefits provided by VR in public speaking training, various training tools have been designed to help people overcome their feelings of anxiety and improve their speaking skills [114], [118], [120], [121], [123], [149], [161], [162]. VEs made from 360-degree videos of lecture rooms have been shown to be effective in stimulating anxiety in certain people [149]. However, VEs using lower polygon versions of people did not elicit any unusual feelings in more confident people [114]. These participants possibly ignored the virtual audience like participants did in a similar study [140]. This might be caused by a lack of connection between the participants and the virtual audience. Given this possibility, we propose the following three research questions: (**RQ1**) *Can copies of familiar people increase the participants' engagement with the simulator (e.g., waiting for their acknowledgment, talking directly to them)?* (**RQ2**) *Does presenting to a real-life audience become more comfortable (e.g., having a better posture or having a lower heart rate) if the participants have presented previously to a virtual copy of this audience?* (**RQ3**) *Can this level of comfort be observed through the participants' speech patterns?*

Realism includes not only the overall look of the virtual audience but also their behaviour [120]. Research has shown that a simple gaze from an avatar can change how people interact with it [96]. Similarly, boredom gestures from an audience can be intimidating [120]. However, when gestures are random, presenters might not notice them [121]. So, avatar gestures should communicate emotion meaningfully and authentically [163].

People expect a virtual character to behave like a real person when it looks lifelike [99]. Inconsistency between a character’s behaviour and its appearance can cut off the emotional connection to a character [141]. However, in simulators, players might not try to develop the connection in the first place [140]. This prior research led us to develop our following two-part question: (**RQ4a**) *In a presentation scenario, do people notice virtual audience movements associated with their performance?* and (**RQ4b**) *Do these movements affect them?*

To find answers to these research questions, we improved the VR simulator from Chapter 4. The simulator penalized negative body language by providing visual feedback (e.g., the audience members crossing their arms). In this simulator, the audience’s heads were either 3D scanned copies of real people or generic stock models.

Section 5.3 The VR Presentation Simulator

We advanced our VR presentation simulator to allow users to practice their presentations and to allow us to investigate the importance of realism and feedback. It simulates a formal environment (see Figure 5-1).

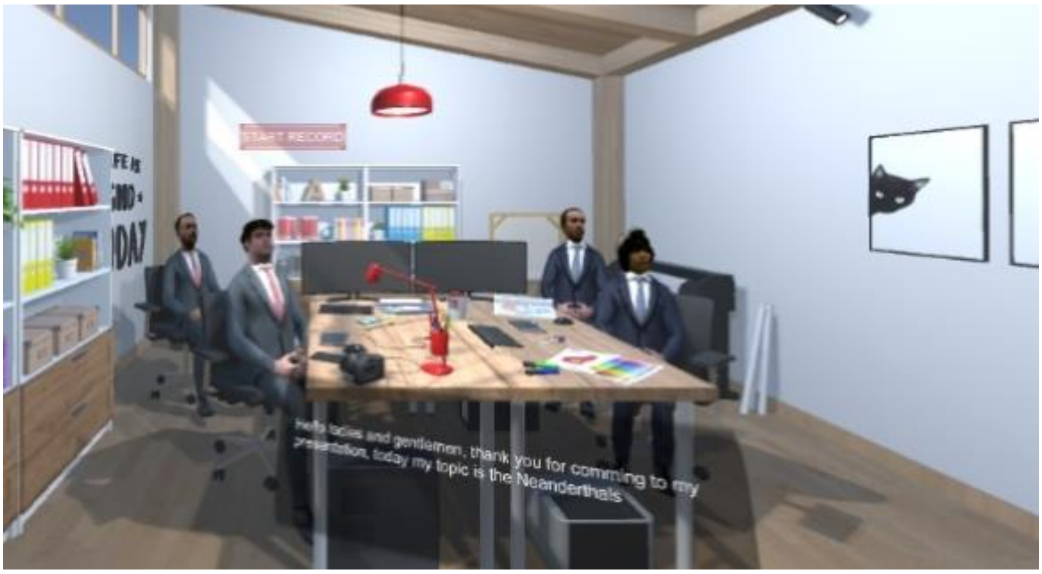

Figure 5-1 A screenshot from the presenter’s point of view. The heads were scanned replicas of the actual people who would be present during the live presentation.

In one version, we used stock models from the Unity asset store as our base models because their animations could demonstrate boredom [153], [154], [163]. In another version, we used customized 3D model heads of the researchers (see Figure 5-1). virtual audience members wore a suit, and we adjusted their size to match the body type of the researchers.

Because we aimed to present the players' scores in-game through the avatars' actions to see if the participants would notice it, the environment was designed to record presentations as they happened. A Kinect was used to identify the series of body language movements exhibited by presenters and penalize or reward them. The point system was based on the amount of time the presenters were in inadequate positions (e.g., crossing the arms, or holding a hand behind the body) [153], [154]. One point was removed for every 5 seconds in which the presenter stayed in an inadequate position. The grade was passively communicated to the presenters in real time through animations of the virtual audience. virtual audience members would cross their arms or otherwise exhibit body language that typically suggest a disconnect between the audience and the speaker. However, the virtual audience always maintained attention, demonstrated by eye contact, with the presenter. The numeric grade (points) for Body Score, and the HR were recorded along with the video of the presentation. This would allow the presenters (players) to re-watch and understand why they lost points and to observe what kinds of movements were more common when they were nervous or felt uncomfortable. The audio was recorded using the microphone embedded in the VR HMD.

The simulator contained a slide presentation behind the player. In addition, there was a prompter in front of the player that displayed notes that reminded the player of the content of the presentation. The prompter worked like the presentation component in Microsoft PowerPoint. Thus, players did not need to completely memorize the content of the presentation, as memorization can be used to change the game difficulty. This feature is both helpful to people who want to play with one less degree of difficulty in their presentation (i.e., memorization) in VR and helpful to the research by removing memorization as a potential confounding factor.

Section 5.4 Experiment Design

To evaluate if using 3D scanned faces in VR would change the participants' responses to the simulator, the participants practiced and then presented in four conditions. The conditions were divided by the familiarity with the audience (familiar and unfamiliar) and by environment (virtual reality and real-life). The conditions are summarized as follows:

- VRF - VR audience with 3D scanned faces of the Familiar audience.
- RLF - Real-Life audience of the familiar faces (actual researchers)
- VRU - VR audience with stock models faces different from the audience (Figure 5-2.)
- RLU - Real-Life audience of the unfamiliar faces (invited audience)

Figure 5-2 A close-up picture of the stock models used as the Unfamiliar VR audience (VRU).

We created two presentations in English with topics that required specialist knowledge and were therefore mostly unfamiliar to lay-persons. This was done to reduce the chance of a participant being familiar with a topic and thus having a different affective reaction. These presentations were then validated by lecturers of the university's School of Languages as being suitable according to the participants' language proficiency levels. The presentations were designed to last approximately 3 minutes.

The participants were given 10 minutes to prepare in a room without people and then present in VR to the virtual audience. Three days later, they presented the same content to a real-life audience. For all the presentations, there were always four audience members, and participants were aware of the different audiences. Because the participants could not see the virtual face of their own avatar, the avatar's face was not changed to match each participant (see Figure 5-2). However, the participants' virtual hands matched their skin tone and were holding controllers that mirrored those of the VR device.

Section 5.4.1 Environment Metrics

Each participant was asked to complete a demographic questionnaire. All participants had their English level pre-screened to be around B1 of the Common European Framework of Reference for Languages [139], as in Chapter 4. Subsequently, they were asked to complete the Personal Report of Confidence as a Speaker [141], which consisted of a series of sentences to be classified as true or false and later summed (maximum 12), which would determine their level of public speaking fear.

For all conditions, to detect anxiety we recorded participants' HR [114] and saved the calculated Body Scores (BS). At the end of the experiment, we asked the participants open-ended questions about their experience in the simulator (see Table 5-1).

Table 5-1 Open questions presented at the end of the experiment.

ID	Question	Type
Q1a	Did you notice any similarities between the VR audience and the real audience?	**Open**
Q1b	How did this affect you while giving a presentation?	**Open**
Q2a	Was giving a presentation to a VR audience different from giving a presentation to a real audience? In what ways?	**Open**
Q2b	How did this affect you while giving a presentation?	**Open**

Section 5.4.2 Apparatus

We used an Oculus Rift CV1 as our HMD, as it is one of the most popular off-the-shelf VR devices. The HMD was connected to a desktop with 16GB RAM, an Intel Core i77700k CPU @ 4.20GHz, a GeForce GTX 1080Ti dedicated GPU, and two standard 27"

4K monitors. The camera was a Microsoft Kinect One, and the HR monitor was a POLAR OH1. We used the Oculus Touch as the input device in both practice and presentation rounds. To replace the VR prompter and projector in real-life, we used a computer monitor. The size of the text notes was similar in VR and in the monitor.

Section 5.4.3 Participants

We recruited a total of 10 participants (five females) from a local university. They had an average age of 19.2 years (*SD* = 0.74), ranging between 18 and 21 years old. All volunteers had normal or corrected-to-normal vision, and none of them declared any history of colour blindness or other health issues, physical or otherwise.

All participants were native Mandarin speakers. None of them had previous VR experience. All participants declared feeling some degree of shyness when presenting in English, even though more than half (60%) declared feeling confident with their speaking abilities outside of formal environments. The participants were each given a 10 RMB gift card for the successful completion of the experiment.

Section 5.4.4 Procedures

The experiment was divided into four days along two weeks. In the first week, half the participants presented to the VRF on the first day and then presented to RLF; the other half presented to the VRU and a few days later to the RLU. In the following week, each half did the opposite (see Table 5-2). This was done to mitigate carry-over effects.

Table 5-2 Experimental order and division by day and week.

	Week 1		Week 2	
	Day 1	Day 2	Day 1	Day 2
Group 1	VRF	RLF	VRU	RLU
Group 2	VRU	RLU	VRF	RLF

A written brief of the experiment was presented to each participant in their native language. They were encouraged to ask for any further clarification before completing the pre-experiment questionnaires and starting the experiment.

Next, all participants were presented with a demonstration introduction of the Oculus Rift and a static virtual environment to get them acquainted with the VR HMD and set up—they were presented to the virtual room and a mock slide presentation. Participants were then introduced to the data collection devices, the region in which they could move around, the rules of the presentation, and the expectation that they would give a presentation to a real audience in the future.

Then, each participant was left alone to complete the training without interference. The participants had 10 minutes to practice on the first day and then up to 5 minutes to present to the audience. If the participants extended their presentation to over 5 minutes, they were cut off. In all presentations, the participants were recorded (see Figure 5-3) and had their HR monitored.

Figure 5-3 Example of a recording of a participant presenting in the RLF condition.

After all conditions had been tested, the last questionnaire was given to the participants. The participants were told that the open-ended questions could be answered in their native language if they felt more comfortable doing so. In the end, the participants were thanked for their time, offered time to rest, and given refreshments and the gift card.

Section 5.5 Results

The participants' scores for the Personal Report of Confidence as a Speaker (PRCS) were normally distributed. We defined the bins between 0 and 4 as not so nervous (NA), which consisted of four participants, and between 5 and 8 as regular nervousness (SA), which had five participants. Any score above 8 was considered extremely nervous. When comparing between subjects we used Mann-Whitney U.

For all of the data, without bin separation, we used SPSS version 24 to do the analysis. We analysed the data both statistically and graphically. The Shapiro-Wilk test was used to test the normality of the average HR in beats per minute (BPM) and the final body score (BS) generated by the system. None of these data was normally distributed. Therefore, Wilcoxon was used to evaluate the significance between the two conditions (*F*amiliar and *U*nfamiliar), while Spearman was used to evaluate the existence of correlations.

Line charts are used to analyse the participants' HR data graphically. We calculated the difference of HR between VRF and RLF (ΔF) and between VRU and RLU (ΔU), to see if the anxiety was similar in those cases. Furthermore, because our graphical analysis diverged from our statistical analysis, we used a multilayer perceptron (MLP) to understand our data in greater depth.

Section 5.5.1 Body Score

A Wilcoxon signed-rank test showed that the face models did not elicit a statistically significant change in Body Score in the VR sessions ($Z = -.890$, $p = 0.373$) nor in the RL sessions ($Z = -.140$, $p = 0.889$). The mean was 94.4 with the VRU and 96.6 with the VRF, while it was virtually identical for the real-life sessions, 92.5 in RLU, and 92.9 in RLF (see Figure 5-4). The results among versions, divided by bins, did not show statistically significant results, as shown in Table 5-3.

Figure 5-5 shows that the mean BS was better for the participants in the NA bin (98.0, *SD* = 2.1) when presenting to the RLF audience than the participants in the SA bin (89.8, *SD* = 13.83), however, the opposite happened when presenting to the RLU audience. In the RLU condition NA scored lower (87.25, *SD* = 14.5) than SA (95.8, *SD* = 8.3). However, none of the results comparing the different bins seem to be significantly different: RLU (U = 17.0, p = 0.081), RLF (U = 7.5, p = 0.535), VRU (U = 17.5, p = 0.061), and VRF (U = 17.5, p = 0.061).

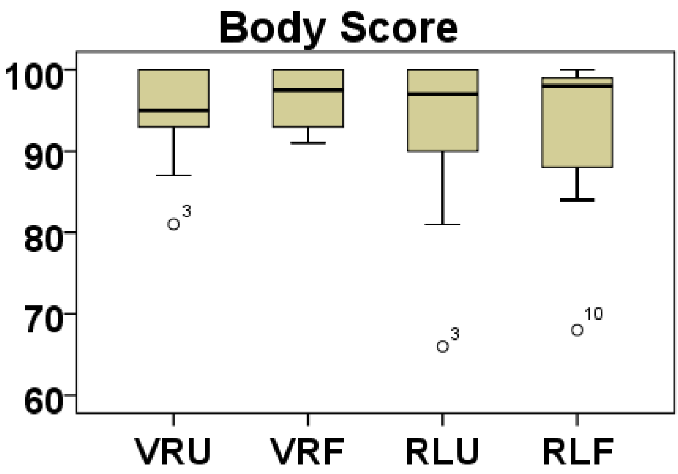

Figure 5-4 Bar graphs of the Body Scores for each condition and phase.

There was no correlation between body scores and the other metrics, within the same version; however, there was a strong negative between BS in VRU and HR in VRF (r = -.971, p=.001).

Table 5-3. Wilcoxon results for the difference in Body Scores between versions separated by Personal Report of Confidence as a Speaker bins, rows with p>.25 are omitted. (NA: Nervous; SA: Regular Nervousness)

	NA		SA	
	Z	p	Z	p
RLF – RLU	-1.826	.068	-1.604	.109
VRU – RLF	-1.826	.068	-1.214	.225

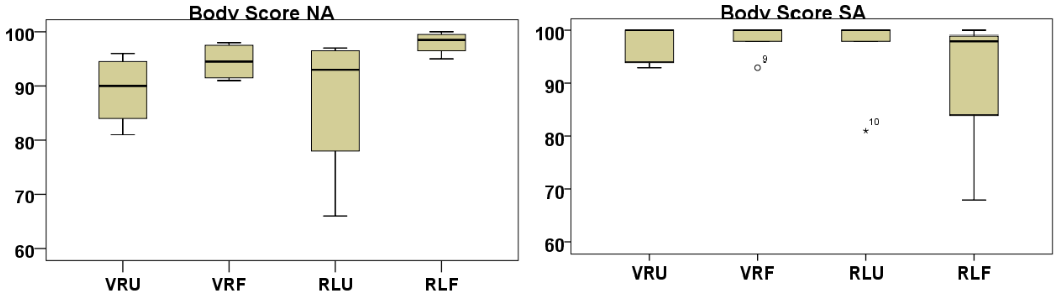

Figure 5-5 Bar graphs of the Body Scores for each condition and phase separated by Personal Report of Confidence as a Speaker bins.

Section 5.5.2 Heart Rate

A Wilcoxon signed-rank test showed that the HR from the VRF was statistically equivalent to that of the VRU in the presentation (Z = -.980, p = 0.327). The mean HR was 102.2 BPM with the VRU and 99.6 BPM with the VRF.

Regarding the difference between RLF and VRF, there was a statistically significant difference (Z = -1.960, p = 0.050); the mean HR in RLF was 107.4 BPM. It was the same case for RLU and VRU (Z = -2.666, p = 0.008); the RLU mean is 113.8 BPM, as shown in Figure 5-6.

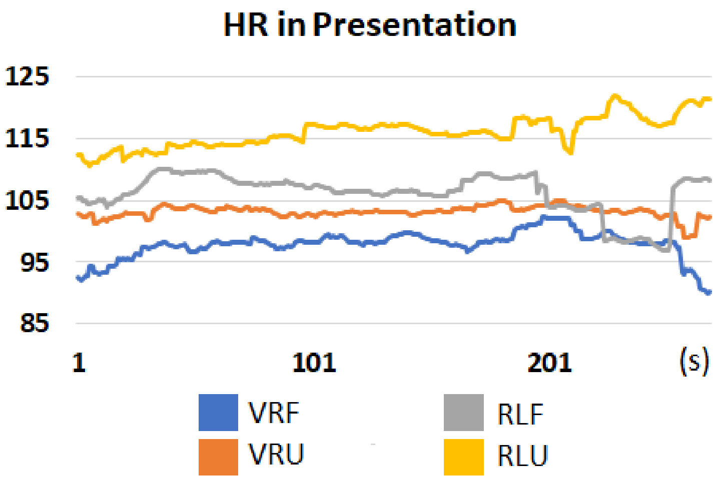

Figure 5-6 The HR in beats per minute (BPM) for the different conditions and phases.

The HR did not differ significantly for the real-life presentations either (Z = -1.400, p = 0.161). Furthermore, the test with all participants showed that there was not a statistically significant difference between ΔF and ΔU (Z = -.840, p = 0.401) not even within bins (p>0.5), as shown in Figure 5-8. The mean difference in HR was 13.4 BPM in ΔU and 9.2 in ΔF (see Figure 5-7).

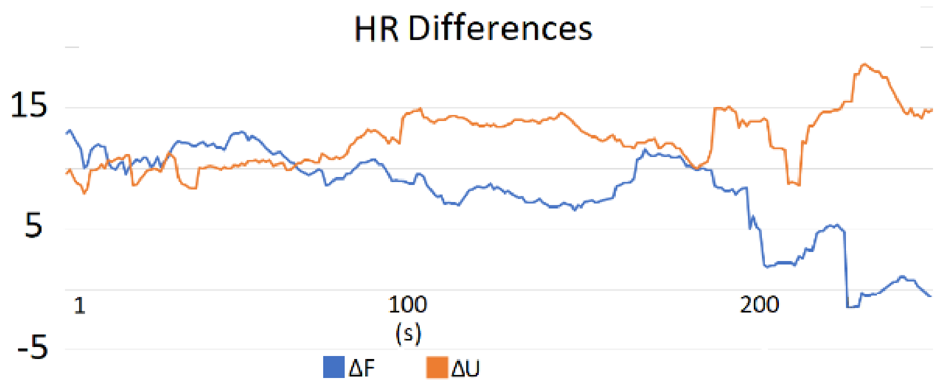

Figure 5-7 The difference in HR between VR and RL. The difference in Familiar condition is smaller overall.

However, a Wilcoxon signed-rank test with only the five participants who scored between 4 and 8 in the PRCS showed that there was a statistically significant difference between for ΔF and ΔU (Z = -2.201, p = 0.028). The mean HR difference was 4.6 for ΔF and 13.6 for ΔU.

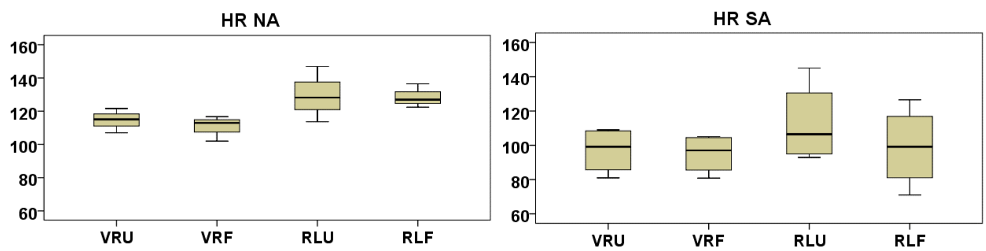

Figure 5-8 Bar graphs of the HR for each condition and phase separated by Personal Report of Confidence as a Speaker bins. (NA: Nervous; SA: Regular Nervousness)

There was a positive correlation in HR between RLU and RLF (r = .762, p = .028) and VRU (r = .893, p = .007), and between RLF and VRU (r = .821, p = .023). VRF did not correlate with the other versions.

Section 5.5.3 Classifying Heart Rate and Body Score with an MLP Model

The MLP model generated by inputting HR and body score could predict which audience the participant was being exposed to. The model had 83.3% accuracy during training and 100% accuracy during the testing phase. The importance of both HR and body score was equivalent in the prediction, as presented in Figure 5-9.

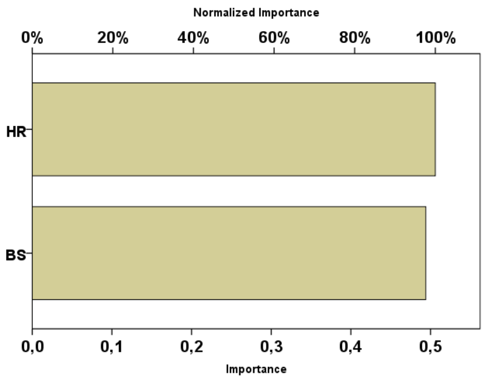

Figure 5-9 The importance of each component in the MLP model to identify to which audience the participants were presenting; BS in the graph stands for Body Score.

Section 5.5.4 Word Count and Speaking Time

The Wilcoxon signed-rank tests revealed a difference in spoken time between most versions. The tests showed that the Median (IQR) time in seconds for VRU, VRF, RLU and RLF were, respectively, 257, 235, 204, and 221 (see Figure 5-10).

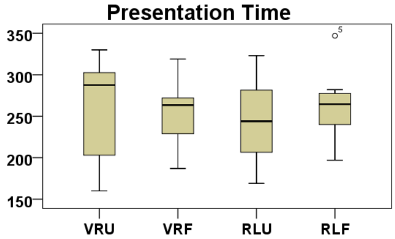

Figure 5-10 Quartiles of the speaking times per condition.

There were significant differences between VRU and RLU (Z = -2.073, p = .038), VRF and RLU (Z = -2.073, p = .017) and VRF and RLF (Z = -2.033, p = .042). When testing only the participants who scored between 4 and 8 in PRCS, no statistically significant differences were found (p>.05).

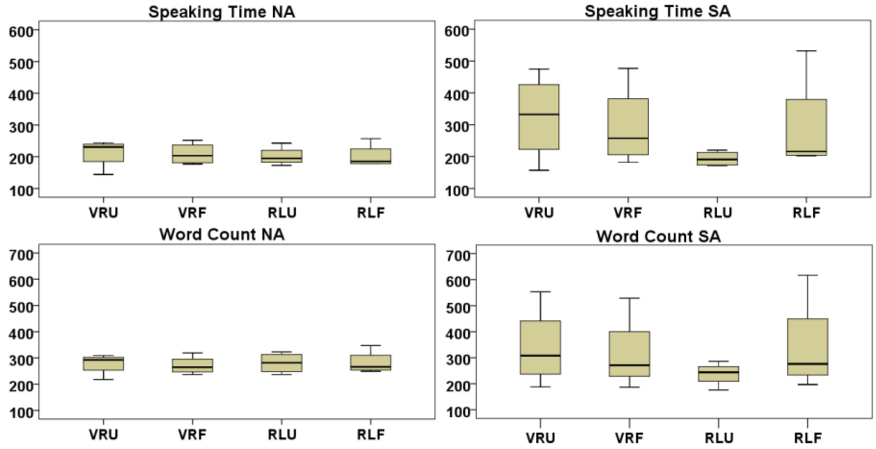

Figure 5-11 Bar graphs of the spoken time and number of words spoken for each condition and phase separated by the Personal Report of Confidence as a Speaker bins. (NA: Nervous; SA: Regular Nervousness)

However, the Wilcoxon test did not find a statistically significant difference between the word count among versions, not even when analysing separately the SA participants (p>.05). The Median (IQR) number of words spoken in each version were VRU = 286.0, VRF = 263.0, RLU = 245.0 and RLF = 270.0. The results among versions, divided by bins, did not present statistically significant results as shown in Table 5-4, neither did the between-subjects comparison between NA and SA (see Figure 5-11).

Table 5-4. Wilcoxon results for the difference in speaking time and word Count between versions separated by the Personal Report of Confidence as a Speaker bins, rows with p>.25 are omitted. (NA: Nervous; SA: Regular Nervousness)

	Speaking Time				Word Count			
	NA		SA		NA		SA	
	Z	p	Z	p	Z	p	Z	p
RLF – RLU			-2.023	.043			-2.023	.043
VRU – RLU			-1.483	.138			-1.214	.225
VRF – RLU			-1.826	.068	-1.604	.109	-1.826	.068

Section 5.5.5 Open Questions

Most (eight out of 10) participants answered *Q1a* positively; they had noted some similarities between the VR audiences and the real-life ones. However, only four of them mentioned that they had noticed the virtual audience's physical appearance to be the same as the real-life audience. One participant did not answer the question, and the other said both virtual audiences were still too abstract. Only two participants mentioned that the virtual audience made them feel nervous (*Q1b*). Most participants (seven out of 10) said the virtual audience did not make them feel nervous. P8, who noticed the physical resemblance between VRF and RLF, said, "*The VR audience were a bit like real ones, this helped me to concentrate more on my presentation. Also, it made me less tense for the real presentation*[RLF]," and P6 commented, "*Yes, The similar classroom and people, it makes a really nice environment for me, it's not making me feel embarrassed to laugh*". On the other hand, P5 did not feel the virtual audience was enough: "*I know the VR audience are trying to make a sense of they are real guys, but I think this can't affect me a lot, because I know they are not real people.*"

All participants felt that presenting to a real audience was quite different from presenting to the VR audiences, regardless of the face (*Q2a*). They mentioned that the micro-expressions during the presentation make more of a difference than the body movements; as P8 put it, "*Of course [*giving a presentation to a VR audience was different from giving a presentation to a real audience]. *For example, there will be eye contact when facing a real audience, which makes me more nervous than facing VR audience*". A similar feeling was echoed by P10: "*we can notice every audience's expression during a real audience's presentation.*" Still, eight of the 10 participants said (*Q2b*) that they were less nervous about presenting to a real audience after presenting to the VR one. However, P2, who scored the maximum in the PRCS, had a particularly intriguing reaction to the virtual audience. P2 complained that "*It affected me a lot because they could not bring me real feedback while giving a presentation*," but had previously mentioned, "*I am not interested to look at them while giving a presentation*."

Section 5.6 Discussion

Our analysis of the qualitative data suggests that using familiar faces does not seem to be enough to engage people with the audience in VR **(RQ1)**. While some people seemed to benefit from it (being less tense during the real-life presentation), still a percentage of the sample population did not engage with them. Our results are aligned with previous research in which some extremely anxious people just disregarded the audience (like our P2) [140]. The confident participants commented more on the movements of the virtual audience; however, they were unaffected by them. The SA (Regular Nervousness) participants (with a medium PRCS) are the ones who seemed to benefit the most from the familiar faces. They commented on their similarities and how real it felt, and how it helped them feel relaxed when presenting to the real audience after.

Our statistical analysis suggests that VR cannot yet elicit the same anxiety level as a real-life presentation, which could be desirable for different levels of difficulty, even when there is negative feedback about the quality of the presentation. According to their verbal responses, the participants did not mention body language as a problem with the feedback. This may indicate that the way body feedback was implemented was suitable or ideal. Though the participants did not correlate their speeches with the avatars' feedback (**RQ4a**), they did not mention it as lacking realism. What failed to generate the same anxiety, according to their answers, was the lack of facial expressions; thus, the perceived lack of facial realism affected them negatively in the feedback sense, but not in making them more nervous (**RQ4b**). Thus, to engage participants and improve realism, the following steps should focus on improving eye contact and other subtle facial expressions, highlighting the importance of works focusing on emotional expression [164].

The differences in body score and HR generated by VRF and VRU were not statistically relevant by themselves. However, since we could see a difference in them graphically, we decided to investigate this further. Using these two sets of data as input, we could classify to which virtual audience the participant was talking to. This indicates that even though the relation is not linear, there is a perceptible difference between presenting to VRF and VRU. Furthermore, the difference in ΔF and ΔU associated with the lower overall HR level of VRF suggests that familiar faces have a soothing effect during training and that they can aid in reducing anxiety for presentations in real-life (**RQ2**) on people with SA. They may feel more comfortable when facing familiar people or being in a virtual environment, which may relieve their anxiety.

Statistically, there were no clear differences between the word count or speaking time from people who experienced VRF and VRU. There was also no statistical difference between RLF and RLU. Unsurprisingly, there were differences between VR and RL. As pointed out by the participants, they felt that the differences between both VR and RL were quite distinct but did not pay much attention to the virtual audience. So, even though it is possible to detect speech differences between presenting in VR and in RL, under the current conditions it was not possible to detect differences between audiences (**RQ3**). Furthermore, again the results indicate that presenting to a virtual audience can be beneficial for the SA participants, given that they were able to keep their speech constant across audiences.

Even though our results between different groups look graphically different, the statistical tests did not reveal any clear difference. This might have been caused by the relatively small sample size, or because the participants who are not phobic, regardless of anxiety level, behave similarly. However, though such hypotheses are an interesting topic of investigation, such deep psychological analysis is beyond the scope of this research.

Section 5.7 Summary and Remarks

This research explored the effects of different face familiarities in a VR presentation simulator. We studied how presenting to a virtual audience with faces scanned from familiar people affected the participants' anxiety in VR and how it posteriorly affected their anxiety in a real-life presentation to the same people compared to unfamiliar ones. To accomplish this, we developed a VR presentation simulator with two conditions: a familiar audience with scanned from the researchers (VRF) and an unfamiliar one using stock, generic models (VRU). Later these participants presented to a real-life audience, comprised of familiar (RLF) and unfamiliar (RLU) audience members. We used the HR, Speech Time, and Word Count as a measure for anxiety during the presentation. We also analysed how long people had been in an inadequate position in a presentation, as a scoring mechanism. Later we used a questionnaire to ask for the participants' opinions about the experience and interaction with the virtual audience.

Our results suggest that familiar faces in VR have an overall soothing effect on people who have a medium level of Public Speaking Fear; however, those who have a high Public Speaking Fear or no Public Speaking Fear do not benefit from it in the short term. Furthermore, from the comments of the participants, we see that the most significant barrier for people to accept that the audience is actively listening to them is the facial expressions. Thus, regarding interaction simulation games, future developments to increase realism in presentation environments should focus on the faces of the virtual audience rather than on the bodies.

Chapter 6 Evaluating Enjoyment, Presence, and Emulator Sickness in VR Games Based on First- and Third-Person Viewing Perspective

Section 6.1 Introduction

The first idea that resembles what we now consider immersive VR was presented in 1965 by Ivan Sutherland [26] as an artificial world that interacted with and immersed all our senses. Due to technological limitations, VR technology was neglected by the masses. In recent years and with the release of affordable immersive VR technologies, VR rose from obscurity and is becoming mainstream. This rebirth left many companies, especially those in the entertainment industry, wanting a market share. However, the best way to implement VR and how to make people less likely to suffer from side effects, like Simulator Sickness [165], are still being researched.

The gaming industry has rapidly moved to leverage VR technologies in order to enhance gameplay experiences that can come with higher feelings of immersion and presence. With the wide variety of genres, formats, and styles, there is a need to know how to best increase gamers' enjoyment and, as such, research on player experience plays a big role in that industry [15].

A factor that is essential for enhancing gamers' experience is the sense of being there inside the game, also known as presence [11], [15], [166]. It is suggested that in games, emotions are stronger with a higher presence [11]. One factor that promotes presence and that is still underexplored in immersive VR is perspective [11], [12]. Perspective not only changes the sense of presence, but also may have an impact on the player's physiological reactions [11], [14]. Such physiological reactions might affect the player's enjoyment of a game. In video games, two perspectives are generally available to the player: 1PP and 3PP [167].

In this study we intend to analyse how the different perspectives in a VR game can influence presence, enjoyment, and Simulator Sickness. We compare gamers playing a VR game in 1PP and 3PP. In addition, we use a traditional display (a TV) as the baseline in our experiment.

Section 6.2 Related Work

Section 6.2.1 Simulator Sickness

Simulator Sickness is a recurring problem in VR, so much so that there are different strategies to overcome this problem. It has been claimed that, by adding movement to VR, it will decrease the likelihood for people to get Simulator Sickness [165]. However, the same researchers conclude that such strategy might not be the best option. One other strategy is demonstrated in [168]: to add a small reticule in which the player can focus their gaze. We hypothesize that this fixed point in which the player can look as a reference point could also be implemented as an avatar in a 3PP game.

In another paper, Medina et al. [14] have presented a study in which they looked at the performance of volunteers when walking in a device called VirtuSphere. This device is a "human size VR 'hamster ball'" used for semi-natural walking. In their study, they verified that users who engage in a virtual environment using 1PP usually suffer from stronger Simulator Sickness effects when compared to those who engaged in the same environment using 3PP.

Section 6.2.2 Immersion and Presence

In [11] the differences between 1PP and 3PP are also described but under a different approach. In this work, Simulator Sickness is not evaluated, nor is any kind of immersive virtual reality. The researchers instead focused on the differences in the feeling of presence when playing in the two different perspectives. In order to perform their experiment, the authors of [11] used a regular PC display and the RPG "Elder Scrolls 3: Morrowind". They found that even though 1PP generated a greater feeling of presence, gamers' experiences in 3PP were more pleasant. This result agrees with those in [15], which also indicated that people were more immersed in the gameplay when viewing the game world through the eyes of the character (that is, in 1PP). According to this research, there is a correlation between viewing perspective with perceived immersion. However, this aspect has not been explored in detail. In a meta-analysis of 83 studies about immersion presented in [12], only two looked at immersion from different perspectives and did not include VR technologies.

Lim and Reeves [169] report a study where players were given a choice of avatar to explore their feelings of presence. It was found that their feelings of presence were on average the same in 1PP and 3PP. In [166], the authors not only studied the differences of perspectives, but did so in an immersive virtual environment. Their results suggested that, regardless of perspective, it was possible to have the feeling of “being there” (i.e., presence). They also investigated components of presence, the sense of ownership, and self-location, and found that 1PP is superior to 3PP. These results contrast with those of [170], but match those presented in [171]. It is hypothesized that the measurements in [170] might lack sensitivity because they are summarized into a single item, whereas the authors of [166] used multiple items to assess each factor.

Salamin et al. [167], [172], took the analysis of 1PP and 3PP in VR using an innovative approach, one in which cameras were strapped in a way that the VR world was the real world viewed from different lenses. The avatars no longer were digital but were the volunteers themselves. For the 3PP the authors strapped a camera to the participants in a way that they would see themselves as avatars. For the 1PP the users wore a kind of mask with a camera attached. In both cases, the participants would see the world only through a head-mounted display (HMD). They would perform a ball-catching activity in the environment. Surprisingly, 3PP facilitated some tasks.

Another experiment involving sports and perspectives in VR was reported in [173], [174]. This setup differs from the others regarding mainly the equipment used for the VE. Instead of using an HMD, the authors used a CAVE-like system. For this experiment, they compared 1PP, 3PP, and 3PP+ (3PP with visual guidance) and involved participants in performing free throws in the three conditions. The results suggest that, for basketball training, it is possible that the 3PP+ is better for novices.

Our work builds on these prior works. To the best of our knowledge, since [14], the effects of perspective in Simulator Sickness have not been studied in VR. With much improvement been made to HMD in the past few years, further investigation is needed given that most papers presented in this session evaluated users' performance in different perspectives, but without giving too much attention to the general enjoyment when games are played in VR under different perspectives. Given the increasing popularity of VR HMD, our work is timely and can inform the design of VR games.

Section 6.3 Study Design

In order to evaluate game enjoyment, immersion, and presence under different perspectives, our study was designed in the following manner. Volunteers would play a selected game under 3 conditions 1PP-VR, 3PP-VR, and a baseline 3PP-CD (3PP conventional display).

Since we wanted to investigate three different factors, we decided to use three questionnaires that analyse different levels of Simulator Sickness, immersion and enjoyment. The first one is the SSQ (simulator sickness questionnaire) [106], [175]. The second one is the GEQ (game experience questionnaire) [64]. For this questionnaire we used only the iGEQ (in-game questionnaire) repeated at every interval, and the PGQ (post-game questionnaire). The third questionnaire we used is the IEQ [176]. We used an adapted version of the IEQ because we applied it in conjunction with the iGEQ. Some of the questions were similar; others were rephrased so they would fit the context of the other questionnaire, and some just did not fit the context of the played game. Overall, the participants answered a 43-question questionnaire with five-point Likert-scale items, and 15 questions from the SSQ.

We recruited volunteers from our university, regardless of their course of study. We had a total of nine volunteers. All participants were male, with an average age of 24.4±4.2 years, with a median age of 23 years old, ranging between 19 and 32. All volunteers had normal or corrected-to-normal vision. None declared any history of mental illness of any kind.

Of the sample, 88.9% of volunteers had experience with the chosen game before. However, none had placed it on a PC emulator, nor on any kind of VR device. However, 66.7% of them had already had experience playing other games in VR, and 77.8% played "regular" games at least once a week. Of that 77.8%, 71.4% played at least 3 times a week.

We used an Oculus Rift CV1 as our HMD shown in Figure 6-1. Our PC was one with 16GB RAM, Intel Core i7-7700k CPU @ 4.20GHz, and GeForce GTX 1080 Ti. As a controller, we used a Betop Pandora 2.4G Wireless gamepad, also present in Figure 6-1. We used it instead of the traditional Rift Touch Controller because this one could be used both with an HMD and without it.

Figure 6-1 Oculus Rift CV1 and Betop gamepad both equipment used in the experiment.

Section 6.3.1 Game Choice

We researched existing games that could be played in both perspectives. Games in which the mechanics would not change considerably independent of point-of-view or kind of display. And since we intend to further our studies, we looked for games that, in our perception, do not openly appeal to a specific gender (future studies will corroborate or dismantle our hypothesis).

For this reason, we focused on Nintendo games that, according to different sources, had attracted a diverse audience, including women and the elderly [177]–[179].

We then researched the possibility of playing Nintendo games in VR. And, within our time frame, we found it playing Nintendo games in VR is possible with the emulator Dolphin VR [180], which is one of the software we used, in its 5.0 version, whose print screen is presented in Figure 6-2.

We then decided to use Mario Kart Wii [181] for this first experiment. The races have roughly the same time length, so it would not affect our study. The game has some easily observable stimuli that can be used for future experiments.

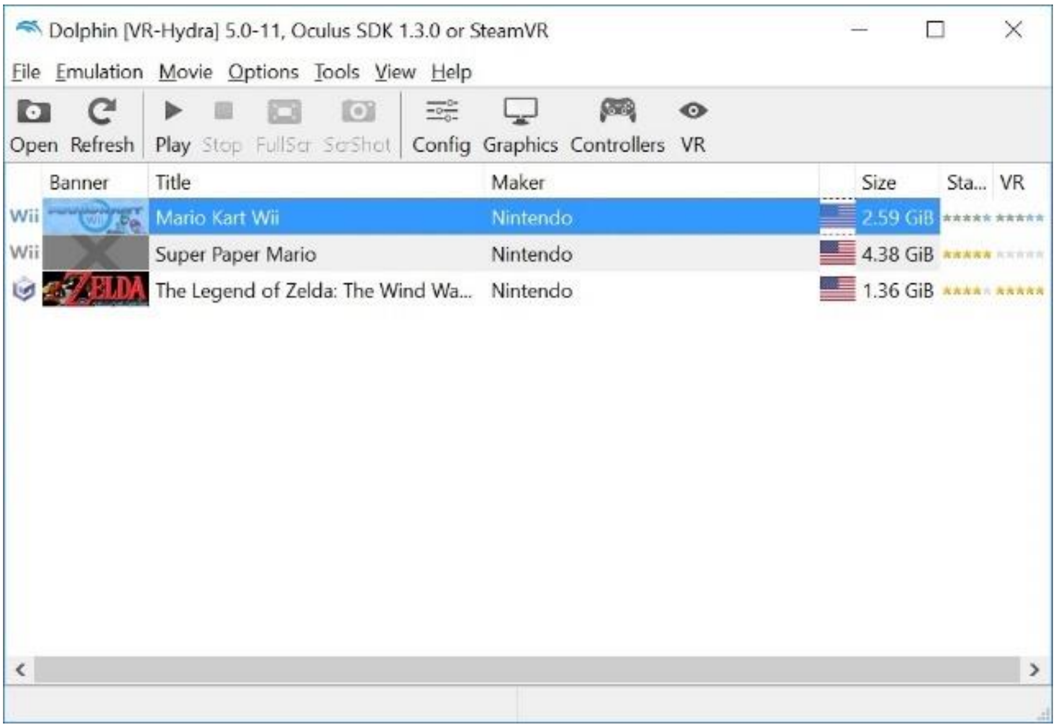

Figure 6-2 Dolphin VR 5.0 main screen, with some of the games we tested.

And, as desired, the game mechanics remain the same regardless of perspective or display. We assume that, since the game does not present a linear story, having played the game before does not have big influence on the subsequent races, in the sense that no surprise or plot point would be "spoiled".

Section 6.3.2 The Game

As stated, the game we chose for the experiment is Mario Kart Wii. It is a racing game developed by Nintendo. The game traditionally is in 3PP. The main goal is to arrive first to the finish line. In the circuit chosen by us, all races have three laps and the circuit takes around 11 minutes in the 150cc difficult setting.

During the race the players can acquire different items, those items vary in power and usefulness. When the player is in a better position the power and usefulness of the item will be lower.

We asked all players to play as Mario in the 150cc difficulty setting, using the standard kart shown in Figure 6-3, because this is one of the most balanced options available. All players played in the Mushroom Cup against 11 computer adversaries. They all used the same configuration for all three cases.

Figure 6-3 Print Screen of the kart selection screen in Mario Kart, showing the selected kart for the experiment.

For the game to be playable in 1PP in Dolphin VR a few adjustments had to be made. The (Heads-up Display) HUD Distance was set to 0.01 meters; the HUD 3D Closer was set to 1; camera Forward was defined as 600 meters; and camera Pitch was set to 15 Degrees Up. For the 3PP, all parameters were kept the same, except camera Forward and camera Pitch, which were both set to 0; the differences are visible in Figure 6-4.

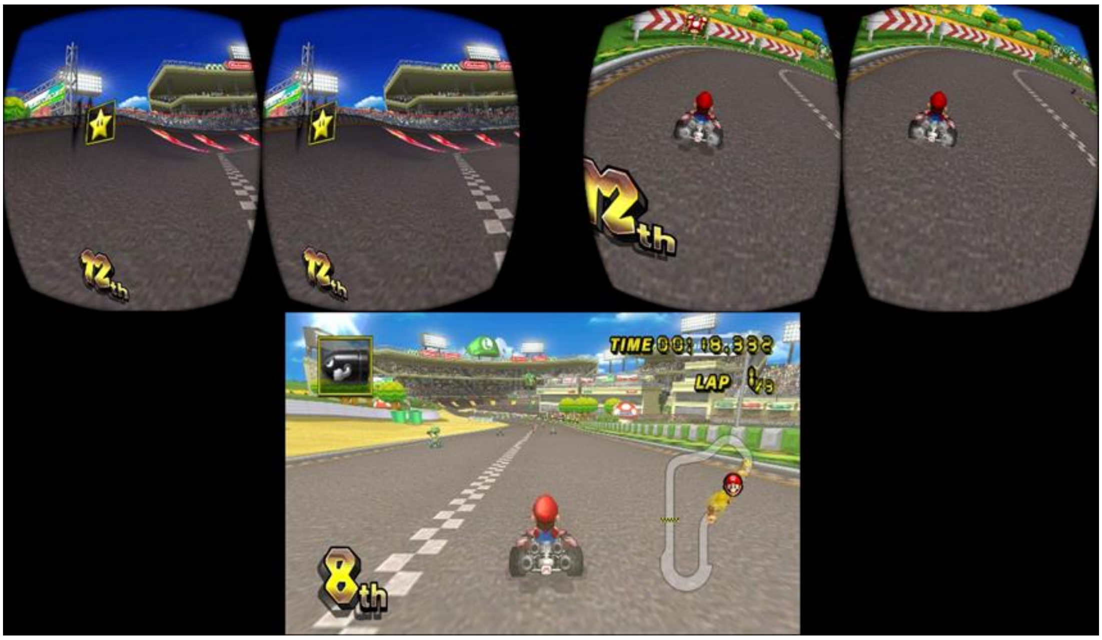

Figure 6-4 Different perspectives, upper-left 1PP VR, upper-right 3PP VR, lower-middle 3PP-CD

The controllers were adapted to match the Betop Controller configuration. Buttons A, B, X, and Y from the emulator were set to match their nominal values on the controller. Control Stick was set to match the left analog in Betop Controller. Buttons L, R, and Z were set as LT, RT, and LB, respectively (see Figure 6-5).

Figure 6-5 Controller and the button layout during game, not all buttons were used.

Section 6.3.3 Procedure

On arrival, all players were assigned to one order in which the perspectives would be played. They were asked to fill a questionnaire with their personal information, such as name age, frequency in which they play games, if they have ever played VR, and whether they have ever played Mario Kart.

They were then introduced to the controller and were asked to sit in front of the monitor and play the circuit once (see Figure 6-6 and Figure 6-7), with the aim that they would get used to it. After this brief introduction, they were asked if they would like to stop for 5 minutes. None of them wanted to stop.

Figure 6-6 Volunteer during controller and circuit adaptation.

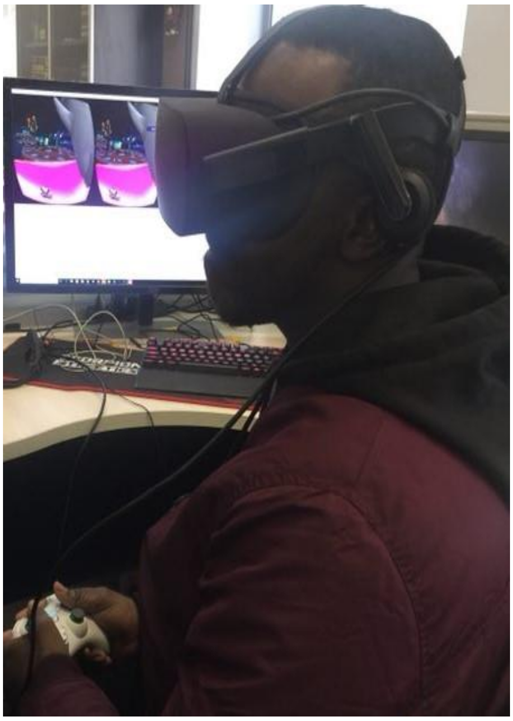

Figure 6-7 A volunteer playing 1PP VR.

After each step the players were asked to fill the questionnaires. This process should take around 10 minutes at least. If the questionnaire was answered faster, we asked the participants to wait a little before going back to the game. This was done not only to have the most precise report we could but also to give the players a break from VR. We wanted to avoid causing Simulator Sickness, or to at least reduce it.

The participants each took around 75 minutes to complete all the processes. At the end they were offered water, fruits, and nuts and were asked to stay in the lab on the couch for a few minutes in case they had any side effects. Before every game session the participants were informed that if they felt any kind of sickness and wanted to stop, they could do so at any moment.

Section 6.4 Results

After the experiment, we had 27 filled questionnaires, three for each participant representing a different perspective that the participant had played in the game in.

We calculated the scores of each questionnaire as explained in their respective sources [64], [106], [175], [176], considering our adaptations. Pearson`s correlations were calculated using the IBM SPSS Statistics Viewer software [182].

In our data, there was a correlation between how many times someone plays games in a week and their positive experience in VR in both 3PP ($r=.703$, $p<0.04$) and 1PP ($r=.768$, $p<0.02$) conditions. This same correlation is not present in 3PP-CD. The authors hypothesize that this might be because VR Mario Kart is a new experience, whereas 3PP-CD Mario Kart is very “common”. We wish to investigate this correlation further and see if participants frequently exposed to the same games in VR present this same pattern of correlations.

Table 6-1: Correlation between Flow in 3PP VR and immersion.

Correlations					
		Flow-3PP VR	Immersion-3PP CD	Immersion-1PP VR	Immersion-3PP VR
Flow-3PP VR	Pearson Correlation	1	.696*	.299	.912**
	P		.037	.435	.001
	N		9	9	9
Immersion-3PP CD	Pearson Correlation	.696*	1	-.042	.832**
	P	.037		.914	.005
	N	9	9	9	9
Immersion-1PP VR	Pearson Correlation	.299	-.042	1	.413
	P	.435	.914		.270
	N	9	9	9	9

Another interesting correlation worth mentioning is that, whereas flow in 1PP VR has a correlation with positive affect in 1PP VR ($r=.818$, $p<0.01$), and flow 3PP VR also presents this correlation with positive affect in 3PP VR ($r=.797$, $p<0.02$), only flow 3PP VR presents a correlation with immersion in both 3PP VR ($r=.912$, $p=0.001$) and 3PP-CD ($r=.696$, $p<0.05$) (for more information see Table 6-1). Since we were unable to generate a 1PP-CD (1PP conventional display), we can only suppose that this correlation is due to the same perspective being applied in both kinds of display, especially because competence in 3PP-CD correlates to competence in 3PP VR ($r=.783$, $p<0.02$).

Regarding the SSQ, our results corroborate those of [14], in that, on average, people felt more nausea with 1PP VR than with 3PP VR, scoring on average 7±7.7 on 1PP VR and 5.3±4.4 on 3PP VR. Furthermore, 33.3% of participants scored higher in 3PP VR and, of those, only two scored meaningfully higher in 3PP VR (for the graphical information of these results refer to Figure 6-8). However, in the oculo-motor session of the questionnaire, there was virtually no difference in average scores: 1.7±1.5 for 1PP VR and 1.4±1.5 for 3PP VR.

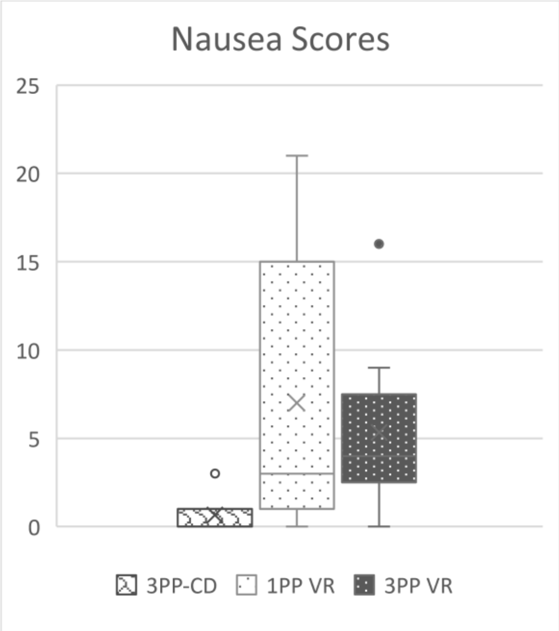

Figure 6-8 Box plot of the nausea scores.

The maximum possible scores were 27 for nausea and 18 for oculo-motor. We suppose that this average difference might be caused by the possibility of looking at a fixed point in 3PP VR (Mario's head). This could be validated by the hypothesis from [168] regarding their fixed point.

When it comes to the overall immersion questionnaire, from a total of 77 possible points, 3PP VR scored 44.3±14.3, being a bit less immersive than 1PP VR, which scored 45.7±12.8. Both surpassed 3PP-CD (39.4±11.0) in immersion. Only one participant scored higher in the overall immersion of the 3PP-CD.

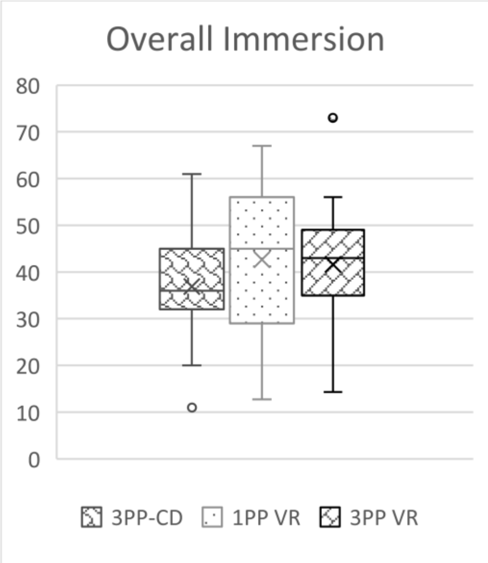

Figure 6-9 Box plot of the overall immersion scores.

We hypothesized that the greater immersion would lead to greater enjoyment. However, this was not clearly apparent in our data (see Figure 6-9). The biggest enjoyment score was the one of the 3PP-CD: 3.0±1.1 versus 2.3±0.6 for 1PP VR and 2.5±0.8 for 3PP VR. 3PP-CD had the lowest average transport 2.1±0.8 and temporal dissociation 0.1±5, versus 3.1±1.0 and 0.9±0.7, respectively, from 1PP VR. What this might indicate is that enjoyment is not so closely related to presence.

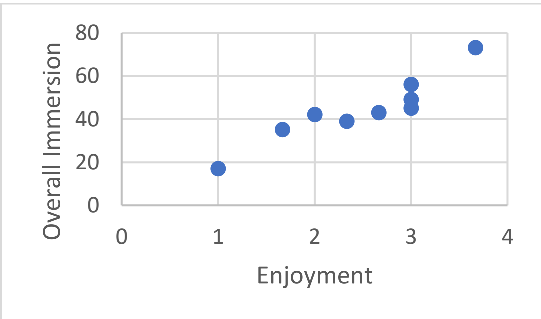

Figure 6-10 immersion vs enjoyment graph in 3PP VR

Nevertheless, in 3PP VR and 3PP-CD there seems to be a correlation between enjoyment and immersion, as presented in Figure 6-10. These seemly divergent results lead us to believe that greater enjoyment might lead to a greater sense of presence, but not the other way around.

Even though the immersion is clearly superior in both VR methods and the sickness tests show better results for 3PP VR than for 1PP VR, no clear preference was found. Exactly one-third of the participants preferred each version, as summarized in Figure 6-11.

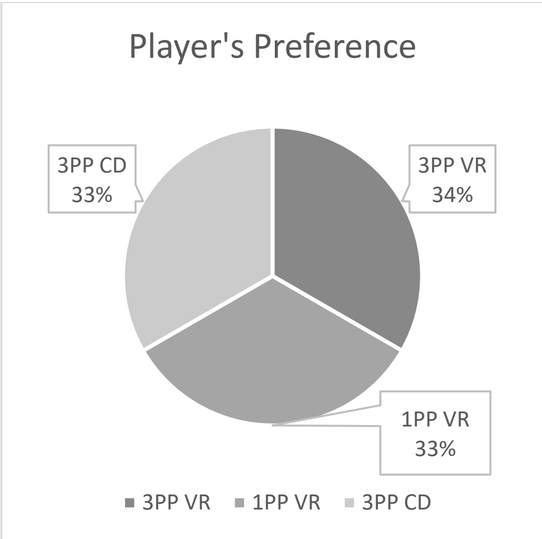

Figure 6-11 Player's preferred perspective and display.

Section 6.5 Summary and Remarks

In this research, we presented a short review of the studies comparing 1PP and 3PP. We executed an experiment to evaluate whether previous studies still held true with the current VR HMD technology. We analysed whether there is a difference in the levels of enjoyment when people played 1PP VR and 3PP VR, as well as compare those data to the level of immersion and Simulator Sickness.

For the time being, we did not find a clearly superior technology, since they have their pros and cons. 3PP VR is less likely to make people VR sick when compared to 1PP VR. However, it is not as immersive, even though the difference is not big. We could not define a clear preference as well, since this data is completely divided.

Some limitations of our study are the fact that we do not have a 1PP-CD to make comparisons between different displays in this perspective, which could shed some light on the results that we have. We did not have female volunteers or a variance in the age groups, making our results less generalizable. Our data sample was relatively small, which makes it difficult to come up with reliable conclusions.

Now that we have formulated some hypotheses based on our data, we intend to continue this study to solve the mentioned problems, be able to come to conclusions with greater statistical relevance, and validate or disprove our newly formulated hypotheses. Besides, our current results are in general in agreement with most similar studies.

As far as we know, this is the first research to compare a fully functional game under different perspectives using the current HMD technology.

Chapter 7 Analysing Subjective Feelings in Virtual Reality Games Based on First- and Third-Person Perspective using EEG and Subjective Metrics

Section 7.1 Introduction

Games can serve many functions beyond entertainment, such as teaching general and specific topics in health education and skill acquisition [183][79], promoting exercising and an active lifestyle [184], and helping people to regulate their emotions [185]. It can be said that virtually everybody has played a game in his or her lifetime [186]. Among games, videogames (which hereafter will be referred to simply as games) have become a mainstream activity, a popular way to spend leisure time [80], and a significant part of the entertainment economy. The increasing popularity of games is reflected in an expanding field of research [45], [83], [187], including the demand for researching what makes a game give a good experience for players [15].

While games have existed for decades, the gaming medium that has been attracting attention is virtual reality, especially VR that makes use of HMD [69]. VR technology has recently moved out of research and commercial labs and is now marketed to the masses. Even though commercially viable VR is a relatively new development, it has already been researched for different purposes and has the potential to enhance several experiences, such as increasing the feeling of presence during training scenarios [188], increasing knowledge retention [130], and facilitating meditation exercises [189]. Some studies have shown that such technology can bring a greater sense of presence [12], [131], a feeling that is often emphasized and which is commonly described as the feeling of "being there" [12]. It is believed that such a feeling might, in turn, bring higher engagement levels [11].

One factor of electronic gaming (hereafter simply gaming) that is not often studied is the difference in viewing perspective—that is, the difference between 1PP and 3PP (short for first- and third-person perspective, respectively) [69]. Despite the limited attention, it is an important topic of analysis because, among other factors, it can influence both the sense of presence, Simulator Sickness, and enjoyment, especially in emergent gaming technologies like VR [11], [14], [69], [71]. Studies demonstrate that 3PP can bring lower levels of Simulator Sickness to players immersed in a game [14], [69], whereas 1PP has been shown to bring a greater sense of immersion when compared to 3PP in an RPG game [11], [15].

Understanding which factors help people transport to and immerse themselves in the VR environment is useful to further leverage the benefits of this technology. Due to its novelty, several researchers have assessed various aspects of playing games in VR using subjective measurements [69], [83], [188], [190], [191]. However, compared to the amount of research using subjective measurements, psychophysiological measurements are rather uncommon, underexplored, and underutilized.

To investigate how different perspectives affect gaming in more depth, we used psychophysiological metrics. Psychophysiological measures such as EEG, electrocardiogram (ECG), and electromyogram (EMG) have received some attention from game researchers who aim to capture not only subjective feelings from players but also their bodily responses to those stimuli [65], [71], [72].

Psychophysiological data is an especially interesting way to understand the influences that different media have on the gaming experience itself. An example dealing with viewing perspective is a study that evaluates whether viewing perspective influences emotion and presence in gameplay [11] where the authors used muscle detection and eye-tracking to assist their evaluation. They did not find a clear pattern of relationships in their work. Because the evaluation is not done in VR, it is not certain if their findings would apply to this technology.

In this research, we use psychophysiological metrics to capture users' emotions through existing metrics during gameplay and explore their association with subjective, self-reported measures. We analyse how those metrics are affected by differences in viewing perspective in a VR game. We also look for correlations and patterns in these metrics, for example, to investigate if they can tell which version (or perspective) the player is enjoying the most.

Section 7.2 Theoretic Foundation

Section 7.2.1 Subjective Metrics

Subjective metrics rely mainly on questionnaires and open-ended questions that allow participants to self-report their feelings and perceptions of an activity or experience—in our case gameplay. Questionnaires are a useful research method because they are simple to use and also provide standardization. Naturally, there are drawbacks to questionnaires. For example, the questions should be carefully worded and phrased so that participants do not get confused or have a biased interpretation [192]. On the other hand, questionnaires can be used to evaluate the most diverse feelings and emotions people have. For example, there is the subjective happiness scale (SHS), the psychological general well-being index (PGWBI), the profile of mood state (POMS), among others [193]. In this work, we are interested in metrics for engagement and enjoyment in games.

Table 7-1 subjective questionnaires and their components

Questionnaire	Components
IEQ	Emotional Involvement Cognitive Involvement Real World Dissociation Challenge Control
GEQ	Competence Flow Challenge immersion Negative Affect Positive Affect
SSQ	Nausea Oculo-motor Issues

Often-used questionnaires when studying games are the Immersive Experience Questionnaire (IEQ) [176] and the Game Experience Questionnaire (GEQ) [64]. These questionnaires aim to understand different aspects of gameplay, such as emotional involvement, focus, and presence (see Table 7-1). Since it is currently regarded as impossible to measure discomfort in an objective manner, we need to use subjective measurements to understand discomfort, such as nausea, fatigue, and eye strain. The SSQ [106] is commonly used to measure levels of discomfort.

Section 7.2.2 Brainwaves and Metrics for Emotions

Neurons are constantly active, emitting small amounts of electromagnetic waves. These waves can be processed to identify well-known frequency bands contained within each one of these waves [194]. There are five well-known frequency bands (Delta, Theta, Alpha, Beta, and Gamma), and each of these frequencies is regularly associated with a number of mental states (see Table 7-2). The frequency bands vary slightly between sources regarding the range of each wave [45], [89], [189], [194]–[196]. Despite these small variations, there is agreement regarding the central frequency components. Game makers and psychologists are interested in understanding those waves [72] because they can be used not only to measure psychophysiological states, but also for indirectly analysing diseases such as depression and Alzheimer's disease [68], [195], and feelings of boredom [189], happiness, sadness, and fear [197] in a non-invasive, safe, and simple way.

Table 7-2 Types of Brainwave Signals

Wave Type	Frequency	Mental State
Delta	1-4 Hz	Deep sleep Unconscious
Theta	4-8 Hz	Creativity Dream sleep Drifting thoughts
Alpha	8-13 Hz	Relaxation Abstract thinking Calmness
Beta	13-30 Hz	Tension Excitement Stress
Gamma	30-50 Hz	Anxiety Nervous Stress

There has been a long interest in discovering brain patterns to characterize engagement, focus, and other emotions [198]–[200]. This is because these kinds of metrics can help workers stay attentive to their tasks [199], can help teachers understand the engagement of their students [194], or even make roads safer by assessing drivers' vigilance and state of alertness [201].

Shin et al. [68] proposed a measurement for engagement that is calculated using only Beta and Theta waves. A metric that has been used by different researchers (e.g. [197], [202]) is that proposed by [199]. This metric calculates a person's level of engagement in a task and takes into consideration three kinds of waves: Alpha, Beta, and Theta. One advantage of those frequencies is that they are less likely to contain artifacts (i.e., noise in EEG data context) from eye blinks and muscle movements because the former usually is present below 4 Hz and the latter over 30 Hz [200].

Even though engagement is one important factor that can be analysed from Beta and Theta waves, they can also be used to determine other emotional states. Other works use Alpha and Beta to identify arousal and valence [197], [200]. "Arousal-Valence" is a way to "dissect" emotions [88]. It states that emotions lay on a plane (see Table 7-1), in which arousal is one axis and valence is the other; arousal is defined as how energetic an emotion is, whereas valence is used to determine whether the emotion is "positive" or "negative" [88].

Despite psychophysiological data providing precise and objective information, its interpretation still relies on subjective interpretation. Thus, in our work, we use not only questionnaires to understand how the participants are subjectively feeling. We also use psychophysiological metrics for comparison.

Based on the above research, we chose metrics that use Alpha, Beta, and Theta as parameters, since those frequencies are less likely to contain muscle and blink artifacts that affect the accuracy of the EEG readings. Thus, we could have a higher degree of confidence in the data. We use these metrics to determine the levels of Arousal-Valence [197] and Engagement [199] because previous research has shown promising results.

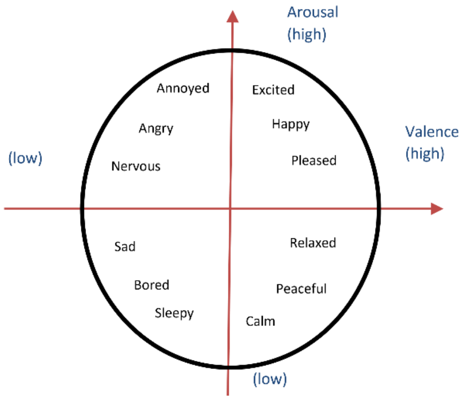

Figure 7-1 Emotional states and their positions in the valence/arousal plane (Image adapted from [89]), where arousal represents how energetic an emotion is and valence shows whether the emotion is "positive" or "negative" [89].

Section 7.2.3 EEG Data from Consumer Headsets

The commercialization of consumer EEG headsets, also known as consumer Brain-Computer Interfaces or EEG headbands, has facilitated research using psychophysiological data for activities such as meditation [189], lecture attendance [202], and relaxation. Although not as precise as their medical equivalents, they do present good precision [203] and have been found to be sufficient for non-critical situations [204]. In their simplest version, they come equipped with a reference electrode and a dry electrode that is connected to the frontal cortex (FP1) [194], [202]. Some consumer EEG headsets that have been used for research are MindWave Mobile [194], [202], MUSE EEG [203], and EMOTIV EPOC [197].

Shin et al. [68] developed an EEG headband dubbed H-Brain. However, it is not commercially available. Their headband compares to MUSE; they are both dry electrode bands; both use a central electrode on the forehead as reference; and two of their other 4 electrodes are on AF7 and AF8 [68], [203]. However, MUSE also has electrodes on TP9 and TP10 (see Figure 7-2), while H-Brain focuses only on the frontal lobes. Given its capabilities, we felt that MUSE was a good choice for this research.

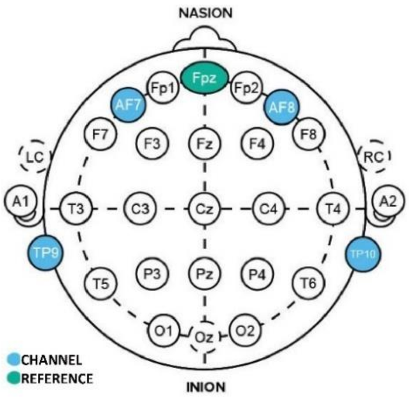

Figure 7-2 The contact points (image adapted from [66]).

Section 7.3 Related work

Section 7.3.1 Viewing Perspective and Presence in VE

There is a belief that VE should be built to increase users' feeling of presence. It is possible to identify three factors that constitute presence: the sense of self-location, the sense of agency, and the sense of ownership of the virtual body [166]. These factors connect deeply presence and embodiment.

Presence is an important topic of investigation in the context of games because the degree of perceived presence might lead to a better gaming experience [11] (or not [69]). Due to its importance, there have been several studies exploring how different factors can influence users' feeling of immersion [12]. These studies have evaluated, among other things, the screen size, image quality, and update rate. However, little research has been done focusing on the differences of viewing perspective in gaming and how it influences both presence and enjoyment, even though viewing perspective is considered to have a great influence on the feeling of presence [166].

One work that has looked perspective and presence is [15]. In their study, they have collected subjective data that suggest that there is a greater sense of presence when people play in 1PP compared to playing in 3PP. They had volunteers play an RPG with the possibility of switching between first- and third-person views. After playing the game, their volunteers would complete an IEQ. Even though they used only a common display for their analysis, one could assume that it is likely that viewing perspective will have a similar influence on different kinds of games played in VR environments.

Using a very similar setup, Kalinen et al. [11] also observed a greater sense of presence in participants who played an RPG in 1PP when compared to those that played in 3PP. However, they did not use only questionnaires for their investigation but also psychophysiological data. Specifically, they used eye movements and facial EMG to observe attention and emotions; however, they could not see a clear correlation between these two metrics and presence.

A recent work that involves both the feelings of presence and a physiological analysis is that of Debarba et al. [70]. In this research, while their participants were shown the floor close to their feet falling down into a pit, the researchers collected participants' galvanic skin responses. Later, participants answered a subjective questionnaire. The volunteers' responses showed a greater level of fear in 1PP.

Other works like [14], [69] present the idea that even though 3PP is not as immersive as 1PP, it has some advantages, such as volunteers presenting a lower level of Simulator Sickness. In both of these studies, volunteers wore an HMD and were presented with the two different perspectives. It is important to note that in [14], participants reported having a better experience in 1PP even though they suffered from greater nausea. The researchers believe that this is because 1PP gives a stronger sense of presence, which might increase engagement.

Interestingly, when we look at the works of Salamin et al. [167], [172], we can see that if we could choose our point of view in Real Life (RL) it might have a positive effect. Their results bring the possibility of changing viewing perspective in RL as it would be in a VE. Their findings suggest that maybe it is possible that 3PP can make spatial location-based tasks easier than 1PP does.

Since it has been suggested by a number of studies that 1PP provides stronger feelings of presence than 3PP, and it has been hypothesized that such feeling might increase engagement, we believe that having psychophysiological metrics would help clarify whether engagement is indeed affected by viewing perspective.

Section 7.3.2 EEG Analysis in Games

Past research has found differences in brainwave patterns when people are playing games or simply resting [186]. In [204], the researchers evaluated brain activity from children aged 8-12 using an Emotiv EPOC+ EEG headset. In this study, they found that during gameplay there could be a significant difference in wave patterns when performing certain game mechanics. This implies that further investigation might reveal even more diverse patterns.

When investigating games using EEG, there is some research that focuses on identifying differences in the use of different controllers [45]. In this study, participants played Resident Evil 4 with a PS2 controller and with a Nintendo Wiimote. They used a Biosemi ActiveTwo system to capture EEG data and found evidence that suggests that more intuitive gaming experiences (using a familiar controller and interaction mechanics) might require less brainpower.

In [205], the post-game effect on sleep was studied. They found that there is a greater sleep latency when subjects play games before going to bed, and some games increase heart rate and decrease Theta wave power. Results from [186] agree with those of [205] in that both PSD (Power Spectrum Density) decreased Theta when the participants were playing a game. What is unexpected is that their work [186] did not find an overall increase in Beta during gaming, which went against their expectations since Beta is associated with anxiety and focus, which they assumed was needed for playing games.

In summary, the above research shows that playing a game can affect brainwave patterns during the activity and even after players stop the game. As stated earlier, there is little research on the effect of viewing perspectives on players' brainwave patterns, especially in the context of VR gaming.

Section 7.3.3 EEG Analysis in VR

In a very recent work, [131] investigated possible differences in learning with a VR-HMD and a regular 2D display. In their experiment, participants performed some actions in a virtual lab where they studied university-level biology. In order to assess learning, presence, and workload, the researchers used a test, a survey, and an EEG system. For their EEG readings, they used the channels C3-C4, Cz-PO, F3-Cz, Fz-C3, and Fz-PO. Their results show that the volunteers had a greater feeling of presence with the VR-HMD; however, according to their EEG data, they scored lower in the tests and their workload was higher in this condition.

In [206], researchers investigated the difference in brain activity between immersive scenarios and desktop applications. In their study, they compared a Single-Wall-VR system with a computer monitor. In the Single-Wall-VR system the participants saw a projection (often three-dimensional) on a wall, normally with a 1:1 ratio between projected objects and their real counterparts. Their volunteers were divided into two groups (immersive and desktop) and had the task of crossing a maze environment as fast and accurately as possible. No significant differences were found in speed and accuracy between the two groups. However, like in [131], participants also had a greater subjective sense of presence when experiencing the more immersive maze. In addition, on the psychophysiological front of their research, they investigated Alpha waves, which presented a considerable difference between the two scenarios in the parietal brain areas.

A somewhat similar study to [206] has been reported in [207]. In this study, the researchers used a wall-like VR system. However, they did not compare the differences between VR and non-VR but, instead, compared Theta wave differences between two VR-WALL modalities: 2D and 3D. Their research found the participants had a greater subjective feeling of presence when exposed to a 3D environment as opposed to a 2D one. In addition, Theta waves were stronger in the 3D condition. This is of interest to us because Theta waves are present in the Engagement metric.

One of the first studies to use a consumer EEG headset for research analysed both the Alpha and Theta bands [208]. Like [206], [207] this experiment used as base for comparison a desktop and a VR-WALL to check for psychophysiological markers. They did find greater activation of the Theta and Alpha bands in some regions of the brain.

The above research sheds some light on the use of EEG data for VR environments but does not explore gaming scenarios specifically. Based on the above research, there is a difference between images shown in 2D and 3D for a wall VR system. Although this is somewhat peripherally related to viewing perspectives, their findings cannot be translated because there are more elements involved in how users perceive things in 1PP and 3PP. Our motivation to explore the effect of 1PP and 3PP on users' perceived level of immersion and engagement comes from this lack of research in the area.

We next describe the study design first and then the results of the study and a discussion based on these results.

Section 7.4 Study 1 Design

To evaluate the possible differences in psychophysiological response between VR / Non-VR gaming and 1PP / 3PP, we devised the following study. Volunteers would play a game under the following conditions: First-Person Perspective VR (1PP-VR), Third-Person Perspective VR (3PP-VR), and a baseline Third-Person Perspective Common Display (3PP-CD).

Participants wore a commercial EEG headset while they played the three versions of the game, and also answered our questionnaire, which contained a combination of three slightly-modified subjective Likert-scale questionnaires regarding gameplay experience, IEQ, GEQ and SSQ, following the steps laid out in [69].

Section 7.4.1 Volunteers

We recruited volunteers from a local university, regardless of their course of study. We had a total of 13 volunteers (one female) with an average age in years of 23.69±3.76, and an age range between 19 and 32 years. All volunteers were in good mental and physical health (that is, they were not aware that they had any neuropsychiatric disorders) and had normal or corrected-to-normal vision. All volunteers were briefed about the procedures and decided to participate in the study on a voluntary basis.

Based on the data from the pre-experiment survey, 84.6% of the participants had experience with the chosen game before but had never played it on a PC emulator nor in any kind of VR system. Of the participants, 53.8% of them had already had experience playing other games in VR and 69.2% played “regular” games at least once a week and, of these, 44.4% played at least three times a week.

Section 7.4.2 Apparatus

We used an Oculus Rift CV1 as our HMD, as it was one of the most popular VR systems available. It was connected to a desktop with 16GB RAM, an Intel Core i7-7700k CPU @ 4.20GHz, and a GeForce GTX 1080Ti dedicated GPU. We used a Betop Pandora 2.4G Wireless gamepad as the controller, which resembled any game controller. We used it instead of the traditional Rift Touch Controller because the Betop could be used both with an HMD and without it. Finally, we used the MUSE MU-02 from Interaxon Inc. as the EEG headset.

The MUSE collects EEG data at 256 Hz, uses machine learning to remove blinks from the data, and computes the power spectral density from each band on each channel using Fast Fourier Transform (FFT). Frequency transform is performed in 1 second windows, with the window stepping of 10ms. In the version we used, the output frequency ranged from 0 to 128 Hz [209]. The headset had an indicator to inform on the quality of the signal for each contact point. The quality was divided into “Good”, “Mediocre”, and “Bad”, which could be seen in real-time on the collection software and was marked on the collected data.

Section 7.4.3 Choice of Game

When choosing the game, we had a few parameters to consider. First, the game would have to have easily reproducible phases between players (and across plays). Second, the game should not have a deep storyline that could be "spoiled" after the first gameplay which we thought might interfere in the overall experience. Third, the game should have easily distinguishable events, such as clear victory points and obstacles. Fourth, we looked for games whose mechanics would not change considerably regardless of viewing perspective. And fifth, we looked for games that were not classified as gender-specific.

Given the above conditions, we focused on Nintendo games, which tend to have a balanced number of players of each gender [177]. We then researched games that fit our other requirements and that could be emulated effectively in VR platforms using the Dolphin VR emulator (version 5.0). Ultimately, we found that Mario Kart Wii would meet all the 5 criteria and was our game of choice.

The main goal of the Mario Kart game is to cross the finish line before the other racers. The circuit we chose (Mushroom Cup) normally takes around 11 minutes to finish in the 150cc difficulty level and each race had three laps. During the races, the players could acquire different items, which would give them extra power, weapons, and so. Under all viewing perspectives, we asked each participant to play as Mario using the standard kart (one of the most balanced options available) to keep consistency. All games were played with the participant racing against 11 computer adversaries. The configuration followed that one used successfully in [69].

Section 7.4.4 Procedure

On arrival, each participant was assigned a specific order of viewing perspectives in which he or she would play the game. The orders had been formed through a Latin Square design to mitigate carry-over effects. The participants then filled a questionnaire to collect demographic and past gaming experience information, such as age, how often they played games, whether they had experience with VR or Mario Kart.

The participants sat on the chair in which they would play the game and a researcher came and explained how the EEG data collection worked, so the participants would be more comfortable wearing the device. After the explanation, the researcher positioned the sensor on the participants' heads, checked the signal, and repositioned the device if necessary.

In the next step, the participants were introduced to the controller and they played the circuit once in 3PP-CD, so they would get familiar both with the controller, the game, and the EEG headset. During this phase, the researcher observed the EEG readings to make sure that the signal quality was good. The participants were instructed to move the least they could, if possible, during gameplay.

After the introductory circuit, if the participant's first condition was a VR one, then he or she would have a researcher positioning and adjusting the VR-HMD so that the EEG headset would not move. The researcher would ask the volunteer to make sure he or she could see clearly while wearing the headset. The participant was asked to sit comfortably and would then receive the controller; this was then followed by the researcher guiding the participant on which difficulty level, character, kart, and circuit to choose. The EEG and a video of the races would then start recording. All the steps were the same for the 3PP-CD apart from the HMD not being placed on the participants' heads.

We required the participants to take 10-minute breaks between versions. Within these 10 minutes, they filled the subjective questionnaires. We wanted to ensure comparability between participants and greater accuracy in the questionnaires. That is, regardless of how fast they answered the questionnaires, they still had to wait for 10 minutes.

Overall, the experiment lasted around 75 minutes for each participant. Towards the end, they were offered refreshments such as water, fruits, and nuts and were asked to stay in the lab a few minutes to make sure they were not impaired or uncomfortable due to possible Simulator Sickness. Finally, they were informed that they could leave the study lab at any moment if they so desired and feel comfortable in doing so.

Section 7.5 Results

At the end of the experiment, we had a total of 39 answered questionnaires, three for each participant, each representing a different perspective. We had around 6.5 hours of EEG-monitoring data. However, we had to discard two EEG analysis of two participants, who moved excessively during their gameplay. One of these participants suffered from severe nausea and had to stop the experiment earlier. We ended up with a dataset from 11 volunteers to work with. We calculated the scores of the subjective questionnaires following the same procedure described in the original sources [64], [69], [106], [175], [176].

We separated the EEG data by viewing perspective. We then filtered out roughly 10% of the data in which the quality was marked as either "Mediocre" or "Bad" and later computed the engagement, arousal, and valence indices for each participant. All calculations were performed following the steps in [71] using the bandwidths according to the respective source literature [197], [199]. We then calculated the average of each of these indexes by participant and viewing perspective. We used IBM SPSS software (IBM Analytics) to calculate a within-subjects MANOVA, multiple ANOVAs, a Friedman Test, and a Spearman Correlation.

Section 7.5.1 Questionnaires

Because the scores for the questionnaires were normally distributed, we performed a repeated measure MANOVA to verify their relevance. Table 7-3 shows the results of the means of the most relevant variables. There are differences in the subjective feelings of enjoyment, Return to Reality (sub-scales of the GEQ [64]), and Transport (which indicates presence [176]). The MANOVA and the multiple ANOVAs also suggested they were significant. Enjoyment, for instance, was greater in 3PP-CD (see Table 7-3 and Table 7-4).

Another noteworthy result was that the score for Simulator Sickness in 3PP-CD (0.4±0.7) which was lower than the two VR versions. Also, participants on average felt sicker in 1PP-VR (6.2±6.5) than 3PP-VR (5.45±5).

Table 7-3 Mean and standard deviation of some of the analysed items

Measure	Mean	Std. Deviation
Enjoyment 3PP-CD	3.2	1.3
Enjoyment 1PP-VR	2.4	0.8
Enjoyment 3PP-VR	2.5	0.9
Transport 3PP-CD	2.0	0.9
Transport 1PP-VR	2.8	1.1
Transport 3PP-VR	2.7	1.1
Return to Reality 3PP-CD	1.3	0.4
Return to Reality 1PP-VR	2.0	0.5
Return to Reality 3PP-VR	2.6	0.8

Table 7-4 Results of the MANOVA tests

Measure	df	F	Sig.
Enjoyment	1.185	4.802	0.044
Transport	2	6.564	0.006
Return to reality	2	14.667	0.000

Section 7.5.2 EEG Metrics and Questionnaire Association

After calculating the scores for the EEG data, we observed that the indexes were not normally distributed, and then ran a Friedman Test as a result. At first glance, we believed we could observe some differences in the means of the EEG indexes (see Table 7-5). However, further analysis indicates that these differences were not significant (see Table 7-6).

Table 7-5 Mean and standard deviation of some of the analysed items

Measure	Mean	Std. Deviation
Engagement 3PP-CD	1.2	2.3
Engagement 1PP-VR	0.5	0.3
Engagement 3PP-VR	0.5	0.4
Arousal 3PP-CD	1.0	1.5
Arousal 1PP-VR	2.1	3.2
Arousal 3PP-VR	1.4	2.1
Valence 3PP-CD	1.5	3.1
Valence 1PP-VR	11.0 / 0.2*	35.7 / 0.5*
Valence 3PP-VR	1.2	5.1

We also investigated how the EEG metrics related to the subjective ones. We first analysed their correlations (see
Table 7-7). The most relevant correlation that we found was that of Engagement and Transport that were present clearly in both VR versions and arguably in the CD version.

Table 7-6 Results of the MANOVA tests

Measure	df	F	Sig.
Engagement	1.019	0.936	0.358
Arousal	2	0.768	0.477
Valence	1.020	0.844	0.382

Table 7-7 Spearman Correlation Between Transport and Engagement

		Transport		
		3PP-CD	1PP-VR	3PP-VR
Engagement	3PP-CD	r = -.169 p = 0.619	r = -.566 p = 0.069	r = -.567 p = 0.069
	1PP-VR	r = -.609 p = 0.047	r = -.767 p = 0.006	r = -.862 p = 0.001
	3PP-VR	r = -.568 p = 0.069	r = -.671 p = 0.024	r = -.654 p = 0.029

Further analyses showed that even though there was no strong linear correlation between Simulator Sickness scores and EEG metrics, they seemed to have some sort of relation. After looking closer into the data, we observed that all 8 volunteers whose highest Simulator Sickness score was above 4 (from a total of 27) presented their greatest level of Simulator Sickness matching their lowest valence levels. For the others, the lowest valence was associated with lower levels of perceived challenge.

Section 7.6 Discussion

A few factors might have influenced the discrepancy between subjective enjoyment scores and arousal-valence indexes. It is possible for example that the Simulator Sickness has had a negative impact on the arousal-valence indexes. Alternatively, there is significance, but it was not visible due to the sample size (though our sample size falls within the norm based on similar experiences by others in the field). Another possibility has to do with the examined frequency bands which have not changed much, and the feelings of enjoyment and arousal are only associated with other bands or are just completely subjective. It is possible that differences are more expressive on specific events, such as going through the finish line or falling down a cliff. The differences could be "watered down". When we analysed only the overall indexes, especially because differences are visible in our work that focuses on specific game events [71].

We had expected to find an increase in engagement and arousal because the literature shows cases of Alpha and Theta changing during gameplay [194]. This might indicate that the current metrics are not appropriate to compare perspectives in VR gaming. Therefore, the utilization of alternative metrics might show different outcomes. Nevertheless, we emphasize that it was important that we have experimented with these metrics because prior research (e.g., [194]) overlooked a component of the metrics, namely Beta.

In a deeper analysis, we observed that on average the results for engagement, arousal and valence are higher when the participants are neither in the first position nor below the fifth. In a further investigation, we observed that, although not linearly correlated, for most participants (8 out of 11), the valence is the highest in the versions they have rated as more challenging. We infer from this that in the first positions the player does not feel challenged enough, and in the last positions the player feels frustrated.

One of the participants mentioned after the tests when being escorted out of the premises that he "played the new version of Mario Kart every day, and it has a more difficult level. The races in this one [Mario Kart used in the experiment] were a piece of cake. The VR-mode[1PP-VR] is quite cool though". From his EEG results, we could see a clear discrepancy in arousal and valence between VR versions—1PP-VR had a positive response 10 times greater than 3PP-VR. We speculate that, since he is accustomed to the game, there were no "real challenges" to him in the 3PP; however, 1PP felt somewhat like a different experience altogether. This volunteer is the only one to finish all races in 1st place, and his data were a clear outlier in Valence 1PP-VR—though his data were consistent with the subjective report.

Interestingly, positive valence seems to be a good indicator of which version is the favourite of the volunteers. The version which has presented the highest level of valence has been chosen by seven of the volunteers. Of the other participants, two have their worst performance in the highest valence version; both are regular daily gamers and have self-declared themselves as “competitive”. The other two have had their engagement, arousal, and valence indexes similar in all versions, which suggests that maybe these volunteers chose their favourite versions without a clear parameter.

We have found it relevant that the engagement metric correlated with Transport and are somewhat surprised that the correlation is negative, since Engagement is thought to relate to presence [189] and transport is a sub-component of the latter [176].

Section 7.6.1 Limitations

As mentioned earlier, one of our limitations is our population sample. We had a very homogenous group and a single kind of game, which made it harder to extrapolate the results. However, this does not invalidate our study since it can be used to guide future analyses and inform on the selection of metrics to evaluate engagement in VR games.
Another limitation of our study is that we could not include a 1PP-CD due to technological limitations—the Dolphin emulator did not support this mode. However, this does not affect our experiment or results since we are mainly comparing 1PP and 3PP in VR gaming.

Finally, we have two factors inherent in games: winning/losing and previous gaming experience. We could not control these factors and had to rely on the players’ expertise. However, most of our volunteers presented consistent results.

Section 7.7 Study 2 Design

Based on the reviewed literature, in this research, we want to test the event-related potential (ERP) of specific events in a VR game and how they compare to a regular screen, and how changing perspective in VR affects the ERP.

We decided that Mario Kart Wii was a suitable game to use, because of its gender neutrality [177], how easily reproducible phases between players can be, its “spoiler” resistance, and because it is relatively easy to reproduce events [69]. To convert the game to VR we used the Dolphin VR emulator (version 5.0) [180], and the configuration follows what was reported in [69].

We then chose the events to be evaluated. The game has many possible events to be evaluated. We decided to focus on three: One ubiquitous to every interaction of the game and two that the researchers observed on previous iterations to be frequent. Event 1 is Crossing the Finish Line; Event 2 is Collecting an Item; Event 3 is Falling off Map. We hypothesized that all events present greater indexes of Arousal-Valence and Focus/Engagement in VR when compared to a Common Display (CD), and also that indexes in 1PP-VR would be even greater than in 3PP-VR.

We recruited volunteers from a local university, regardless of their course of study. We had a total of 13 volunteers— (one female) with an average age of 23.69±3.76, median of 23, mode of 22, and range between 19 and 32 years. All volunteers had normal or corrected-to-normal vision, and none declared any history of colour blindness or mental illness. Of the participants, 84.6% of the participants had experience with the chosen game before but had never played it on a PC emulator nor in any kind of VR system. 53.8% of them had already had experience playing other games in VR, 69.2% played “regular” games at least once a week and, of these, 44.4% played at least three times a week.

All the volunteers received an explanation about the experiment, particularly the collection of their psychophysiological data. After understanding the process and what psychophysiological data would be collected, there was verbal agreement to continue with the experiment.

We used an Oculus Rift CV1 as our HMD. Our PC had 16GB RAM, an Intel Core i7-7700k CPU @ 4.20GHz, and a GeForce GTX 1080Ti dedicated GPU. We used a Betop Pandora 2.4G Wireless gamepad as the controller. We used it instead of the traditional

Rift Touch Controller because the Betop could be used both with an HMD and without it.

The standard data collection procedure was to introduce the players to the game and controllers using the 3PP-CD. We had the players play the game once under this perspective, while the EEG device calibrated on the volunteers' heads. This part of the process took approximately 11 minutes for each volunteer. The races were recorded so we could observe the events later.

The players played one of the three conditions, followed by answering questionnaires about their experience. The process was repeated for all three conditions, and the order of the conditions was determined by a Latin Square design. Due to space limitations, the questionnaires are not discussed in this research.

Section 7.8 Results and Discussion

We collected data from 13 participants. However, the data of two participants were discarded because they had stopped the experiment before it finished due to excessive Simulator Sickness.

During the analysis, we evaluated Arousal-Valence and Focus. We calculated the scores as presented in the relevant literature. In order to evaluate those components for each of the predetermined events, we graphed the scores starting one second before and finishing one second after the event had happened. In total, we had over 360 graphs characterising the events.

The analysis first identified that one of the initially chosen conditions had to be ignored, due to it occurring less commonly than expected, namely condition 3 - "Fall". Based on our previous experience, we expected participants to fall off the chosen map with a certain consistency. However, the results showed that was not the case. Even though people did fall, they did not do so consistently across versions, which means there was no solid basis to make comparisons.

Events 1 and 2 happened constantly and thus were evaluated in this research. Both of these events had relatively consistent graphs across versions, meaning that while each volunteer showed a different response to each of these events, the same volunteer presented similar responses across different versions to similar stimuli. This suggests that the impact of the events themselves had a higher impact on the players' minds than the display or point of view the event is presented in.

The data was challenging to compare across volunteers, with most of them showing changes in different indexes and in very dissimilar ranges. However, for Event 2, general patterns within volunteers could be observed. However, this event presented another challenge: to isolate it from other forms of interaction.

Event 1 almost ubiquitously led to a spike in valence, generally a few milliseconds before the actual crossing of the finish line. However, the peak either turned into a plunge if the ranking was not, we hypothesize, the desired position for the player, or flattened straight away. We could not find a specific pattern to determine what a desired position was, as this would require complex investigation. Thus, we are going to present a number of selected cases we felt were particularly interesting.

Subject A finished the first game (3PP-CD) in 1st place, and it seems his positive feelings peaked; however, on the next version, when obtaining the same result, even though the level of positive feelings increased, it was closer to apathy. Finally, when obtaining a worse result than in the previous versions, subject A presented a plunge in positive feelings, which could have been caused by either the result or the experiment being over. This is shown in Figure 7-3.

Figure 7-3 Graphs of ERP of Subject A's Event 1 under different conditions.

Subject B presented mixed feelings regarding his positions; we can see that in all his interactions there is an alternation between peaks and valleys. However, his data shows very clearly that patterns are closer within volunteers than across volunteers (Figure 7-4).

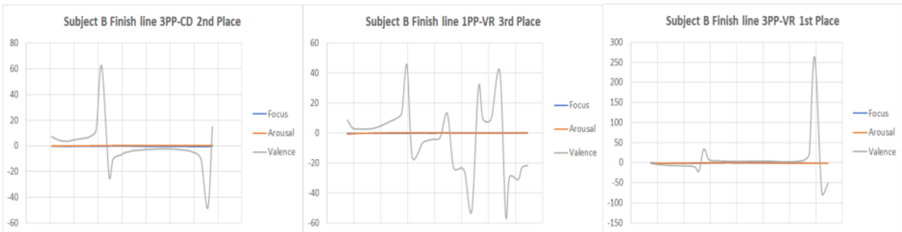

Figure 7-4 Graphs of ERP of Subject B's Event 1 of under different conditions.

With regards to Event 2, Figure 7-5 represents the first time Player A gets an item in each version. On the first graph (which represents the Subject's first race), Valence shows a very high peak when compared to the other two graphs. The range that the indexes varies in the later graphs is closer than the variation of the first. We speculate that this occurred because the rewards caused by getting an item were greater on the first iteration; and, based on the law of diminishing returns, the rewards in later iterations of getting an item probably became lower and lower (the player had acquired over 12 items by then). Another plausible explanation is that it was the beginning of the first race and expectations and the feelings of excitement may have been higher.

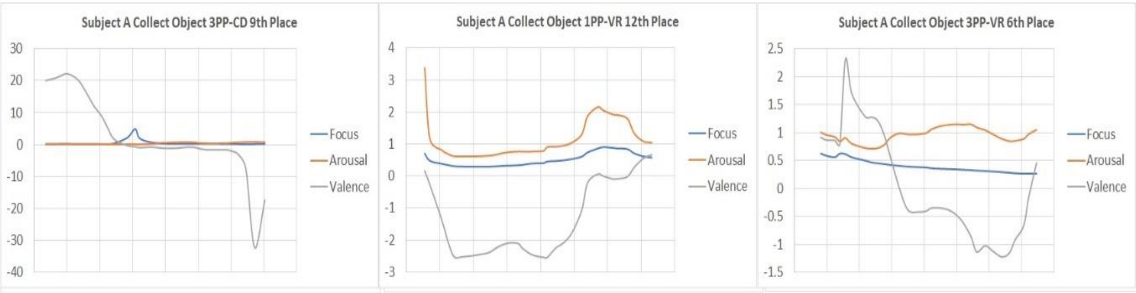

Figure 7-5 Graphs of ERP of Subject A's Event 2 of under different conditions.

Our hypothesis that events 1 and 2 in VR would cause greater indexes of Valence and Arousal was not confirmed, nor was the hypothesis that 3PP would have lower indexes than 1PP. Our hypothesis that "Falls" would have a greater impact in VR was not verified due to challenges with data collection.

It is possible that events such as falling, being hit by an object, or hitting another driver might elicit stronger reactions since the brain might interpret it as something likely to have actually hit the player; whereas a victory is regarded as a victory, regardless of medium. The polluted data of events 1 and 2 provided some supporting evidence of this when these events happened simultaneously with hitting another player, causing different reactions.

One possible direction we could take for future work is to try to summarize the patterns and identify interruptions using techniques similar to those used in activity-driven video summarization (see [210]–[212]).

Section 7.9 Study 3 Design

In order to find the possible correlations between subjective feelings and brainwave patterns when playing a VR game, we devised the following experiment. We had volunteers play a game under three different perspectives: 1PP-VR, 3PP-VR, and 3PP-CD while their brainwaves were recorded. In between each perspective, they answered subjective questions regarding the last perspective they played in.

The game chosen for the volunteers to play was Mario Kart Wii. This game was relatively easy to replicate the stimuli, did not present a storyline that could be "spoiled" after the first gameplay, and was easily playable and accepted by gamers of both genders [177].

We recruited volunteers from a local university, regardless of their course of study. We had a total of 24 volunteers (six were female). Their age ranged between 18 and 32 and had an average age of 22.75±3.76. All volunteers had normal or corrected-to-normal vision and did not have any history of mental illness.

We first introduced the participants, or players, to the game, controller, and the EEG device using the 3PP-CD setup. We had the players play the game once under this perspective, while the EEG device calibrated on the volunteers' heads. This part of the process took approximately 11 minutes for each volunteer.

We then had the players play one of the three conditions followed by answering the questionnaires from [69]. The process was repeated for all three conditions and the order of conditions was determined by a Latin Square design, which we used to mitigate carry-over effects.

All volunteers played one circuit, the Mushroom Cup, in the 150cc difficulty level and drove Mario's standard kart. We followed the configurations of Monteiro et al. [69], who measured the level of enjoyment, presence, and emulator sickness in the VR game based on 1PP and 3PP. Our apparatus was composed of the HMD Oculus Rift CV, and the consumer EEG device Interaxon MUSE MU-02.

Section 7.10 Results

The MUSE collected data at 256Hz, and the process of transformation to standard brainwaves bin (see Table 7-2) was done by the embedded software using an FFT which would produce brainwaves PSD every one-tenth of a second. The embedded software also classifies jaw clenching and blinking [209].

After the experiment, we removed data that our device classified either as blinks/jaw clench as bad (i.e., data produced when the contact points were not well-positioned). This resulted in the removal of data from two volunteers who experienced severe nausea, left the experiment earlier, and presented excessive movement, and other four who either just moved too much or the device did not fit well on their heads. Further, the data was filtered using a bandpass filter removing the 50Hz frequency and frequency below 1Hz. We calculated the results of the questionnaires following the procedure of Monteiro et al. [69]. We calculated the average of the PSD generated by our device; across the time the volunteers were playing. We used these averages and questionnaire results to calculate Pearson's correlation.

In 3PP, both CD and VR competence presented a negative correlation with Beta in the frontal lobe. In two modalities Gamma was negatively associated with positive affect. The feeling of being transported was negatively correlated with both Beta and Gamma (associated with attention and stress) in the VR versions. Overall immersion was positively correlated with Alpha in both VR versions but was not in the CD version, whose only correlation was a negative one with Beta. For more details on the Pearson correlation see Table 7-8.

In a follow-up analysis we did a series of linear regressions using the brainwaves from all different combinations as our independent variable and the values from the questionnaires as our dependent variables. We did it to see if the brainwaves could be used to predict the scores of some of the items. For Overall immersion, Beta and Gamma TP9 were the most relevant, both needing similar coefficients in both 3PP-CD and 1PP-VR.

We performed multiple within-subject ANOVAs with the average of the PSD. The ANOVAs presented borderline significance (p between 0.075 and 0.16). Thus, we did not consider them for analysis.

Section 7.11 Discussion

Section 7.11.1 General Observations

Traditionally Beta is associated with concentration levels, and as such we find it interesting that it correlates with competence in a similar fashion when the same perspective is applied regardless of the level of immersion of the display (see green cells on Table 7-8). This result is aligned with previous studies that suggest that perspective is more influential for gameplay than immersion [69]. This also indicates that, the easier people believe the game is, the less they have to concentrate.

The fact that competence presented a positive correlation with Alpha in the 3PP-VR version is especially interesting to us (see orange cells on Table 7-8). Alpha can be associated with both Flow and image processing and previous studies have shown that VR can cause Alpha to reduce [206]. We expected that Alpha would have a negative correlation with competence in the parietal region. However, it is interesting to see that Alpha does not present a correlation with overall immersion on the other conditions, which may be due to the lack of extra visual stimuli in the CD condition and perceived difficulty in the 1PP-VR.

Furthermore, it is interesting to see that the 3PP-CD and 3PP-VR present Beta and Gamma correlations not exclusively on the same side of the brain, which might suggest that they are not being processed in the same manner (see green and purple cells in Table 7-8).

Overall, immersion does not seem to present a clear correlation among versions. However, when doing a linear regression, it seems that the interception of Beta and Gamma are good indicators for the level of immersion across versions.

Table 7-8 Pearson Correlations between subjective feelings and brainwaves.

	Alpha AF7	Alpha AF8	Beta TP9	Beta AF7	Beta AF8	Gamma TP9	Gamma AF7	Gamma AF8	Gamma TP10
Competence-3PP-CD				-.566*			-.607*		
Competence-1PP-VR			-.528*						
Competence-3PP-VR	.485*	.702*	-.566*	-594**	-.624**	-484*	-.522*	-.691**	
Overall immersion-3PP-CD			-.476*						
Overall immersion-1PP-VR	.483*		-.634**		-.562*		-.609**		-.515*
Overall immersion-3PP-VR		.513*							

Green	Negative correlation to Beta 3PP across versions
Light yellow	Negative correlation to Gamma 3PP across versions and brainwaves across versions
Orange	Possible Flow (intense focus/enjoyment)

*p<0.05, **p<0.01

Our analysis is based on the literature and is in accordance with most similar researches. Nevertheless, we are aware that the brain waves can have further interpretations. Such interpretations, however, are beyond the scope of this research. We emphasize that what has been reported in this research can serve as a useful guide for future research.

Given the number of variables we have, we have not adjusted the Alpha-level using Bonferroni. Despite this, most of our variables presented correlations and we believe that most of the results are not produce by chance.

Finally, based on the analysis of the data, we can suppose that there is a physiological difference between gaming in VR and with a common display. For a more immersive first experience, it is likely that players should not be excessively challenged.

Section 7.12 Summary and Remarks

In this chapter, we present a theoretical introduction to a type of psychophysiological measurement for VR gaming. We have argued why it is important to study the effects of viewing perspective in VR gaming. We also discuss subjective and psychophysiological metrics for measuring gameplay experience. We describe some related work that uses those metrics. Later we report the results of an experiment to see what those metrics can tell us about the difference in viewing perspective in VR gameplay; Their EEG data are then compared with the quantitative data to answers collected from a subjective questionnaire completed by the same participants.

We have found that the current metrics for engagement, arousal, and valence might conflict with the subjective metrics for gameplay experience when analysed through a linear model. However, what they can tell us are user preferences and suggest perceived levels of difficulty within a game. They show no clear pattern when comparing viewing perspectives in VR gaming.

Our findings indicate that some waves correlate to subjective feelings regardless of viewing perspective (e.g., Gamma and Positive Affect), whereas others are more reliant on which perspective the gamers are playing in (e.g., Alpha). The results of this research might be applicable to the design of VR games aimed at improving players' health, especially from a psychophysiological perspective.

Even though we did not confirm our initial hypothesis, these results are relevant. They suggest that certain feelings are not so connected to how the experience is being seen (i.e., 1PP or 3PP) but rather to what it means to the gamers. Therefore, we can extrapolate that this lack of difference may also be behind why people feel so involved with sports in which they are simply spectators.

To the best of our knowledge, this is the first work that investigates the differences of viewing perspective in virtual reality gaming using those metrics. We believe this work can provide some foundations for other studies of psychophysiological analysis of viewing perspectives in VR gaming. Further studies, maybe using simpler games, will help shed some light on the reasons why the two kinds of metrics used in the study have not converged in our study.

Our sample data is not big but this is by design. We wanted to examine the data for each subject in detail. There was no analytical framework from prior research that we could follow in our analysis. As such, it was vital to perform this experiment with a small sample size. Data from subject A makes us believe that for a game to be enjoyable, it should increase in challenge gradually, since after the first victory, the only strong feeling present is that of disappointment.

To the best of our knowledge, this is the first work that investigates the differences between viewing perspective and display type in virtual reality gaming using ERP-based metrics. We believe this work can provide some foundations for other studies of psychophysiological analysis of viewing perspectives in VR gaming. Further analyses of our data and future studies may help shed some light on why we could not identify specific patterns in the data.

For future work, we plan to conduct a long-term study with various games to see whether the results remain the same or how they differ when players become accustomed and familiar with the game, gaming in VR, and playing in first- and third- person perspectives in VR.

Chapter 8 An In-depth Exploration of the Effect of 2D/3D Views and Controller Types on First-person Shooter games in Virtual Reality

Section 8.1 Introduction

VR has been proliferating rapidly in the last few years, especially with the advent of mass commercial VR HMDs. Due to their rapid growth, there are still aspects that are not well understood, including the influence of different types of controllers in virtual interactions [44], [213]–[215], and how they affect immersion [216]. Similarly, the effects of viewing types on SS during fast-paced applications, such as FPS games [116], [217], [218], are also underexplored.

Simulator Sickness (SS) or kinetosis has been studied even before the arrival of VR HMDs [106]. Because it is often believed that SS is caused by the brain receiving conflicting information from the senses, some researchers assume that the closer we get to the real physical environment, the less likely a person is to feel the symptoms of SS [190]. Although some researchers have suggested that such an approach is adequate [223] or at the very least indifferent [224], Dziuda et al. [165] have re-evaluated this assumption to identify some sort of uncanny valley of SS. They showed that a simulator that included a moving platform associated with the visual stimulus generated higher SS levels. Their results would imply that trying to go for the highest possible realism might not be an effective way to prevent SS. Because the actual mechanisms for SS might be different from those of motion sickness, an earlier study found it hard to correlate motion sickness with SS [225]. Thus, in this context, SS can also be known as (also referred to as cybersickness or visually induced motion sickness)

However, players pay attention to several factors during gameplay. In VR, the level of playability and immersion are important considerations because they affect enjoyment and performance, especially while players are (or are not) feeling SS symptoms. Playability and immersion can be affected by how the VR environment is controlled [16],

[219]. Nevertheless, we must note that research shows that immersive technologies can hinder performance when compared to a traditional setup (i.e., monitor, keyboard, and mouse) [220], [221]. From these studies, it is not clear if it is the keyboard and mouse combination as the input control responsible for the positive results. Likewise, it is not possible to know whether the enhanced results are the product of the reduced depth or visual detail of the 2D display that allows FPS players to perform better (e.g., having a better aim). After all, real-life marksmen close one eye when shooting, which effectively renders their vision to 2D mode as a filtering mechanism. Finally, it is not clear if better performance is caused by the combination of visual reduction of details and how players control or navigate the environment.

In this research, we first investigate existing techniques to mitigate SS and based on these results we explore the effect of 2D and 3D views and the controller types in immersive VR for FPS gameplay. To this end, we aim to analyse: (1) the effect of variations of those views on immersion and performance in VR; (2) the impact of using a standard keyboard and mouse combination compared to hand-held controllers that are used in today's VR systems.

To study the effects of the 2D views, we created a 2D visualisation technique, which we call PlaneFrame. It is a technique that slightly alters how the VR environment is viewed by users. Through three experiments, we found that this technique can improve performance when playing an FPS game in VR and decrease the level of SS with little effect on the users' perceived level of immersion. The results of our studies show that PlaneFrame could be a useful visual technique for FPS games and other VR applications.

Our main contributions in this research are a deeper understanding of the impact of 2D/3D views on VR FPS and the influence traditional input methods have on these kinds of VR applications. Furthermore, we found that a 2D view in VR HMDs can be a good compromise between performance and immersion and help reduce SS symptoms. Finally, we introduced PlaneFrame. This versatile and cost-effective visual technique can be used for VR applications, including FPS games and exploratory environments.

Section 8.2 Simulator Sickness Review

Intricate techniques to reduce SS, such as galvanic vestibular stimulation, require extra hardware and might not be usable by all users (e.g., users wearing pacemakers cannot use them) [226], [227]. Other techniques that seem to yield promising results are the ones that use vibration to stimulate the sense of movement, such as bone-inducted vibration [228] or producing small strikes on the region behind the ears [229]. Nevertheless, these techniques are still in their early stages, needing further development and evaluation to assess their real effectiveness, suitability, and long-term effects.

One technique that is simple but has yielded great success is the use of a gazing point in front of the user's view. The gazing point can be a circular point [168] or the head of a character in the case of the third-person perspective (3PP) view [14], [230]. Furthermore, this is somewhat a diegetic technique and can be easily inserted as the gun's aim and can be quite useful. The counterpoint is that gazing cues can attract the player's eyes [231], which could affect users' ability to see objects that fall outside of the area of the gazing point and are located in their peripheral vision.

Recently, techniques that are developed primarily for VR HMDs (rather than all kinds of simulators) have been proposed. These techniques focus more on how virtual worlds are presented rather than the movement or navigation technique alone. For example, some have attempted to apply blur during movement according to the distance of objects to the user [232]. Others deformed the peripheric region in which the user is moving towards to create the illusion of less movement or reduced FOV [233], [234]. Other techniques involve removing visual information from the user [23], [235]. While these techniques seem to present positive results in lowering SS, it is unclear if these optical illusions affect user performance and their perceived level of immersion in the VR environment. To the best of our knowledge, only two other studies have evaluated the effect of their mitigation technique on performance [236], [237]. This involves mainly changing the rotation aspect, either reducing the speed during rotation or removing information. Their techniques interfered little with the users' performance in their respective applications. Thus, it is likely that mitigation techniques are a promising path for VR games, especially for FPS games that require quick head movements [237]. Before we advance our study, it is important to understand the techniques aimed especially at those environments and that do not affect users' performance. In order to do so, we perform a systematic review.

Section 8.2.1 Contribution

The main contribution of this systematic review is to analyse the latest, state-of-the-art research on the use of visual techniques to reduce SS in VR HMDs. We have been able to extract key details of the study design, main types of applications used to test the results, most common measurements, and results from relevant papers. Although there are several review papers about SS [63], [238]–[241], with the most recent in 2020 [239], they focus on the factors associated with SS rather than on specific types of techniques that are designed to prevent them. Because of the rapid advancement of the technology, it is important and timely to catalogue and analyse existing techniques to build on them, identify gaps and shortcomings, and detect future trends. Our review led to 330 articles published after 2014 (to June 2020). In addition, a set of recommendations are provided that can help frame future studies.

Section 8.2.2 Materials and Methods

The structure and sequence of this review follow the PRISMA guidelines [242]. First, we need to define what we classify as a visual technique to reduce ambiguity and clarify the choices made further into the paper. For this paper, we classify a visual technique as one that changes how the virtual world is viewed without altering its physics or how navigation is done. That is, visual techniques can add or remove visual elements and still allow users to navigate the environment based on how it is done in the real world. We did not consider techniques that alter the movement in the virtual environment, such as teleport, require changes in velocity, or distortions in the virtual environment, as visual techniques. These are techniques that can alter the mechanics of the environment. For example, it is possible to add a grid to a racing game keeping the same rules, but teleporting would require evident changes. Furthermore, for this study, they must not involve specialized hardware, such as LED attachments, vibration motors, or any other improvements and changes added to the HMD.

We classify the visual techniques based on their main components. We define components as the parts shared by the techniques that can be altered within the articles but remain recognizable and are considered by their authors as the reason for the possible reduction in Simulator Sickness. For example, we consider FOV restriction as the main component of some techniques. It is present in different articles with variations in shape and size that define the specific techniques in each article; however, it is a recognizable component because it hides part of the FOV.

Section 8.2.3 Eligibility criteria

To be included in this systematic review, the publication was required to: (1) have used a visual technique to mitigate SS; (2) have been published after the end of 2013; (3) not be a locomotion technique (e.g., teleport, changes in velocity); (3) have used an HMD and no telepresence device; (4) have been written in English and appeared in a professional peer-reviewed journal or conference or have been cited by at least other 2 papers from the list; (5) have recruited at least 12 participants; (6) have involved both genders for analysis to reduce bias; and (7) have been rated at least to be of "moderate quality" based on the Quality Assessment Tool for Quantitative Studies [243]. No additional exclusionary criteria were employed.

Section 8.2.4 Information sources and search strategy

Key information technology databases were utilized to ensure we accessed the most current and relevant journal and conference papers (see Table 8-1). Articles were extracted in June 2020 by using four databases. The systematic review began with broad search terms, including 'virtual reality' and 'cybersickness'. Then we followed a more focused approach with well-defined inclusion and exclusion criteria (see Table 8-2 and Table 8-3).

Table 8-1 Bibliography databases used in the search and filtering criteria.

Bibliography Databases	Microsoft Research, ScienceDirect, IEEE Xplore, ACM Digital Library, and Google Scholar
Search Engines	Google and Google Scholar
Article Types	Journal Articles and Conference Papers
Search On	Title, Keywords and Abstract
Sorting on Returns	Relevance
Language	English
Publication Period	Starting 2014

Table 8-2 Inclusion and exclusion criteria used in the databases

Criteria 1(Inclusion)	(cybersickness OR “Simulator Sickness” OR “SSQ”)
Criteria 2 (Inclusion)	(Mitigate OR mitigation OR reduction OR reduce OR prevent)
Criteria 3 (Inclusion)	“virtual reality”
Criteria 4 (Exclusion)	-sound -training -acrophobia -education -navigation -CAVE -mobile -training -AR -"Augmented Reality" -Estimate -Classify

Table 8-3 Boolean search as input in the databases

(cybersickness OR “Simulator Sickness” OR “SSQ”) AND (Mitigate OR mitigation OR reduction OR reduce OR prevent) AND “virtual reality” -sound -training -acrophobia -education -navigation -CAVE -mobile -training -AR -“Augmented Reality” -Estimate -Classify

Section 8.2.5 Selection of studies

The selection was performed through five rounds (see Figure 8-1). In the first round, duplications were identified and removed. We then screened their abstract in the second round; articles that did not meet the eligibility criteria (1–6) were removed. Eligibility criteria (1–6) were used again in the third round, where articles were assessed through reading the whole article. In the fourth round, a quality assessment for each article was conducted using the Quality Assessment Tool for Quantitative Studies, where articles rated below “moderate quality” were excluded. For the quantitative analysis, only main components that had been tested in at least three similar experiments were evaluated so it meets the threshold to assure meaningful comparisons [244].

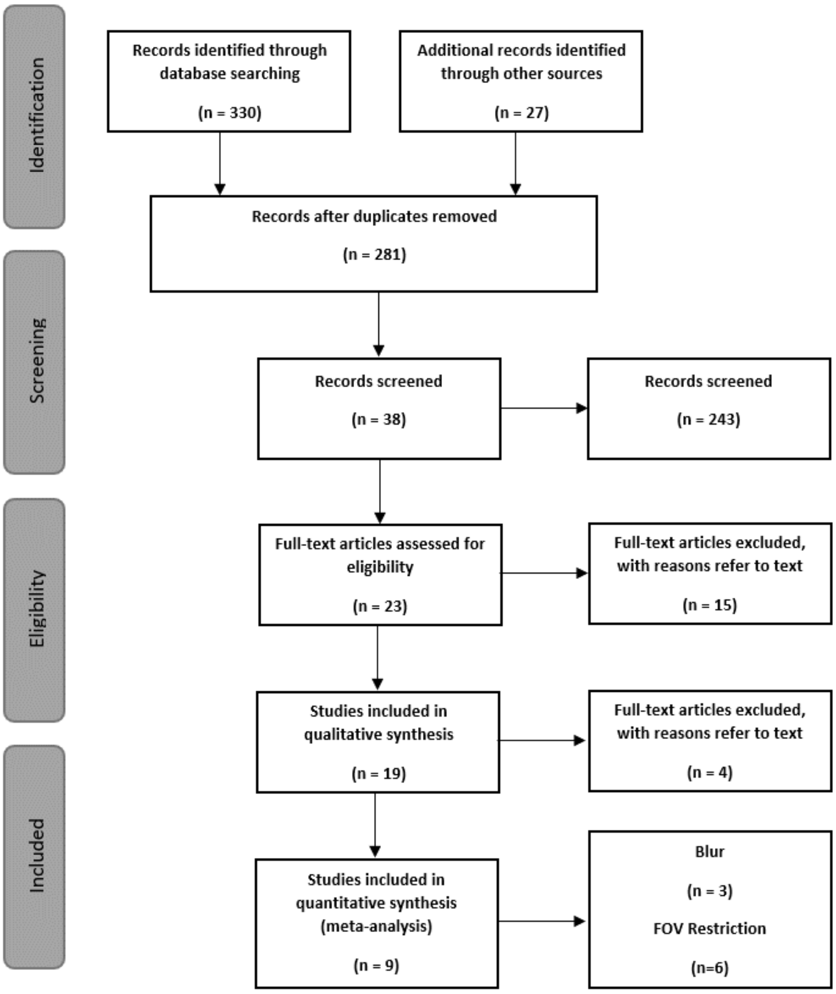

Figure 8-1. Summary of the discovery and selection process of the papers used for this review.

Section 8.2.6 Data items

Each row in Table 8-4 and Table 8-5 refers to a different study. The studies were summarized based on first author and year, study design (design, duration of sessions, and the interval between sessions), the testing environment (software, hardware, and controller), sample-related issues (sample size, dropout rate, age, gender), the main component of the visual technique, measurements of the study relevant for SS, and finally their results (Table 8-5).

We determined the quality of the studies by the Quality Assessment Tool for Quantitative Studies. The quality assessment is comprised of five components: (1) selection bias, (2) study design, (3) confounders, (4) data-collection method, and (5) withdrawals and dropouts. The bias risks are summarized regarding these criteria in Table 8-4. Each component was rated as weak, moderate, or strong, and a final rating was made for each study. A study was rated as "low quality" if it had received two or more weak ratings; "moderate quality" if it had received one weak rating or the majority as "moderate quality", and "high quality" otherwise. Any discrepancies in terms of rating were resolved between two reviewers. At least one other researcher helped reassess all items.

Section 8.2.7 Results

Section 8.2.7.1 Study selection

The systematic inclusion process included 330 papers (21 from Microsoft Research, 66 from ScienceDirect, 52 from IEEE Xplore, 78 from ACM Digital Library, and 113 from Google Scholar). Another 27 papers were added through a snowball process. After distilling them through inclusion and exclusion criteria, 19 papers were included in the final review and summarized in Table 8-5. The study selection process is outlined in Figure 8-1. After removing duplicates (N= 76), 281 articles were obtained. In the first round of screening the abstract, we excluded 243 papers because they were out of the scope of current review (e.g., they were applications, used AR instead of VR, involved hardware, etc.). In total, 38 articles were included after screening their titles and abstracts. Further exclusion included the removal of studies that (1) involved locomotion techniques as previously defined (N = 3); (2) did not present or study a mitigation technique (N=6); (3) did not use a VR HMD (N=2); (4) involved telepresence (N=1); and (5) used additional hardware (N=2). In the last round, we excluded four articles that were rated "low quality". The resulting studies can be found in Table 8-5.

Section 8.2.7.2 Study characteristics and risk of bias within studies

The first article identified was published in 2015, the number of papers peaked in 2018 with seven papers and declined again in 2019 to 3 papers. The origin of these studies was primarily from the United States of America (N=8), followed by Germany (N=4), China (N=2), Canada (N=2), and the UK, Australia, and Korea each with one publication. The quality of research bias of individual trials was rated either "moderate quality" (N=6) or "high quality" (N=13). The results are summarized in Table 8-4.

Table 8-4 Summary of the quality assessment of the papers selected for the review.

Study description	Selection Bias	Study Design	Confounding Factors	Data Collection	Withdraws	Overall
Isaza et al., 2019 [245]	***	***	***	***	***	***
Buhler et al., 2018 [246]	***	***	**	**	**	**
Farmani and Teather, 2018 [235]	***	***	***	***	***	***
Farmani and Teather, 2020 [237]	***	***	***	***	***	***
Nie et al., 2020 [232]	***	***	***	***	***	***
Jorge and Fumanal, 2017 [22]	***	**	**	**	**	**
Carnegie and Rhee, 2015 [247]	***	***	**	**	***	***
Budhiraja et al., 2017 [23]	***	**	***	***	***	***
Zhou et al., 2019 [248]	***	***	**	***	***	***
Fernandes and Feiner, 2016 [234]	***	***	***	***	***	***
Kim et al., 2018 [249]	***	**	**	***	**	**
Basting et al., 2017 [250]	***	**	**	**	**	**
Al Zayer et al., 2019 [251]	***	***	**	***	***	***
Adhanom et al., 2020 [21]	***	***	***	***	***	***
Norouzi et al., 2018 [252]	***	***	***	***	***	***
Zielasko and Freitag, 2018 [253]	***	***	***	**	***	***
Cao et al., 2018 [254]	***	**	**	***	***	***
Yu et al., 2017 [60]	***	**	**	*	**	**
Wienrich et al., 2018 [255]	***	***	**	**	**	**

Section 8.2.7.3 Components of the Techniques

In total, we identified six main components used in the techniques to reduce Simulator Sickness: Mono-Stereoscopic View (N=1), Resting Frame (N=4), Blackout Screen (N=2), Scenic Interruption (N=1), Blurring (N=4), and FOV Restriction (N=7). Figure 2 shows a summary. In the figure, these components are abbreviated to 2D, RF, BO, SI, Blur, and FOV, respectively

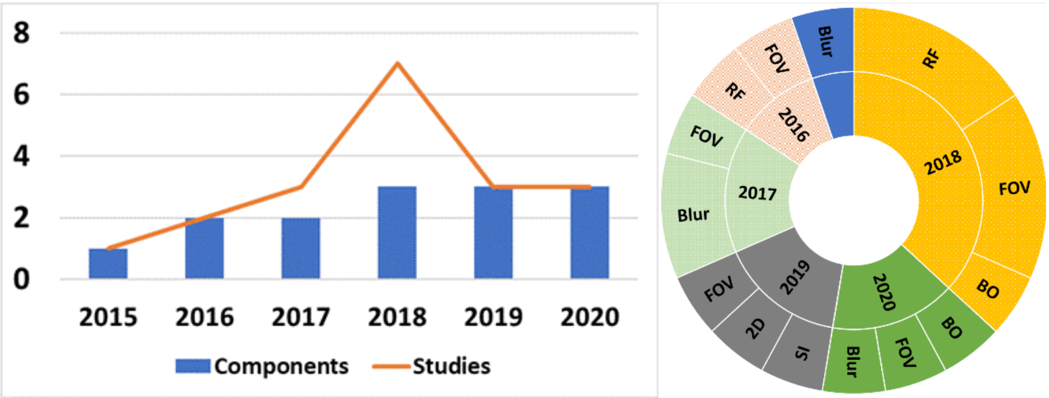

Figure 8-2 (Left) Graph showing the relation of main components evaluated and the number of studies performed. (Right) Graph showing the number of studies using each technique per year.

The different number of components studied each year did not increase much over the years, with its current peak value at 3. Even though the number of studies seems to have declined between 2018 and 2019, the number of techniques kept stable. Only in 2017 and 2018 were the number of studies higher than that of the main components.

Resting Frame was evaluated often (N=3) in 2018, but it was not a common occurrence in other years. The most consistently studied technique has been Field-of-View reduction, as it was studied at least once every year from 2016. Blurring is also a component consistently evaluated though not as often as FOV. As Figure 8-2 shows, it was evaluated in 2015, 2017, and 2020.

Section 8.2.7.4 Measurement Tools

In total, we identified four measurement tools used to assess Simulator Sickness. These tools can be divided into Reported (*N*=18) and Device Measured (*N*=1). Only one study used Body Sway as a measurement to gauge SS. Meanwhile, the SSQ was the most popular measurement tool (N=16). The studies that did not use the SSQ (N=2) and the studies that used more than one measurement tool (N=5) opted to use interactive scales such as Discomfort Score (DS) [234], Fast Motion Sick (FMS) [256], and Health Score (HS) [253].

From all the sixteen studies that used the SSQ, less than a third (N=5) presented the sub-scores. The sub-scores in the SSQ were divided into three categories, caring for nausea, oculomotor-discomfort, and disorientation. From the three categories, a total score could be calculated. The Discomfort Score could be presented in three different ways: (1) according to the previous DS called Relative Discomfort Score (RDS), according to the "current" DS defined as Absolute Discomfort Score (ADS), or according to the final value Ending Discomfort Score (EDS).

Section 8.2.7.5 Duration of Experiments

The majority of experiments had sessions that lasted between 10 and 25 minutes (N=10), three lasted for less than 10 minutes each session, and four lasted more than 25 minutes each session. The others (N=2) did not disclose the time for each session. Almost half of the experiments (N=8) had intervals of at least 24 hours between sessions; one had a pause of less than one hour; the others either did not pause or did not disclose it (N=9).

Even though FOV reduction was studied more often than the other techniques, its studies presented a relatively consistent duration, with a single outlier. Resting Frame times were the lowest among the studied techniques. Blurring studies varied in times, presenting an upper limit higher and lower limit smaller than three other techniques (see Figure 8-3).

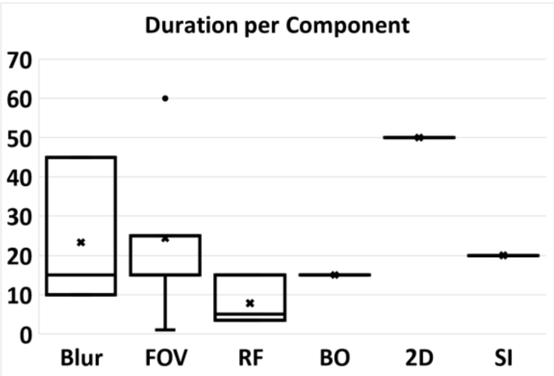

Figure 8-3 Graph showing the quartiles of the times the studies in minutes of each technique used. The "x" in the middle represents the mean. When times were expressed using a limit, we considered the limit as the standard time).

Section 8.2.7.6 Testing Environment

Games were a popular testing environment for these techniques (N=8), followed closely by scenic environments and virtual towns (N=7); moving vehicles were used two times. The most different testing environments were tunnels, and a graph path each used only once. The most popular kind of game was First-Person Shooter games (*N*=4), racing games were mentioned once, and the others were undisclosed.

Section 8.2.7.7 Device and Controller

The most common types of device used were various versions of Oculus Rift[16] (N=9), followed by HTC VIVE[17] (N=6); three experiments used a smartphone; FOVE VR[18] was used once. The versions of the Oculus Rift were divided into a majority of *consumer versions* (N=5), three DK2 and one DK1. Only three studies used the original controllers that came with the HMD, the same number of studies that relied on *head movements*. The most common form of controlling the virtual environment was using a Microsoft X-Box-like gamepad (N=8), and the other four used the keyboard and mouse combinations. Two papers either did not use any form of control or did not disclose it.

Section 8.2.7.8 Participants

Around 501 people participated to completion in the experiments of the 19 papers. The mean number of participants in each study was 23.9 participants, and the median number was 22. They ranged from 14 to 40 participants, with 15 studies having fewer than 30 participants. Only Blurring and Rest Frame studies had more than 30 participants and only in 2016 and 2020. The number of participants was fairly consistent among the years and techniques (see Figure 8-4). We run a Spearman correlation analysis and found no

[16] https://www.oculus.com/
[17] https://www.vive.com/eu/
[18] https://fove-inc.com/

clear correlation between the number of participants and the duration of each experiment (p >.05).

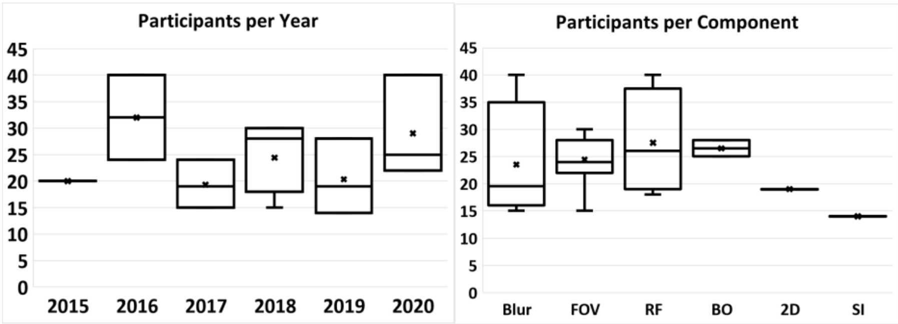

Figure 8-4 (Left) Number of participants used in the studies over the years; (Right) Number of participants used in the studies based on the main components.

Most participants were young adults; the average age of participants within the studies that allowed this calculation (N=10) was 25.2 years old. The other studies offered the age information in a *range* (N=6), with a lower limit (N=1) or not at all (N=2). The age range for all the studies combined was from 18 to 53 years old. However, participants aged 50 years or older can optimistically only account for just 7.4% of the sample population (these studies were range only). Between 54.2% and 62.9% of the participants in these studies were male. Some studies did not make the ratio clear (N=2).

A total of 48 participants were declared as withdrawn from the experiments. The withdrawal rate from the experiments that reported this information was between 6.3% and 21%. Given the nature of the experiments, the number 48 is somewhat low, probably caused by studies that omitted those results. From the studies that declared participants who dropped out, only six declared the gender ratio; these studies account for 141 participants. From these 15 left the experiment earlier, and at least six were men. That is, at least 40% of the accounted gendered withdraws were male.

Section 8.2.7.9 Summary of Study Results

Overall, the studies show a positive trend in SS mitigation. Sixteen studies observed some form of improvement after using their techniques; however, almost half (N=7) declared that their results were only applicable in some cases (N=4), while for the others (N=3) the results were not statistically significant or were inconclusive. Two observed no change, and one observed a worsening in the condition. FOV restriction was responsible for divergent results; that is, it produced no change or sometimes even negative results. All the other techniques only observed improvements (some not statistically significant).

There was no clear pattern to the analysis of studies that had statistically significant differences. From the studies that analysed Blurring, half found statistically significant differences. However, it is important to note that in these studies some people got worse after experiencing it. Similarly, FOV restriction had people who got better and worse spread around studies. The others only presented positive results.

Even though there were no clear patterns among the techniques that did find significant differences each presented slightly different methodologies from the majority. In the FOV restriction, one that found significant differences divided the data into participants who improved with the technique and the ones who had their sickness worsen with it. Another used an "objective" indicator rather than relying on the SSQ. And one other article that declared significant differences presented the SSQ as a function of nausea, dizziness, and oculomotor subscales and used more measurements beyond the standard SSQ.

Section 8.2.8 Meta-Analysis

For our meta-analysis, we divided the eligible studies between Blur and FOV Restriction. For Blur studies we utilized three papers that accounted for three studies [23], [232], [247], whereas for studies dealing with FOV restrictor we had six papers that accounted for 23 studies.

The homogeneity test of the Blur studies revealed a lack of heterogeneity in our data (Q (2) = 1.122, p = .571). We then used a Fixed-Effects Model, with mean d = -.554 (p=.004). Our Rosenthal Fail-Safe *N* revealed that six non-significant studies would be necessary to make our overall effect non-significant. This number of required studies is double of the ones currently available. This result shows that Blur-based techniques appear to have a positive impact on reducing SS.

For the analysis of the FOV restrictor-based techniques, the homogeneity test revealed the sample to be heterogeneous ($Q(22) = 112.329$, $p = .000$). We used a Random-Effect Model; however, the use of FOV to reduce SS was not significant ($d= .247$, $p= .103$). Thus, with the current studies, we were unable to see an effect of FOV restriction on SS.

Section 8.2.9 Discussion

This article has summarized and analysed the findings of 19 research articles, with 13 rated as “high quality” and 6 as “moderate quality” based on the Quality Assessment Tool for Quantitative Studies. Even though the effects of applying techniques could be seen in techniques that have used within- and between-subjects design, SS is very reliant on individual differences. Because of the highly variable and cumulative nature of SS, we recommend that (***R1***) more studies resorting to using a within-subjects experimental design that is spread across multiple days. The high variability among participants could be seen with at least four studies having people who were positively affected by the technique, whereas others did not see any advantage. We suggest that (***R2***) future studies present the data for SS variation in two groups: (1) those that did improve and (2) those that did not. Some techniques might have been discarded as ineffective because they were not adequate for all participants, when in fact they could have been highly effective for others, as some studies we have analysed.

Furthermore, we recommend that (***R3***) the duration of each experimental session lasts between 10 and 15 minutes. This length will help avoid participants’ adaptation to the VR environment and still being long enough to produce the SS symptoms to be avoided. Studies that used longer experimental times did not necessarily present more significant (or non-significant results). In addition, based on our revise, we suggest that (***R4***) a period of at least 24 hours should be observed before the next session to avoid cumulative effects.

Our results show a slightly higher chance of dropouts from women in SS experiments. However, these results need further analysis. We recommend that (***R5***) future studies be more detailed about their demographics so that the effects of gender and age can be better understood, as they seem to be impactful. Furthermore, because the influence of gender is not fully understood, we recommend that (***R6***) studies should present the effectiveness of their technique(s) divided into *general*, *male*, and *female*, and if necessary, use a *non-binary* designation. Moreover, we observed that no study had involved participants who do not identify as either male or female. Given the slight imbalance in dropouts, this might be an influencing factor.

We observed a lack of reproduction between studies. The analysed techniques were tested only between a control group and themselves and, as such, it is difficult to compare differences among techniques to determine the adequate ones. Moreover, there is a lack of objective measurements when analysing these techniques. This improvement could help comparability among techniques in the future. Furthermore, one of the studies that reported significant differences among techniques utilized at least one objective indicator. Because many studies that used the SSQ reported improvements but are non-significant, we hypothesize that: (1) this tool is based on subjective data and is not precise to detect the variations in sickness when studying mitigation techniques; and (2) some components of the SSQ might dilute the variations in sickness caused by the different techniques. Thus, it is important to (***R7***) present the subscales of the analysed measurements. Considering the small consequences of a false positive and the typical sample size of 20 participants, we recommend that (***R8***) for this type of experiment, an alpha of 0.10 and beta of 0.20 could be adopted as the standard [143], [257]. Otherwise, the chance of false negatives might be too high, resulting in fewer publications and thereby limiting future meta-analyses that could be used to validate effective techniques.

Moreover, the meta-analysis results suggest that the effect size of techniques to reduce SS using blur is medium (Cohen's $d \sim 0.5$). This result means that when evaluating if a technique is indeed useful, we should see (***R9***) if the difference between group means is at least 0.5 standard deviations.

Even though the number of studies decreased after 2018, the number of components studied remained stable and new main components were introduced in 2019. In 2020, we did not see any new, novel main components, but those reported in this year's studies mainly expanded on existing ones. On the other hand, the number of different main components studied remained at its peak. This suggests that a relatively high number of different research groups are studying the topic, and thus we can infer that it is still an open problem. Finally, we recommend that (***R10***) more studies focus on investigating the different effects of Mono-Stereoscopic Rendering systems in VR and Scenic Interruption. These were the least explored topics in this review but also have the potential to help new VR users. Moreover, they were the latest techniques to be introduced in the fight against SS.

The following are the aggregated lessons or recommendations that can be distilled from this review: Studies evaluating visual techniques for Simulator Sickness mitigation should take (***R3***) 15-minutes per condition with a (***R4***) 24-hour break, with a (***R1***) within-subjects design. The analyses should rely on at least two measurement tools, and the results should be present in detailed regarding both demographics and subscales (***R5, R6, R7***), and separated in groups by positive or negative results (***R2***). Studies on the topic should observe effect size (***R9***) and relax the Alpha (***R8***). And evaluate the techniques that have been studied the least (***R10***).

In summary, we list the recommendations as follows:

(***R1***) Prefer a within-subjects experimental design;

(***R2***) Present the data for SS variation in two groups: (1) those that did improve and (2) those that did not;

(***R3***) Prefer experimental session that lasts between 10 and 15 minutes;

(***R4***) Observe a 24-hour resting period between sessions;

(***R5***) Detail the demographics (e.g., gender, age);

(***R6***) Divide the results by demographics, especially regarding gender;

(***R7***) Present the subscales of the analysed measurements;

(***R8***) Adopt an alpha of 0.10 and beta of 0.20;

(***R9***) Observe if the difference between group means is at least 0.5 standard deviations.

Section 8.2.9.1 Limitations

This review has two limitations. First, there could have been some publication bias given that quite often positive results tend to be accepted for publication or presentation than studies with negative results. Second, we only considered published studies in English and restricted our search to publications available in the main research databases and references within the same. Finally, our selected keywords and inclusion criteria could have hidden some relevant studies, though we attempted to be as inclusive as possible within the topic.

Section 8.2.10 Remarks on Simulator Sickness

This review demonstrates that using visual techniques is possible to mitigate SS without extra hardware (e.g., galvanic skin stimulation) and without changing the movement patterns in VR like teleporting. The results have been observed through different devices and tested via within- and between-subjects experiments. We found that most studies have overlooked the gender differences when evaluating the effectiveness of a technique to mitigate SS, which contributed to our inability to establish the frequency of these differences. We also found that, even though the most investigated techniques are based on FOV reduction, this type of visual technique requires further investigation because their results are not consistent among reported studies. On the other hand, the other techniques require further investigation to increase the power of the findings. Overall, there have been many published articles regarding VR but not many techniques to stop SS from using visual techniques. Finally, we observed that when analysing SS it is important to have at least one other measurement method to validate the results and that ideally any investigation regarding SS should be divided into those who benefit from the technique and those who do not.

Table 8-5 Summary of the papers used in this review associated with their methods and results.

Study description	**Research design (design; duration of sessions; interval)**	**Test Environment: Software; Hardware; Controller**	**Sample Details: Final Size (dropout); Range Age (mean, s.d.); Gender M:F**	**Main component**	**Condition**	**Measurement; Value**	**Findings**
Isaza et al, 2019 [245]	Within Subject design; < 50 minutes; 24 hours	VR Moving Vehicle; Merge VR; Head Movement	19 (ND); 20-30 (ND); ND	Mono-Stereoscopic View	Control Mono-Stereoscopic View	SSQ ND SSQ ND	The results indicate that cybersickness decreased amongst users with little to no VR experience hindered when using the Dynamic Mono-Stereoscopic Rendering System.
Farmani and Teather, 2018 [235]	Between-Subjects design; 10-20 minutes; NU	Own shooting Game; Oculus CV1; Keyboard and Mouse	28 (3); ND (26.4, ND); 17:11	Black out Screen	Control Viewpoint Snapping	SSQ 48.1 (s.d. 19.83) DS (0 - 10) 5.5 (s.d. 0.5) SSQ 29.8 (s.d. 17.25) DS (0 - 10) 3.6 (s.d. 0.3)	SSQ was reduced by 40%; Disorienting at first; Nausea scores increased as a function of time however in a lower level; No difference in performance.
Farmani and Teather, 2020 [237]	Between-Subjects design; 10-20 minutes; NU	Own shooting Game; Oculus CV1; Keyboard and Mouse	25 (3); 18-35 (26.4, ND); 14:11 20 (2); 18-39 (28.7, ND); 14:6	Black out Screen	Control 1 Rotational Snapping Control 2	SSQ 48.1 (19.83); DS (0 - 10) 5.5 (+/-0.5) SSQ 29.8 (s.d. 17.25) DS (0 - 10) 3.6 (s.d. 0.3) SSQ 52.1 (46.5) ; DS 4.1 (0.75)	Rotational snapping is the same as Viewpoint Snapping; Translational reduced SSQ by 47%.

					Translational Snapping	SSQ 27.1 (s.d. 24) DS (0 - 10) 2.2 (s.d. 0.3)	
Nie et al., 2020 [232]	Between-Subjects design; 45 minutes; NU	Own Racing Track; HTC Vive; Gamepad Joystick	40 (2) ; 18-32 (25.7, 2.9) ; 20:20	Blur*	Control Dynamical Blurring	SSQ 27.47 (11.03) SSQ 17.79 (13.01)	Experimental data support the notion that dynamic non-salient area blurring can be used to reduce VR sickness in asymptomatic individuals.
Jorge and Fumanal, 2017 [22]	Within Subject design; ND; 24 hours	Unspecified Game; FOVE VR; ND	19 (5); 22-53 (ND); 14:5	Blur	Control Dynamical Blurring	SSQ ND SSQ ND	Sickness reduction in both oculomotor and nausea categories was not statistically significant despite a drop in the mean.
Carnegie and Rhee, 2015 [247]	Within Subject design; 15 minutes + 15 minutes; 24 hours	Two scenic Views; Oculus DK1; Keyboard and Mouse	20 (6); 18-50 (ND); 14:6	Blur*	Control Dynamic Depth of View	SSQ 21.46 (16.5) SSQ 13.58 (13.5)	6 withdrew from the experiment because of nausea before using the technique; Most felt better with the technique; two felt no difference.
Budhiraja et al., 2017 [23]	Within Subject design; 10 minutes; 24 hours	Own shooting Game; Oculus DK2; Keyboard and Mouse	15 (3); 18-26 (ND); 12:3	Blur	Control Rotation Blur	SSQ 51.36 (36.67) MS Scale (0 - 6) 3.8 SSQ 42.14 (27.61) MS Scale (0 - 6) 3.1	Blur helped 8 out of the 15 participants, was indifferent to 3 of them, and worsened symptoms to 4 of them; Participants who saw an increase were less prone to sickness.
Zhou et al., 2019 [248]	Mixed design; 20 minutes; >24 hours	Own Virtual Town; Oculus CV; Oculus Move	>=14 (2); >18 (ND); ~8:6	Scenic Interruption	Control Cognitive Distraction	SSQ 60.9 (44.2) SSQ 44.35 (33.9)	Results were not statistically significant; However, results seem promising; women were not helped as much as men.

Fernandes and Feiner, 2016 [234]	Mixed design; < 25 minutes; > 24 hours	Oculus Tuscany Demo; Oculus DK2; Gamepad Joystick	24 (2); 20-27 (22.1); 16:8	FOV Restriction	No Restrictor (1st Exposition) With Restrictor (1st Exposition) No Restrictor (2nd Exposition) With Restrictor (2nd Exposition)	SSQ 61.1 (30.4) ADS 4.95 (1.35) RDS 7.33 (1.92) SSQ 57.7 (17.2) ADS 2.97 (1.24) RDS 5.18 (3.01) SSQ 48.4 (44.4) ADS 3.16 (1.94) RDS 4.63 (3.07) SSQ 48.1 (30.2) ADS 3.93 (2.25) RDS 5.05 (2.86)	Participants who use the Restrictor first are affected less by the No Restrictor condition. However, differences were not statistically significant otherwise.; Only one participant was able to follow through the whole experiment N1.
Kim et al., 2018 [249]	Within Subject design; 1 minute; ND	RollerCoaster; Samsung Gear VR; Head Movement	28 (ND); ND; ND	FOV Restriction*	Control Technique	Body Sway 1.78 Body Sway 1.21	The result shows that the proposed method is significantly effective to reduce VR sickness by 31.4% ($p < 0.005$).
Basting et al., 2017 [250]	Within Subject design; 25 minutes; ND	Cavernous tunnel; HTC Vive; Vive Controller	24 (ND); ND; ND	FOV Restriction	110º 77º 55º	Vection Rate (0 - 5) 2.8 Vection Rate (0 - 5) 2.2	Greater FOV increased SS more.

					33º	Vection Rate (0 - 5) 1.6 Vection Rate (0 - 5) 1.4	
Al Zayer et al., 2019 [251]	Mixed design; 25 minutes; 10 minutes	Rocky Hills Environment; HTC Vive; Gamepad Joystick	28 (2) ; 18-33 (23.0, 3.6); 14:14	FOV Restriction*	Control (All) Technique (All) Control (Men)	SSQ – Total 51.29 (40.0) SSQ – Nausea 42.25 (40.0) SSQ – Oculomotor 39.52 (28.5) SSQ – Dizziness 56.67 (48.7) ADS 1.64 (1.5) EDS 2.86 (2.4) SSQ – Total 35.53 (23.4) SSQ – Nausea 30.66 (28.4) SSQ – Oculomotor 27.61 (16.1) SSQ – Dizziness 36.79 (29.2) ADS 0.98 (1.1) EDS 1.82 (2.1) SSQ – Total 49.69 (46.6) SSQ – Nausea 42.25 (42.1) SSQ – Oculomotor 36.28 (33.1) SSQ – Dizziness	FOV restriction was shown to be effective in mitigating VR sickness symptoms in both sexes; did not find a significant effect of FOV restriction on spatial navigation performance analysis did not find a significant sex difference in any of the VR sickness measures.

						56.67 (58.4) ADS 1.36 (1.4) EDS 2.64 (2.8)	
					Technique (Men)	SSQ – Total 29.92 (19.8) SSQ – Nausea 26.58 (21.9) SSQ – Oculomotor 23.82 (16.8) SSQ – Dizziness 28.83 (27.6) ADS 0.85 (1.1) EDS 1.29 (1.4)	
					Control (Women)	SSQ – Total 52.89 (34) SSQ – Nausea 42.25 (39.4) SSQ – Oculomotor 42.77 (23.9) SSQ – Dizziness 56.67 (38.8) ADS 1.93 (1.6) EDS 3.07 (2.1)	
					Technique (Women)	SSQ – Total 41.14 (25.9) SSQ – Nausea 34.75 (34) SSQ – Oculomotor 31.4 (15.1) SSQ – Dizziness 44.74 (29.5) ADS 1.11 (1.1)	

						EDS 2.36 (2.8)	
Adhanom et al., 2020 [21]	Within Subject design; < 60 minutes; 24 hours	Windridge City environment; HTC Vive; Gamepad Joystick	22 (3); 20-32 (23.7,3.9); 11:11	FOV Restriction	Fixed Restrictor (All) Foveated Restrictor (All) Fixed Restrictor (Men)	SSQ – Total 17.51 (19.5) SSQ – Nausea 13.44 (16.0) SSQ – Oculomotor 14.82 (15.6) SSQ – Dizziness 18.35 (22.78) ADS 0.62 (0.7) EDS 1.23 (1.8) SSQ – Total 24.48 (27.6) SSQ – Nausea 16.04 (20.9) SSQ – Oculomotor 21.36 (22.2) SSQ – Dizziness 28.47 (36.8) ADS 1.14 (1.5) EDS 1.64 (2.2) SSQ – Total 13.26 (10.9) SSQ – Nausea 12.14 (12.1) SSQ – Oculomotor 9.65 (9.0) SSQ – Dizziness 13.92 (17.6) ADS 0.38 (0.5) EDS 1.18 (2.2)	Results did not find a statistically significant difference between the fixed and foveated restrictors for VR sickness as measured.

					Foveated Restrictor (Men)	SSQ – Total 22.78 (25.1) SSQ – Nausea 15.61 (20.1) SSQ – Oculomotor 21.36 (19.7) SSQ – Dizziness 22.78 (33.1) ADS 1.27 (1.9) EDS 1.91 (2.7)	
					Fixed Restrictor (Women)	SSQ – Total 21.76 (25.3) SSQ – Nausea 14.74 (19.7) SSQ – Oculomotor 19.98 (19.3) SSQ – Dizziness 22.78 (33.1) ADS 0.87 (0.9) EDS 0.87 (1.4)	
					Foveated Restrictor (Women)	SSQ – Total 26.18 (31.0) SSQ – Nausea 16.48 (22.6) SSQ – Oculomotor 21.36 (25.3) SSQ – Dizziness 34.17 (41.0) ADS 1.00 (0.9) EDS 1.36 (1.6)	
Norouzi et al., 2018 [252]	Within Subject design; < 20 minutes;	Own Virtual Forest; HTC Vive;	15 (1); 19-24 (21.7);	FOV Restriction¬	Control (G1)	SSQ – Total 12.65 (30)	Majority of participants (11 of 15 in total) experienced

	24 hours	Head Movement	~9:6			SSQ – Nausea 9.42 (28.53) SSQ – Oculomotor 10.29 (25.59) SSQ – Dizziness 14.12 (64.12) RDS (0-10) 1.13 (0.35)	significantly more VR sickness if the velocity or acceleration-based vignetting methods were used than in the control condition; However, 4 participants had improved scores with the vignetting, suggesting that it can be positive for some.
					Vignetting Velocity (G1)	SSQ – Total 24.41 (52.94) SSQ – Nausea 14.12 (26.76) SSQ – Oculomotor 22.35 (43.82) SSQ – Dizziness 44.41 (79.41) RDS (0-10) 2.49 (0.7)	
					Vignetting Acceleration (G1)	SSQ – Total 41.18 (53.53) SSQ – Nausea 33.82 (42.06) SSQ – Oculomotor 31.76 (42.94) SSQ – Dizziness 44.41 (79.41) RDS (0-10) 2.90 (0.61)	
					Control (G2)	SSQ – Total 50 (45) SSQ – Nausea 43.24 (29.12) SSQ – Oculomotor 42.06 (35) SSQ – Dizziness 45.59 (64.12) RDS (0-10) 4.74 (1.42)	
						SSQ – Total	

					Vignetting Velocity (G2) Vignetting Acceleration (G2)	35 (36.76) SSQ – Nausea 24.12 (31.76) SSQ – Oculomotor 32.64 (50.29) SSQ – Dizziness 35 (31.4) RDS (0-10) 3.54 (1.21) SSQ – Total 35 (60.29) SSQ – Nausea 31.47 (55) SSQ – Oculomotor 25 (60.59) SSQ – Dizziness 35 (60.88) RDS (0-10) 3.84 (1.13)	
Zielasko and Freitag, 2018 [253]	Between-Subjects design; 15 minutes; NU	Graph path; Oculus CV1; Gamepad Joystick	30 (3); ND (27.3, 4.7); 24:6	FOV Restriction	Control Technique	SSQ 47.6 (38.5) HS (0-10) 4.7 (2.4) SSQ 31.7 (29.3) HS (0-10) 3.4 (2.9)	Participants who left were from the control condition; In contrast to existing research, they found no correlation between the subjective vulnerability to motion sickness and measured cybersickness. None of the sickness measures showed a significant difference. However, the reduction group moved more.
Cao et al., 2018 [254]	Mixed design; 15 minutes; 48 hours	VR Apocalypse; Oculus CV1; Gamepad Joystick	22 (0); 18-39 (22.6, 5.2); 16:6 20 (2); 18-49 (22.4, 6.4); 15:5	Resting Frame	No Rest Frame (1st Exposition-G1) Static Rest Frame (1st Exposition-G1)	SSQ 65.96 DS (0-10) 6.59 SSQ 68.57 DS (0-10) 4.44 SSQ	SSQ did not change; DS was worse without anything, but results are inconclusive; Pairwise comparison revealed both groups had significantly higher overall SS in their first sessions.

					No Rest Frame (2nd Exposition-G1) Static Rest Frame (2nd Exposition-G1) No Rest Frame (G2) Dynamic Rest Frame (G2)	38.89 DS (0-10) 4.05 SSQ 43.52 DS (0-10) 3.55 ND ND	
Yu et al., 2017 [60]	Between-Subjects design; 3.5 minutes NU	Own Game; ND Smartphone; Head Movement	40 (ND); 21-28 (ND); 27:13	Resting Frame	Control Cross Minion Border Ring	SSQ 22 (13) SSQ 15.5 (10) SSQ 7.5 (9) SSQ 22 (10.5) SSQ 15 (10)	The researchers tried to cause sickness through blurring; Their research demonstrates that the existence, the eye-catching position and the abundant texture of static symbols can decrease Simulator sickness to some extent; symbols with abundant colors and texture are more effective in the reduction.
Wienrich et al., 2018 [255]	Between-Subjects design; ND; NU	Own Jumping Game; Oculus DK2; Gamepad Joystick	30 (ND); 24-36 (29,ND); 15:15	Resting Frame*	Control Virtual Nose	SSQ 25 (2.5) SSQ 20 (1.5)	The virtual nose reduced simulator sickness without negatively interfering the experience of the game.

Buhler et al., 2018 [246]	Within Subject design; 5 minutes; ND	Own Virtual Town; HTC Vive; Vive Controller	18 (ND); 19-38 (ND); 12:6	Resting Frame	Control Circle Effect Dot Effect	FMS (0-20) 3.056 (s.d. 3.208) FMS (0-20) 1.556 (s.d. 3.434) FMS (0-20) 2.333 (s.d. 2.425)	Fewer users reported high impact on sickness with the technique; however, the results overall were not statistically significant.
ND – Not Disclosed; NR – Not Relevant; NU- Not Utilized; SSQ – Simulator Sickness Questionnaire; FMS – Fast Motion Sickness Scale; DS -Discomfort Scale; HS – Health Score							

Section 8.3 Related Work

Our research is related and informed by three main themes: (1) Simulator Sickness; (2) gameplay performance; and (3) view modality. First, in the previous sessions we have reviewed and summarize previous studies on mitigating SS associated with higher level of immersion, we analysed the data and propose some guidelines for future studies. Now, we discuss viewing technologies and how they affect immersion and enjoyment. Last, we present previous studies that have dealt with the trade-off between immersion and performance.

It is common to separate the terms immersion from presence. However, due to their close connection, they are often used interchangeably [222]. Hence, we also adopt this approach.

Section 8.3.1 2D/3D Views in VR

While commercial VR HMDs are relatively new, techniques that focus on increasing the sense of immersion are not. One such example is commercial 3D TVs, which have been used to demonstrate that having a 3D image changes how information is processed in the brain. They can elicit not only a greater sense of presence but also greater vection (the sensation of movement of the body in space produced purely by visual stimulation) when compared to 2D images [207]. Such findings are consistent with research that has evaluated 2D/3D VR when comparing a 3D wall display against a 2D monitor [206]. This research seems to indicate that people process 2D and 3D views differently. However, both of these studies were conducted in CAVE-like displays to navigate mazes. Hence, it is still unclear if the same results would apply to other tasks, like looking for fast-moving opponents in a game, or whether more immersive technologies like VR HMDs [258] would produce the same results.

One trait that seems consistent among the different studies is that people often seem to yield the highest enjoyment from the most immersive experiences, especially in HMD type of displays [220], [221], [230], [258]–[260]. This is the case even though HMDs tend to cause nausea and other kinds of Simulator Sickness afflictions [207], [260].

Most of these studies compared HMD VR, PC, or a CAVE in some combination. However, to our knowledge, there has not been a study that has compared 2D and 3D views within an HMD. As such, it is unclear whether the higher level of satisfaction comes from the HMD display, regardless of whether it shows 2D or 3D views.

Section 8.3.2 FPS Performance

FPS games are a well-established genre, with reasonably simple gameplay mechanics that well-suited to being translated into VR due to their inherently first-person view format. Unavoidably, there are challenges to bring an established format (in 2D displays) to a new platform like VR. One possible problem is the loss of player performance. For example, in a study that compared a CAVE to a PC FPS, participants performed considerably worse in CAVE; however, the same participants declared having enjoyed more the CAVE experience [220], [221]. It could be likely that VR HMDs may suffer from the same issue, trading immersion for performance [261]. However, not all studies agree with this view; [222], when trying to replicate a study that compared an HMD and a desktop display for an RPG [262], found no significant differences between the two in terms of user satisfaction and presence. One reason for the discrepancy is because FPS is naturally more engaging. A second explanation might be the kind of controller that players used, namely keyboard and mouse in all platforms, which could potentially break the immersion.

One issue with these studies is the type of controller used because the studies had their controllers match the technology (i.e., keyboard for the desktop and a game controller for VR). As such, their results might have had more to do with the controller than the display. After all, earlier studies on the influence of the type of controller had found a difference in presence when comparing a more natural Wiimote to a PlayStation game controller for a tennis game [216]. There are even differences in brainwaves when comparing both controllers [45].

For FPS games, [44] found that there is no difference between the Oculus Touch dual hand controller and the Xbox game controller [44]. However, because the Oculus

Touch is more closely related to the Xbox controller than to the keyboard and mouse, it may not be possible still to discard the possibility that the type of controller could affect immersion and performance. In another study, it was found that the keyboard and mouse led to lower performance [215]. Like other studies, the authors also matched the controller to the technology. With new proposed designs emerging for VR FPS controllers [219], the search for a suitable controller is an ongoing process [261] and an important aspect to be explored to seek ways to improve both immersion and performance for FPS VR games.

In short, prior research shows that the closer we get to simulating a real environment, the lower the performance can get. Also, it is not clear how this affects SS. This research aims to fill this gap by investigating the following research questions:

RQ1. Is performance loss caused by the VR environment or by the type of controller used?

RQ2. Do visual stimuli play a more prominent role in in-game subjective feelings than how we control them?

RQ3. Is the stereoscopic 3D view a significant factor within the HMDs?

To explore these three questions, we devised three studies. To our knowledge, there was not a technique or tool that allows rendering a 2D view inside the Stereoscopic display, like VR. Thus, we developed one in-house. In the next section, we present this technique, which creates a customizable 2D Frame, which we dubbed PlaneFrame.

Section 8.4 PlaneFrame – 2D in VR

To study the effect of a 2D screen-like view for HMDs for FPS VR games on performance, level of immersion, and degree of control, we developed PlaneFrame (PF). PF displays a copy of the 3D VE within a user's field-of-view (FOV) through a 2D rectangular plane (see Figure 8-5and Figure 8-6). When navigating in the VE, the 2D plane appears in front of the user at a fixed distance (see Figure 8-6a). The user can perceive the 3D VE through the 2D plane, like looking at 3D views through a 2D display. The main difference between PF and the gazing point technique [190] is that PF carries a 2D copy of the 3D VE region at which the user is looking. As such, with

PF, there is no blind spot or hidden peripheral regions. This technique also differs from Slave visualization [23], a technique not aimed at changing the whole VR environment but at sharing 3D views among VR users. As we show in Experiment C, the size of PF is configurable. It also works as a fixed gazing point to help avoid or minimize SS.

To make PF easily configurable and fit diegetically in different kinds of application scenarios, we developed two variations, called Mask and 2D-3D Extreme Mask.

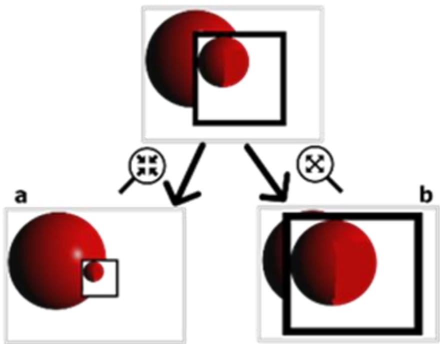

Figure 8-5 The concept used in the PlaneFrame technique and the configurable size of its FoV. (a) A smaller-sized 'mask.' (b) A bigger-sized mask.

Section 8.4.1 Mask View

A 'Mask' can have an area (A) that occupies between 2.5% and 100% of the FoV (height × width) of the HMD (see Figure 8-5 and Figure 8-6). Mask is a rectangle plane in which the center is in position (0,0), with a given height (H) and width (W). We define the PF FoV (PFOV) based on H and W, as given in equations 1 to 4, where *s* is the desired size of PF.

Equation 8-1 Constant calculation based on the percentage of the view to be covered

$$k = \sqrt{s}, 0.025 < s < 1 \quad \textbf{(1)}$$

Equation 8-2 Height of the frame

$$H = height \times k \quad \textbf{(2)}$$

Equation 8-3 Width of the frame

$$W = width \times k \quad \textbf{(3)}$$

Equation 8-4 Frame dimensions

$$P_{FoV} = H \times W \qquad \textbf{(4)}$$

Section 8.4.2 2D-3D Extreme Mask

2D-3D Extreme Mask View (or simply Extreme Mask), like the Mask, would allow the user to see the area in front of him/her and converted the 3D stereoscopic view into a 2D view. The size of PF was also configurable. However, in this version, the user can only see the area within the Mask, and the 3D VE is hidden (see Figure 8-6c and Figure 8-6d). One main difference of Extreme Mask is that it was continuously activated, whereas, in the other versions, it was only activated during motion. This could lead to a more immersive experience because it would remove all external distractions and changes in view.

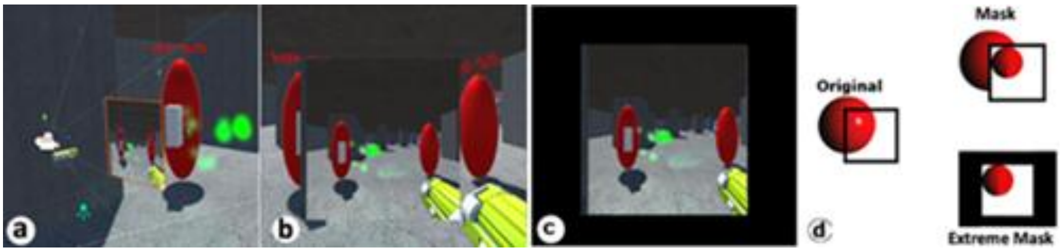

Figure 8-6 (a) PlaneFrame (PF) is a technique that displays a copy of the 3D VE on a rectangular plane which is placed in front of the camera (or users' view). (b) A demonstration of the PF Mask version used in a first-person shooter (FPS) game. (c) A demonstration of the PF Extreme Mask version used in the game. (d) Illustrations of the concept of Mask and Extreme Mask: they both display a screenshot of the 3D VE within a user's field-of-view through a 2D rectangular plane. Mask keeps the 3D VE behind the plane, whereas Extreme Mask hides the 3D VE completely.

Section 8.5 Game Environment and Experiment Metrics

We conducted three experiments to understand how the PF would affect our participants' level of immersion, SS, and gameplay performance. Experiment A collected keyboard data and compared different views, including unmodified VR 3D view and regular 2D monitor. In Experiment B, we compared PF with another existing technique popular in FPS games, namely "Gazing point" [168]. Experiment C evaluated the performance of three variations of the PF to help us understand in more depth the effect of different sizes of PF on induced SS, user performance, and immersion. Experiments A and B are also used to understand the effects of controller types (i.e., keyboard and controller) for FPS games in VR from a between-subjects perspective.

The FPS game used as a testbed for our experiments was developed in-house. This allowed us to avoid or minimize any co-founding factors and to implement the different versions of PF and game to fit the specific controllers and displays. In the game, players have to navigate rapidly while performing other actions typical of an FPS game [235]. Next, we describe the game environment.

Section 8.5.1 Game Environment

The goal of the game was for players to traverse a maze with hallways, open spaces, and chambers and try to survive and destroy (kill) easily distinguishable non-player characters (which we called adversaries). There were three types of adversaries with three different behaviours. One adversary, Sniper, was stationary. The second, dubbed Patroller, could only walk along a fixed route. The third adversary, the Hunters, could hunt the player (based on an AI algorithm provided in [263]). To make them easily distinguishable, they were presented in three different colours: green, blue, and red (see Figure 8-7). The adversaries were positioned throughout the maze, progressively increasing in numbers until the final chamber (see Figure 8-8).

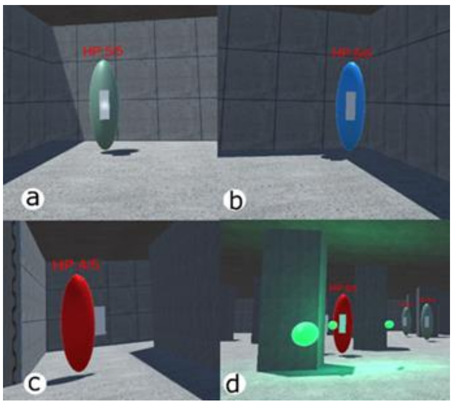

Figure 8-7 Three kinds of adversaries or enemies present in the game environment: (a) green adversary, a Patroller; (b) blue adversary, a Sniper; (c) red adversary, a Hunter; and (d) a player is inside the main chamber of the maze which contains several adversaries of each kind.

Figure 8-8 Bird's-eye view of the maze used in the experiment. A player must eliminate all adversaries going from point A to point B (the final chamber).

The maze was designed with tall grey walls and turning points, which forced the player to move in a non-straight path. This design slowly introduced the adversaries to the players. The lack of visual cues was meant to reduce path memorization, and the turns were meant to assess the occurrence of SS, as the act of rotating around the x-axis tends to increase the levels of SS during gameplay [228], [264]. This kind of turning also allows us to verify the accuracy of the different controller types (see next section). Like any FPS game, there were many positions where the player could hide, but to finish the game, the player had to reach the end of the maze (see Figure 8-8). There were no ambiguous paths or splits; thus, all the players would follow roughly the same path.

Section 8.5.2 Evaluation Metrics

For all the experiments, we use the SSQ [106], and IEQ [176]. The SSQ contains three parts, measuring the level of Nausea, Dizziness, and Oculomotor issues. These three parts are combined to give the overall weighted level for SS (Total Severity). Each symptom uses a 5-point Likert Scale, ranging from least severe to most severe. Each symptom might count points towards more than one part.

The scoring system for the game was as follows: when a player gets hit, he or she loses 10 points; when a player hits a target, he or she gets 10 points. Negative punctuation is possible. Based on this system, we define Accuracy (Acc) as the number of shots that hit the target divided by the number of shots performed by the player. The formula for Accuracy can be found in equation 5, where i is the total number of shots and j is the shots that did not find a target

Equation 8-5 Formula to calculate accuracy shots, hits divided by total

$$Acc = \frac{i-j}{i} \text{Acc} = \text{(i-j)/i} \ (5)$$

Even though a player could be killed, the player's health points were set to be able to endure at least 4 minutes of straight shooting. This life span guaranteed that the players could advance through the maze long enough to allow collecting meaningful results. The IEQ also uses a 5-point Likert Scale for each question. In the end, all questions are summed up.

Section 8.6 Experiment A

Experiment A was designed to (1) compare how an immersive environment compares to a regular computer screen; (2) determine the influence of a 2D view; and (3) collect the gameplay data for when players use the keyboard only. In short, we compared induced SS and user performance with or without PF when playing an FPS game in VR HMD systems. We also included a desktop version as another comparative baseline. Thus, this experiment evaluated three versions: conventional PC display (CD), HMD VR display without PF (RVR), and HMD VR display with PF (VRPF). For this experiment, we only used Extreme Mask.

Section 8.6.1 Participants

We recruited 18 participants (six females; 12 males) from a local university. They had an average age of 18.89 years (*SE* = 0.99), ranging between 17 and 21 years. All volunteers had a normal or corrected-to-normal vision, and none of them declared any history of colour blindness or other health issues, physical or otherwise. Sixteen participants (88.9%) had experience with VR systems before the experiment. Three participants (11.1%) declared having already felt a certain degree of sickness when playing FPS games in the common types of display (laptop, desktop, or TV).

Section 8.6.2 Apparatus

We used an Oculus Rift CV1 as our HMD, as it is one of the most popular off-the-shelf VR devices. The HMD was connected to a desktop with 16GB RAM, an Intel Core i7-7700k CPU @ 4.20GHz, a GeForce GTX 1080Ti dedicated GPU, and a standard 21.5" 4K monitor. We used a mechanical keyboard and a high-resolution gaming mouse as input instead of the traditional Oculus Touch, which is one factor we are collecting data in this experiment.

As a side note, even though players could not see the keyboard during the VR versions, demographics indicated familiarity with keyboard and mouse input when playing FPS games. Nevertheless, to diminish the likelihood of it being a co-founding factor, the player was not required to move their fingers to different keys on the keyboard after the initial positioning. Furthermore, it is essential to note that the position of the monitor was calculated to emulate the size of the VRPF, in order to eliminate FOV as a co-founding factor. Our VRPF configuration in Unity was rotation: (90,180,0) scale: (0.1, 0.1, 0.1) and position: (0, 0.687, 1.663).

Section 8.6.3 Experimental Procedure

On arrival, each participant was assigned a specific order of the three versions in which he or she would play the game. The order was counterbalanced with a Latin Square design approach to mitigate carry-over effects. The participants then filled in a pre-experiment questionnaire that collected demographic and past gaming experience information.

After that, participants were presented with a demonstration of the Oculus Rift and a virtual environment without walking to get them acquainted with the VR HMD and setup. Next, we introduced the input device and the rules of the game to each participant.

Participants played the three versions in a pre-defined order and, after completing each version, were asked to answer the questionnaires mentioned above. The participants were required to rest (and as much as they wanted) before they could go to the next version.

Section 8.6.4 Results

The data were analysed using both statistical inference methods and data visualizations. We conducted a Shapiro-Wilk test to check the normality of the data. For those that were classified as normal, we used parametric tests; for the others, we used non-parametric ones.

For normally distributed data, we conducted Mauchly's Test of Sphericity. We also employed Repeated Measures ANOVA (RM-ANOVA) using Bonferroni correction to detect significant main effects. If the assumption of sphericity was violated, we used the Greenhouse-Geisser correction to adjust the degrees of freedom. Partial Eta-Squared is declared for the ones with significant main effects. We conducted Friedman tests for non-normal data. When there was a detectable significance, we ran separate Wilcoxon signed-rank tests on the different combinations to pinpoint where the differences occur.

To make a consistent presentation of the plots, in all of them, we used separate colours and patterns to represent the three versions: CD (blue-checked), RVR (red-diagonally stripped), and VRPF (green-plain). The data were presented with outliers for a deeper understanding of the results.

Section 8.6.4.1 Simulator Sickness Questionnaire (SSQ)

Figure 8-9 summarizes the results of the SSQ data and shows that RVR presented the most significant change in the level of SS than the other versions. The data did not follow a normal distribution. The Friedman test showed that there was a statistically significant difference in Oculomotor, Dizziness, and Total Severity (hereafter Total) based on which version the game was being played, $\chi2(2) = 11.815$, $p = .003$, $\chi2(2) = 16.618$, $p = .000$, and $\chi2(2) = 16.395$, $p = .000$ respectively. Nausea did not appear to differentiate significantly according to the test $\chi2(2) = .323$, $p = .851$.

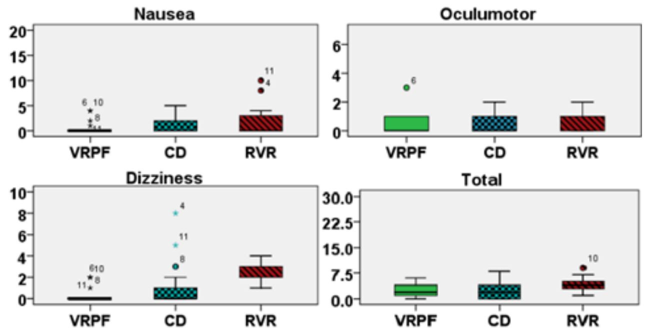

Figure 8-9 The results of SSQ based on the four sub-scales in Experiment A. RVR presented the most significant change in the level of SS compared to the other versions. VRPF and CD did not differ significantly between themselves.

During our post-hoc analysis with Wilcoxon signed-rank tests, we observed the Median (IQR) Oculomotor levels for CD, VRPF, and VR were 0, 0, and 1, respectively. After correction there were no significant differences between them, except between RVR and CD ($Z = -2.333$, $p = 0.020$). The lack of significance is clear between CD and VRPF ($Z = 0.000$ $p = 1.000$), but less so between RVR and VRPF ($Z = -1.684$, $p = 0.092$).

The post hoc analysis further showed significant differences in Dizziness between VRPF and RVR ($Z = -3.299$, $p = 0.001$). However, there was no significant difference between VRPF and CD ($Z = -1.476$, $p = 0.140$) or between RVR and CD ($Z = -1.764$, $p = 0.078$). Their Median (IQR) were CD = 0, RVR = 2 and VRPF= 0.

What is important to note from this analysis is that Total diverged significantly between RVR and CD ($Z = -2.759$, $p = 0.006$), and between RVR and VRPF ($Z = -2.831$, $p = 0.005$). The other pair did not have significant difference ($Z = -.224$, $p = .823$). 2, 4, and 2 were the Median IQR of CD, VRPF, and VR, respectively.

Section 8.6.4.2 User performance

Figure 8-10 shows the data for Score, Time, and Accuracy across the three versions. The Shapiro-Wilk test revealed that the Score measurements did not follow a normal distribution, but Time and Accuracy did. The RM-ANOVA suggested that there were significant main differences in Accuracy ($F_{2, 34} = 7.304$, $p = .002$, $\eta_p^2 = .313$) among the three versions. The post-hoc analysis showed that Accuracy in FPS-VR was significantly lower than the other two versions ($p=.013$ and $p=0.14$). No other significant relations were found for Time ($F_{2, 34} = 1.448$, $p = .250$); on average each condition lasted about 251 seconds (4.2 minutes).

The Friedman test revealed no significant difference for Score $\chi^2(2) = 3.254$, $p = 0.197$.

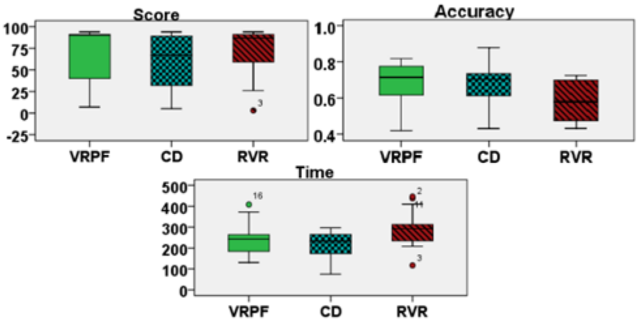

Figure 8-10 The results of Score, Accuracy, and Time. Accuracy in RVR was significantly lower than the other two versions.

Section 8.6.4.3 Subjective Immersion

Figure 8-11 shows the data for Overall immersion across the three versions. The RM-ANOVA suggested that there were significant main differences ($F_{2, 32} = 3.852$, $p = .032$, $\eta_p^2 = .194$) among the three versions. Post-hoc analyses using Tukey's HSD indicated that Overall immersion was greater on VRPF when compared to VR ($p = .030$).

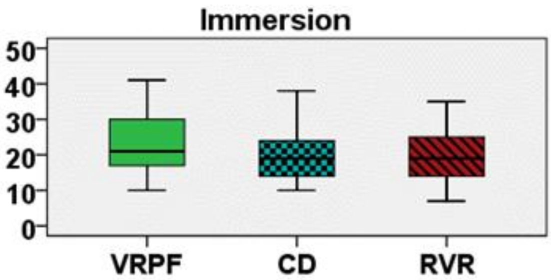

Figure 8-11 The results of Overall immersion. Overall, immersion did differ significantly in VRPF compared to the other versions.

Section 8.6.5 Discussion

The results show that PF was a promising approach to lower the level of SS for FPS VR games. Results from SSQ showed that, with PF, participants had significantly lower levels of Oculomotor, Dizziness, and Total Severity than playing in the original, unmodified VR environment. It seemed that by projecting the 3D VE onto the 2D replica, we could decrease players' level of SS significantly while allowing them to navigate in the VE and did not lead them to lose the context of the environment. Our results also showed that when using PF, the VR environment might lead to similarly low levels of SS, as does the CD. This offers the possibility that users could play or, more generally, navigate in VEs with a low level of SS in VR systems with PF.

Concerning user performance, the results indicated that PF led to a significantly higher accuracy than the original VR version and had similar performance with the PC version. We could infer that because, with a lower SS, the players were able to shoot at the targets more accurately in the FPS VR game. This result is significant because this could mean that the lower SS could lead to a more precise view of the elements of the environment, and this could have applications outside of FPS games.

Furthermore, what was surprising was that VRPF was statically more immersive than RVR. We had expected that the free nature of RVR to be completely immersive and allow players to forget they were in a VE. Based on the results, we were able to identify two hypotheses for why RVR suffered in this experiment:

HA1. Playing with the keyboard did not allow the users to be fully immersed in the experience.

HA2. Given the total lack of other stimuli in the VRPF, the players got solely focused on the game and thus felt immersed.

In summary, when compared to a regular FPS in VR, PF presented positive results in the reduction of SS and supported users to achieve high accuracy. Since the standard display is currently the baseline for most FPS games, we can infer that PF represents a viable technique to reduce SS in VR while allowing users to achieve performances that are comparable to the current best environment.

Section 8.7 Experiment B

Experiment B aimed to compare PlaneFrame (PF) and Regular FPS VR with a diegetic technique that can potentially improve Simulator Sickness and increase Accuracy during FPS gameplay in VR HMD systems: gaze point [168]. Three versions were compared: HMD VR display without PF (RVR), HMD VR display with PF (VRPF), and Gaze, which gave the users a point to stare at continuously (see Figure 8-12). Our Gaze point was a circular green ring (radius 15 Unity Units) located at the center of the camera. It was permanently on the screen, and it followed the player's head movement. We clarify that the Gaze point was not used as a game control in any capacity.

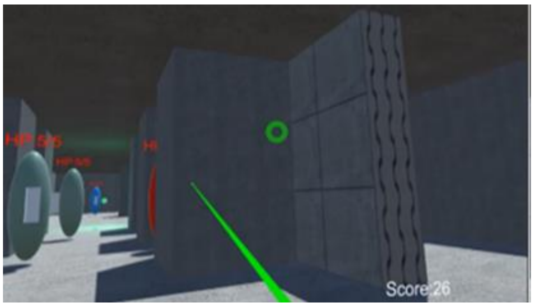

Figure 8-12 A screenshot during gameplay using the Gaze technique which provides a gazing point represented by the green ring. The player could stare at the point during navigation to mitigate SS effect and potentially improve aim.

Section 8.7.1 Participants, Apparatus, and Experimental Procedure

Another 18 participants (equally divided between males and females) were recruited from a local university with an average age of 20.0 years (SE = 1.90), ranging from 18 to 27 years. All volunteers had a normal or corrected-to-normal vision, and none declared any history of colour blindness or health issues. Thirteen participants (72.2%) had experience with VR systems. Two (11.1%) reported feeling sick during regular FPS gameplay. We used the same VR device and the desktop PC configuration as in the first experiment. Participants also followed a similar procedure to complete the three versions offered in this experiment.

Section 8.7.2 Results

We employed a similar procedure to analyse the data as the first experiment. We used the following colours to represent the three versions in the plots: Gaze (blue-checked), RVR (red-diagonally stripped), and VRPF (green-plain).

Section 8.7.2.1 Simulator Sickness Questionnaire (SSQ)

Given that the SSQ data were not normally distributed, we performed a series of Friedman tests, which showed that there was a statistically significant difference in all sub-scales and Total (see Table 8-6). The results of the SSQ in terms of sub-scales and Total are summarized in Figure 8-13.

Table 8-6 Summarized results of Friedman tests of SSQ, significant differences in all subscales

Scale	χ2(2)	p
Nausea	16.618	.000
Oculomotor	16.745	.000
Dizziness	11.577	.003
Total	11.789	.003

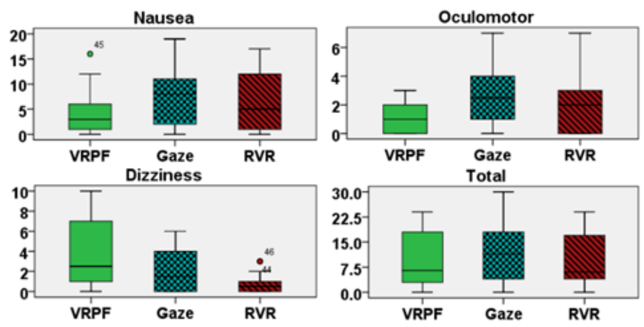

Figure 8-13 The results of SSQ concerning the four sub-scales of Experiment B. VRPF presented positive results in Nausea and Oculomotor but underperformed in Dizziness.

The post-hoc analysis with Wilcoxon signed-rank tests showed that the Median (IQR) Nausea levels for the Gaze, VRPF, and RVR were 8, 3, and 5, respectively. After correction there were significant differences between VRPF and Gaze (Z = -3.428, p = .001) and VRPF and RVR (Z = -1.960, p = .050), but not between Gaze and RVR (Z = -.828, p = .408).

Post-hoc analysis of the Oculomotor component was similar to that of Nausea, revealing significant differences between VRPF and Gaze (Z = -3.325, p = .001) and VRPF and RVR (Z = -2.183, p = .029), but not between Gaze and RVR (Z = -1.592, p = .111). Median (IQR) Oculomotor levels for the Gaze, VRPF, and RVR were 2.5, 1, and 2.

Interestingly, all comparisons presented significance in the post-hoc analysis of the Dizziness component: VRPF and Gaze (Z = -1.968, p = 0.049) and Gaze and RVR (Z = -2.365, p = 0.018). The strongest effect was between RVR and VRPF (Z = -3.051, p = 0.002). Median (IQR) Dizziness levels for the Gaze, VRPF, and RVR were 1.5, 2.5, and 0.5.

Finally, our post-hoc comparison showed that Total diverged significantly between Gaze and VRPF ($Z = -3.190$, $p = .001$). the other pair combinations did not have significant difference: RVR Gaze ($Z = -1.709$, $p = 0.087$) and VRPF and VR ($Z = -1.156$, $p = .248$). The Median IQR were Gaze = 11.5, RVR = 6 and VRPF= 6.5, respectively.

Section 8.7.2.2 User performance

The RM-ANOVA indicated that there were no statistically significant differences in Accuracy ($F_{2, 34} = 2.605$, $p = .089$) or Time ($F_{2, 34} = .493$, $p < .615$); on average each version lasted 280 seconds (4.6 minutes). However, the Friedman test revealed significant differences for Score $\chi^2(2) = 9.662$, $p = 0.008$. Figure 8-14 shows the results.

Post-hoc analysis revealed that the Score in VRPF was significantly lower than RVR ($Z = -2.369$, $p = .018$) and Gaze ($Z = -3.247$, $p = .001$), while the Median (IQR) Score levels for the Gaze, VRPF and RVR were 76.5, 60.5 and 74.5. Gaze and RVR did not present significant differences ($Z = -0.327$, $p = .744$).

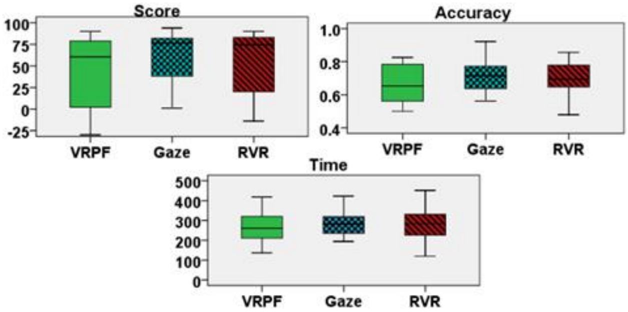

Figure 8-14 The results of Score, Accuracy, and Time for Experiment B. VRPF performed worse than the two other versions in Score.

Section 8.7.2.3 Subjective Immersion

Figure 8-15 shows the data for Overall immersion across the three versions. The RM-ANOVA showed that there were significant main differences in Overall immersion ($F_{1.508, 25.635} = 7.849$, $p = .004$) among the three versions. Post-hoc analysis showed that Overall immersion in Gaze was significantly higher than in the VRPF (p=.04). VRPF and RVR (p=.02) also presented significant differences. Gaze and RVR (p>.05) did not present significant differences.

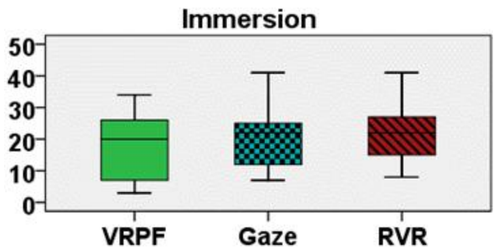

Figure 8-15 The results of Overall immersion in Experiment B. VRPF presented significant differences compared to the two other versions.

Section 8.7.3 Discussion

From our analysis, we observed a clear difference among the three versions regarding induced SS. The Total Severity was significantly different between VRPF and Gaze, but not between VRPF and RVR. These results are somewhat surprising. We expected Gaze to perform better than no technique at all. VRPF performed visibly better in most subscales except in dizziness. The results suggest that the FP 2D view helps to mitigate much of the symptoms of Simulator Sickness but can also cause disorientation, which should be investigated further (see Experiment C, in the next section).

The counterpoint to this distinctly lower level of Simulator Sickness is the lower levels of immersion generated. This result is especially interesting since, in Experiment A, we observed VRPF generating the highest levels of immersion. This divergence might be caused by the influence of the controller. According to the results, if the users use a keyboard, they will feel more immersed in a 2D view while, when using the controller, using a 3D view will likely increase their level of immersion.

This effect extends further as VRPF has led to lower Score results and can have implications on the controller aspect as well. The results together seem to imply that the type of controller could affect the speed of navigation, which in turn could have a disorienting effect when looking at 2D and 3D views. The effect could be enhanced further in the latter. Disorientation could invariably affect performance and lower scores.

Section 8.8 Experiment C

The goal of this experiment was to compare different variations of PF regarding induced SS, user performance, and immersion when playing an FPS game in VR HMD systems only. We varied the size of PF to 30% (SM) and 95% (BM) of the HMD's FoV. We were interested in exploring if the size of PF would influence the dependent variables. Therefore, in this experiment, we compared three versions in total: VRPF, SM, and BM.

Section 8.8.1 Participants, Apparatus, and Experimental Procedure

We recruited another 18 participants (five females, twelve males, and one non-binary) from a local university with an average age of 19.7 years (*SE* = 1.73), ranging from 18 to 24 years. All volunteers had a normal or corrected-to-normal vision, and none declared any history of colour blindness or health issues. Thirteen participants (72.2%) had experience with VR systems. Three (16.7%) declared feeling sick during FPS gaming sessions with regular displays. We used the same VR device, and the desktop PC as in Experiment A. Participants followed a similar procedure as the other experiments to complete the three versions for this experiment.

Section 8.8.2 Results

We employed a similar procedure to analyse and represent the data as the first experiment. We used the following colours and patterns to represent the three versions in the plots: BM (blue-checked), SM (red-diagonally stripped), and VRPF (green-plain).

Section 8.8.2.1 Simulator Sickness Questionnaire (SSQ)

Results of the SSQ in terms of sub-scales and total are summarized in Figure 8-16. Because of the non-normality of the data, we performed a series of Friedman tests, which showed that there was a statistically significant difference in Oculomotor and Dizziness, $\chi 2(2) = 14.292$, $p = 0.001$ and $\chi 2(2) = 8.667$, $p = 0.013$ respectively. On the other hand, Nausea and Total did not appear to differentiate significantly according to the test $\chi 2(2) = 1.755$, $p = 0.416$ and $\chi 2(2) = 5.729$, $p = 0.057$, respectively.

During the post-hoc analysis with Wilcoxon signed-rank tests, we observed the Median (IQR) Oculomotor levels for the VRPF, BM, and SM were 0, 2, and 1, respectively. After correction there were significant differences between VRPF and SM (Z = -2.365, p = .018) and VRPF and BM (Z = -3.169, p = .002), but not between BM and SM (Z = -1.512, p = .131). We further observed significances in Dizziness, VRPF and SM (Z = -2,337, p = .019) and VRPF and BM (Z = -2.209, p = .027), but not for BM and SM (Z = -0.686, p = .492). Median (IQR) Oculomotor levels for the VRPF, BM, and SM were 3, 2 and 1.

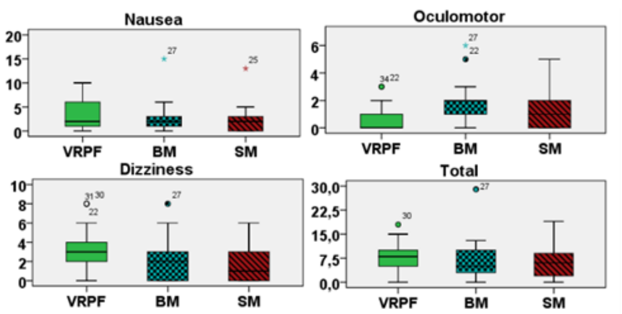

Figure 8-16 The results of SSQ for four sub-scales in Experiment C. VRPF generated more Dizziness than BM and SM, but less in Oculomotor distress.

Section 8.8.2.2 User Performance

Figure 8-17 shows the data for Score, Time, and Accuracy across the three versions. Like Experiment A, Shapiro-Wilk revealed a lack of normality for Score, but Time and Accuracy data were normal. The RM-ANOVA suggested that there were no significant main differences in Accuracy ($F_{2, 34} = .677$, p = .515) among the three versions. No significant relations were found for Time ($F_{2, 34} = 0.26$, p = .974) either. On average, each condition lasted 307 seconds (5.1 minutes).

Unlike in Experiment A, the Friedman test revealed a significant difference for Score χ2(2) = 8.121, p = 0.017. The Wilcoxon test revealed that the difference was significative between BM and SM (Z = -2,182, p = .029) and between BM and VRPF (Z = -2,535, p = .011). Their medians were BM = 74.9, SM = 69.8, and VRPF = 61.2.

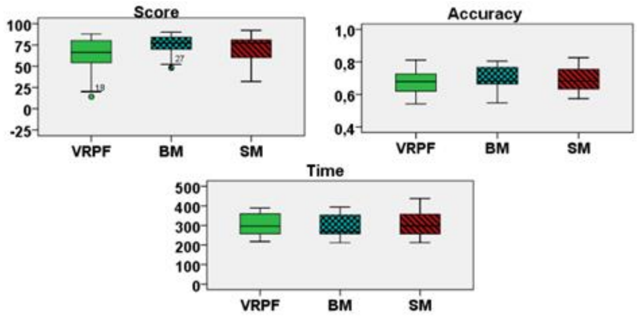

Figure 8-17 The results of Score, Accuracy, and Time in Experiment C. BM presented a higher Score than the other versions.

Section 8.8.2.3 Subjective Immersion

Among the three studied versions, Overall immersion did not present significant differences according to the RM-ANOVA (F2, 34 = .535, p = .591). Figure 8-18 shows the data for Overall immersion across the versions.

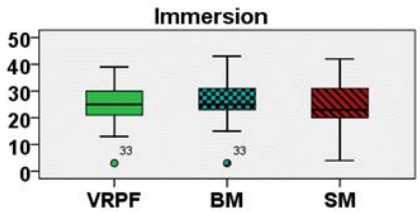

Figure 8-18 The results of Overall immersion for Experiment C. Overall immersion across versions did not differ significantly.

Section 8.8.3 Discussion

In this experiment, we observed that different variations of PF resulted in similar levels of user performance and immersion. BM and SM presented virtually no difference except in the oculomotor component of SS, which was also statistically insignificant, suggesting that our results are not the fruit of the FOV, unlike the work of [225] that used some kind of visual technique to hide only the outer region without a 2D effect.

Both SM and BM presented reduced levels of Dizziness when compared to VRPF but also presented increased levels of Oculomotor discomfort. The darker outer region could have allowed for the eyes to rest. However, this was a liability when associated with the fast movement of the controller.

From the results, it seems that it can be advantageous to keep some level of the 3D view besides the PF view when using a controller because it might aid in the Dizziness factor. If the oculomotor strain is the most significant concern, opting for the darker outer layer can be of more considerable aid, similar to [225].

Since there was no notable difference among these techniques in this experiment, it can be concluded that all of them are somewhat viable for FPS games, even considering the SS. However, considering the results from Experiment B, it may be useful to investigate SM further since it generated less Nausea than VRPF and lower levels of Dizziness as well.

Section 8.9 Between-Subjects Analysis (Experiments A and B)

The goal of this analysis was to compare how vital the controller is for user performance and immersion when playing an FPS game in VR HMD. We compared the keyboard against the recommended VR controller. We were interested in exploring if the controller would influence the dependent variables. Therefore, in this analysis, we compared four conditions based on two inputs and two views: VRPF-Keyboard (1), VRPF-Controller (2), RVR-Keyboard (3), and RVR-Controller (4).

We compared the data of participants from Experiments A and B. However, to avoid co-founding factors, we only compared the data from the first version played by each participant. As such, we have a total of 24 participants, 6 played VRPF and 6 RVR using a keyboard and 12 participants who did the same using the Oculus Touch Controller. We performed two between-subjects analyses using the data, one comparing the differences between keyboard and controller in VR regardless of the view and one only between subjects who had experienced the same view.

Section 8.9.1 Results

We employed a similar procedure to analyse the data as the first experiment, first checking for normality, followed by investigating significant differences with ANOVA for parametric data; otherwise, Kruskal-Wallis and Mann-Whitney was used. We used the following colours and patterns to represent the four versions in the plots: 1 (light green – vertical stripes), 2 (dark green - diagonal stripes), 3 (light red – vertical stripes), and 4 (dark red – diagonal stripes).

Section 8.9.1.1 Simulator Sickness Questionnaire (SSQ)

Figure 8-19 shows the results of the SSQ in terms of sub-scales and Total. Because the data were not normally distributed, the Mann-Whitney test was used and showed that there was a significant difference in Nausea (Mann-Whitney U = 24.5, P < 0.05), Oculomotor (U = 28.0, P < 0.05) and Total (U = 22.0, P < 0.05) based on the input alone. No significant difference was found in Dizziness (U = 66.0, P > 0.05).

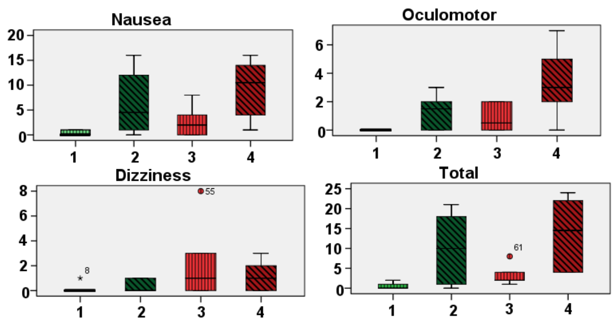

Figure 8-19 The results of SSQ concerning four sub-scales in the between-subjects analysis. The input method was hugely influential.

Comparing inputs within a view Mann-Whitney tests also presented significant differences in Nausea for both RVR (U = 5.0, P < 0.05), and VRPF (U = 6.0, P < 0.05); Oculomotor for both RVR (U = 5.5, P < 0.05), and VRPF (U = 6.0, P < 0.05). Total had similar results for RVR (U = 3.0, P < 0.05) and VRPF (U = 5.5, P < 0.05).

Section 8.9.1.2 User Performance

When analysing input alone, the Kruskal-Wallis H test did not show a statistically significant difference in Accuracy ($\chi2(2) = .654$, $p = .419$), with a mean rank of 13.67 for Controller and 11.33 for Keyboard. However, there was a significant difference in Score ($\chi2(2) = 6.321$, $p = 0.012$), with a mean rank of 8.88 for Controller and 16.13 for Keyboard.

Results from a one-way ANOVA showed no significant effect for Time (F1, 22 = 0.013, p = .911); on average each condition lasted 293 seconds (4.9 minutes). The medians can be seen in Figure 8-20.

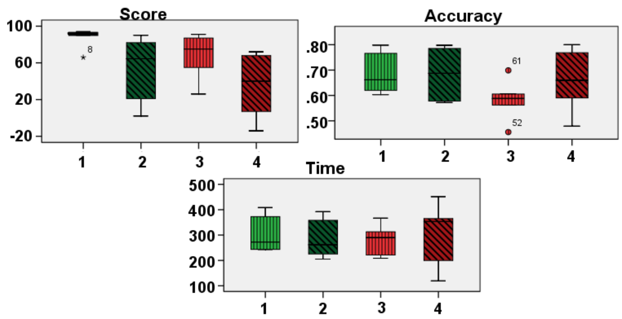

Figure 8-20 The results of Score, Accuracy, and Time in the between-subjects analysis. Keyboard yielded better results in Score; Accuracy relied on the input method matching the view.

The input conditions presented differences in Score when the players were using the VRPF (U = 3.5, P < 0.05 two-tailed) with a mean rank of 8.92 for 1 and 4.08 for 2.

Section 8.9.1.3 Subjective Immersion

When comparing all views, the Kruskal-Wallis H test did not show a statistically significant difference ($\chi2(2) = 4.984$, $p = 0.173$), with a mean rank of 16.75, 7.83, 13.58 and 11.83 for conditions 1-4, respectively. When checking without considering the 2D/3D aspect, the difference was still non-existent (Mann-Whitney U = 40, P > 0.05). However, when comparing input in VRPF, there is a difference in immersion (U = 5.0, P < 0.05). Details can be seen in Figure 8-21.

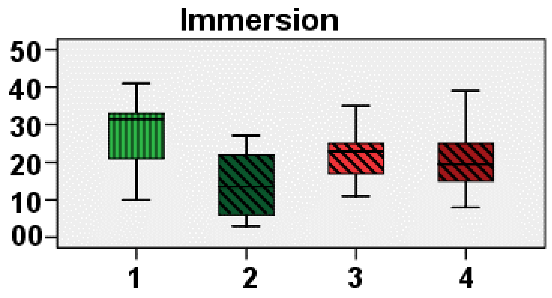

Figure 8-21 The results of Overall immersion for the between-subjects analysis. Overall immersion across versions did not differ significantly.

Section 8.9.2 Discussion

Our results suggest that the kind of controller used can be the most influential aspect of FPS when it comes to SSQ Nausea and Total. The keyboard generated lower levels of Simulator Sickness when compared to the regular Oculus controller when playing the FPS VR game. In addition, the Score measure also presented an improvement when using the keyboard combined with our VRPF, which suggests that the use of controllers might not be beneficial for people who are using a 2D view. Similarly, keyboards might be less precise for those who are using a 3D view.

Interestingly, for Overall immersion, our results did not show significant difference when using the keyboard in RVR. These results somewhat contradict findings from [44], [215], [216]. The cause might be simply due to how the other studies were conducted. For example, in [44], [215], [216], the authors did not count for the cofound variable of view, which the results in Experiments A and B show to affect immersion. Moreover, FPS games might be less susceptible to variations in immersion, as suggested in previous studies (e.g., [222]). *RQ2* is answered with mixed results. Our results support that subjective feelings of physiological background are more affected by the controller than by the view. However, our results suggest that the view can be highly impactful, even if it has an unnatural synergy with the controller.

Our results strongly suggest that using a Keyboard can bring benefits for the overall VR experience. Our results do not support *HA1*—playing with the keyboard is not detrimental to the immersive experience; it yielded a good score performance and was superior in immersion when using a 2D view. Our results support *HA2*—the 2D view can be immersive and it lets the players focus solely on the game; as such, overall, it is not detrimental to the immersive experience. For FPS games, the overall superior playability and the possibility of spending extended periods in the games might make the keyboard a preferred choice. For *RQ1*, our results indicate that the controller seems to have a more significant part in the performance than the view does. However, for the maximum yield of each controller type, it must be harmonized with the adequate type of view, because score performance was the best when the view matched its expected controller type. Based on the answers to *RQ1* and *RQ2*, the answer to *RQ3* appears to be mixed: the stereoscopic 3D view alone is not a significant influence; however, its combination with the control method can be highly significant.

Section 8.10 Recommendations and Potential Applications of PF

Based on our findings, we recommend the use of the VRPF version of our PF associated with a keyboard for VR FPS games. This combination seems to lead players to enjoy an immersive experience while obtaining their best performance scores. Besides, PF players appear to be able to play the game longer because, though they are fully immersed in VR, they are likely to feel less sick.

Figure 8-22 a) Demo of a museum visit using PlaneFrame Mask. The user can appreciate the exposition and travel/look around for a longer time, which is adequate for such a learning environment that requires attention to detail as Simulator Sickness can be reduced. b) Demo of an open house or architectural 3D VR environment. The user can have a complete understanding of the space, due to the linear movement provided by the technique and without having the vision occluded by a gaze point.

In an environment that does not allow the use of a keyboard, we recommend the use of an SM (see Experiment C) because it keeps most of the SS mitigation benefits provided by VRPF, while still letting players immersed and allowing them to achieve high Accuracy.

PF can be inserted diegetically in many games. FPS and Racing game characters can justify the use of the technique through the use of protective gear. In horror games, it can be a device to bring focus to a specific scene.

PF could also be used in applications outside of games that do not require fast responses. For example, educational environments such as museums (see Figure 8-22a) and other forms of VR education and training environments can make use of the technique. It allows students to be immersed for extended periods or let them move more and explore further the environment. Extensive exploration is also useful for other kinds of applications, such as an open-house VE (see Figure 8-22b). In this kind of scenario, the user likely wants and needs a clear understanding of the space. Such clear understanding could be challenging to achieve when users feel sick and want to disengage from the system because further exposition will have adverse effects after every movement.

Section 8.11 Limitations and Future Work

We studied three variations of PlaneFrame (PF). Due to the continuous nature of the mask FoV, we did not explore all possible sizes for the frame. Nevertheless, given that the two sizes we explored did not lead to significant differences in the results, it is likely the case for other possible sizes. Further explorations are needed to ascertain if this is the case.

We explored the use of PF in one specific kind of game (FPS). We chose FPS games because of the fast-moving nature and constant change of viewpoints; they are often regarded as ideal environments to investigate SS mitigation techniques (see [23], [265]). These techniques have also been used outside of the context of VR HMD [265]–[267]. Our results could be applied to other VR applications with less demanding navigation tasks. Nonetheless, it will be useful to explore the applicability and use of PF in other types of applications. For future work, we would like to explore if we can achieve the same results in other types of games.

As the main aim of our study was to identify the effect of view mode and controller type on people's level of immersion, SS, and performance for VR FPS, we have been able to develop a viewing technique that led to improved results in VR FPS. In the future, it will be useful to explore the combination of PF with other types of devices and techniques that also aim to reduce SS [228], [236]. We can evaluate users' preferences and assess their combined effect. Furthermore, we can analyse other tasks like locating objects in 3D VE and spatial memory recall of the location of items in these environments [82].

Section 8.12 Summary and Remarks

This research presents an in-depth exploration of the effect of different view dimensionalities and input methods in VR gaming, with a focus on first-person shooter games. We studied how such interactions affect immersion, accuracy, and Simulator Sickness. To do so, we developed a concept technique PlaneFrame (PF) to study 2D/3D views in VR. In the first of three studies, we investigated the 2D/3D views and display using a keyboard in a first-person shooter game. The results showed that with PF users achieved better performance and had lower SS than playing the game in the original, unmodified VR display. In our second experiment, we compared PF with other VR conditions using a controller (gaze pointing and original, unmodified VR display). In our last experiment, we presented and evaluated three variations of PF. In our final analysis, we observed that even though view did play a factor in all the component the controller type was more influential.

Our results suggest that PF is useful in reducing the levels of SS in VR while maintaining and possibly boosting players' performance. It is a novel technique that is simple to implement with nearly no associated computing overhead. It also does not require additional external devices, is not intrusive, and is relatively easy for users to adapt to and use it. Keyboards are preferred when there is no need for fast turning, and

2D is acceptable. We conclude that further investigations should focus on controllers rather than viewing, because of its considerably stronger influence in the diverse aspects of VR.

Chapter 9 Discussion, Conclusion, and Future Work

Section 9.1 Summary

This thesis examined the factors that influence enjoyment in VR games in HMD Devices. The first part of the thesis outlined the literature of HMDs, existing research on the topic of enjoyment in VR HMDs, and proposed areas of investigation: (1) Emotions in virtual characters; (2) aesthetical realism in virtual-characters; (3) viewing perspective in-game; (4) measuring subjective feelings in VR games; (5) preventing Simulator Sickness without disrupting enjoyment.

In the remainder of the thesis, these areas of investigation were addressed in the context of different research studies. The studies were grouped and presented by theme.

The first group of studies involved exploring how the appearance and emotional expression in non-playable characters affected the gaming experience. In this group of studies, the first was focused on the expression of emotions through a virtual character and their effect on the participants. To do so, we had an experiment in which players interacted with a virtual girl whose emotions were mediated through a state-of-the-art algorithm and expressed them through combinations of different mediums, namely, voice and facial expression. We concluded that not all emotions are translated equally through a virtual character. Furthermore, we observed that adding a single modality of emotional expression is enough to increase the effect it has on players; however, adding more modalities leads to diminishing returns. That is, the increase in realism and effect of having either voice or facial expression over having neither is greater than the increase of having both over just one of them.

The second and third studies proposed to evaluate the effects of an aesthetically realistic audience in a tension-prone environment, in these cases that was a presentation simulator. We developed a presentation simulator with feedback from the audience. We then had participants simulate they were presenting to a 3D scanned audience, to a stock model audience, and to no audience at all. First, we observed that having a virtual audience on itself increased nervousness when compared to no audience at all. Then, the continuation of our study on how to better implement the simulation suggested that a 3D scanned audience is not inherently better than one using standard 3D stock avatar models to engage players.

The next group of studies focused on viewing perspective during Virtual Reality gaming, its effects, and how to measure them. The first study aimed to analyse how the different viewing perspectives affected preferences and other subjective factors. In this study, we used a fully functional racing game for our analysis. We noticed that both viewing perspectives in VR caused sickness; however, first-person perspective could be more sickening. Moreover, there was not much difference in overall immersion, and there was no clear preference among versions. The second study of this group was divided into three parts; we set up to analyse subjective metrics associated with psychophysiological ones under the same context as the previous study. We aimed to explore connections and identify possible measurements psychophysiological measurements that could be used to gauge different levels of enjoyment when experiencing the different perspectives. In the first part of this study, we observed discrepancies between existing psychophysiological metrics and subjective questionnaires because the subjective questionnaires found significant differences among versions, which were not present in the EEG metrics. In the second part of the study, when analysing the psychophysiological metrics as a time series, we observed that being close to crossing the finish line led to peaks in one of the measurements used, but hardly any other patterns were confirmed or found. And finally, in the third part, we observed some physiological differences in VR and a common display but not differences in VR versions. Furthermore, we observed Beta and Gamma brain waves as potential indicators of immersion.

The last study of the thesis focused on techniques to mitigate Simulator Sickness without interfering with gameplay and in-game mechanics. The study began by identifying existing techniques and trends in a systematic literature review followed by a meta-analysis. We identified six main components commonly used to avoid Simulator Sickness in Virtual Reality. Among these components, one of the most common was Field-of-View restriction. Our analysis observed that it (Field-of-View restriction) presented mixed results, even though it is often regarded as a good technique to mitigate Simulator Sickness. On the other hand, Blur, which is a popular technique as well did present positive results. Finally, we observed a lack of variation among techniques, especially regarding the use of 2D elements in VR. We proposed to relieve this lack of variation in the subsequent study in which we presented a novel technique we dubbed PlaneFrame to mitigate Simulator Sickness. The main concept of PlaneFrame is to present a picture of the environment in front of the players as they move in the game. PlaneFrame can be adjusted to best match the players' preferences. In our experiments, we used sizes between 30% and 95% of the screen. We compared it to no mitigation technique at all and another technique in an FPS game in VR. PlaneFrame presented good results to mitigate Simulator Sickness while being minimally invasive in-game performance and immersion. In some cases, it was almost as good at mitigating Simulator Sickness as having a regular computer monitor; however, in this instance, a keyboard was necessary.

In this final chapter of the current thesis, we provided a comprehensive summary of our investigation and their partial results. Now, we will present a discussion of the main topics we studied followed by some design guidelines for future VR games, and finally a list of future works to deepen our understanding of enjoyment in VR.

Section 9.2 Discussion

Throughout the several studies performed, we could observe how different factors affect enjoyment in Virtual Reality games.

Based on our first group of studies, we observed that the graphical style of the character is not as influential as its ability to express emotion in an understandable way. This is further corroborated by the work of Slater et al.[118]. In their work, the audience was a low-polygon one who could express rudimentary but intelligible feelings. Moreover, their results suggest the different kinds of feedback were hugely influential, even though they only added visual feedback. This is further corroborated by our study, facial expression was requested by players from one of our experiments, and well-regarded in the other. In addition, in our study, adding channels to express emotion suffers from diminishing returns. A character expressing emotion through one medium increases considerably the affect it has on the player; however, the second addition does not increase the affect as much.

Throughout our experiments, we could observe that one of the greatest challenges to enjoyment in Virtual Reality is Simulator Sickness. It was present in all the experiments we performed when movement happened. To avoid confounding factors, participants were stationary in our experiments regarding emotions and aesthetics in virtual agents. In our subsequent experiments, we observed that even though participants did suffer from Simulator Sickness in 3PP during a racing game, the level of sickness was lower than that of participants who experienced 1PP. Though our study had a limited sample size which warrants caution for the generalization of the results, it is important to notice that it is aligned with similar independent previous research [14]. Because many existing games naturally have 1PP/3PP camera toggle as one of their default options, this should a relatively inexpensive and well-accepted technique to mitigate Simulator Sickness in VR games.

Additionally, our work indicates that Simulator Sickness can be mitigated through several other techniques and that, though no complete prevention method has been found, many obtained very positive results. One limitation of our analysis is that we focused only on techniques that do not alter movement in VR or require extra hardware. We focused on techniques that change how the virtual world is presented. However, our analysis was still able to observe diverse techniques within these parameters, and we were able to propose a new mitigation technique within this vein.

Our experiments with our new proposed technique showed that it is a good mitigation technique for Simulator Sickness. The experiments also revealed that performance and immersion are not negatively affected. Some researchers have found no particular reason to use a keyboard instead of the controllers; however, they often did not count for the 2D effect caused by our technique [44], [46]. Thus, our study presented a limitation in existing studies of Simulator Sickness. That is, it revealed the importance of counting with the interaction factors when determining the usefulness of a technique against a regular monitor baseline.

Moreover, as we experienced measuring the game experience through EEG we observed great challenges. First, we observed that measuring the game experience through commercial EEGs is extremely tricky with VR HMDs, especially because of the fitting of the two devices. Additionally, the current metrics do not associate properly with the subjective measurements. We were able to identify some correlations and tendencies within the same participant. However, more analysis with more participants is necessary to reach definitive conclusions regarding how to best use EEG to monitor gameplay in VR. Similarly, the effects of HR and Body Score were not detectable through standard linear techniques. However, they could be used to identify which version the player was simulating, which suggests that there is a difference.

Finally, after studying the current literature and having made these novel observations, we propose a series of design guidelines (summarized in Table 9-1 below) for games in Virtual Reality, based on their gender and objective.

Section 9.2.1 Guidelines

Section 9.2.1.1 Virtual Agents

We recommend that in-game virtual agents have expressive faces or voices and that aesthetical realism be implemented but only after the other characteristics have been satisfactorily implemented. This is especially important for games classified as Social Simulators, Sandbox, Role-Playing games, or any others that require the character(s) to be involved with the player. This aspect of gameplay can be often overlooked in Racing and Trivia games, in which external characters have little involvement with the player. For FPS games, the use of emotional characters needs a deeper analysis according to the aim of the game; however, for competitive adult settings, emotional expressions might be interesting, our results suggest they increase competitiveness.

Section 9.2.1.2 Locomotion

We advise that, for long-term exposure games in which dizziness and self-location are not a concern for the gameplay, teleport is to be used. It is a concept that has been experimented in commercial games, is relatively easy to implement, and many studies have shown it helps mitigate other factors of Simulator Sickness. This sort of locomotion is ideal for Detective games and other kinds of Puzzles. However, performance is hindered in FPS games, and traditional Race games are incompatible with teleport. Thus, for games in which self-location is necessary, we advise the use of continuous movement. When using continuous movement, it might be necessary to use other forms of Simulator Sickness mitigation techniques like a 2D view, which do not hide the borders and can help with dizziness as well. For Social games where self-location is not extremely needed, but the perception that other players have of you is important, we recommend the use of techniques based on Out-of-Locomotion [38] or a Semiautomatic Navigation Interface [39].

Section 9.2.1.3 Game Control

Game control, as has been presented in this thesis, is extremely influential to performance and sickness level, and can influence immersion and enjoyment. Thus, choosing an adequate control is essential for the quality of the game experience. We suggest that in competitive First-Person Shooter games, a keyboard and mouse be used, which can increase performance. On the other hand, for casual games, the trackable controller is more appropriate, given that it has been found to generate greater enjoyment. The controller should be preferred in other games that require minimal movement, given that other studies have found them more enjoyable.

Section 9.2.1.4 View

Choosing between a 2D and a 3D view in VR can be affected by the game style and by the controller that will be used. For competitive FPS games, we recommend the use of 2D views hidden as a helmet or other similar device associated with the use of the keyboard. For racing games, the same can be applied as the image in front of the glass. For social games and detective games in which fast movement is not a necessity, the use of the 3D view is recommended because the player can use both the keyboard and mouse or a controller without hindering any immersive perceptions of gamers.

Section 9.2.1.5 Perspective

We suggest the addition of perspective toggle buttons for all the games that require movement and that the perspective is not an essential part of the story and mechanics. The trade-off between immersion and Simulator Sickness in these cases should be determined by the player. However, we suggest the default for racing games and social games be the third person perspective.

Section 9.2.1.6 Virtual Environment Presentation

We recommend that in VR whenever possible creators should avoid maps and designs that offer sharps changes in direction. For instance, offering ramps instead of stairs will help avoid the up and down movement and reduce sickness. In racing games where turns are often unavoidable, it may be useful to try to opt for turns that do not change the direction close to 90 degrees or above, or reserve those for more difficult levels, after the user has gotten used to the VR device. In First-Person Shooter games, using a crosshair at all moments can be a natural way to offer a gazing point, and can improve performance results and reduce SS simultaneously. For horror games, the use of a reduced field of view is advisable since it can be used to create a mysterious atmosphere while preventing SS. For strategy games and platformers, the use of a 2D view during movement is likely not detrimental to the experience, especially associated with the traditional third-person perspective. In social and simulation games, occlusion circumvention techniques can be used.

Section 9.2.1.7 Measurements

Currently, we do not recommend extended use of EEG devices to measure game experience in VR. The positioning of the device is currently limited, and morphological skull differences make inferences exceedingly difficult. The positioning in one user's head might lead to different results in another user's head, leading to measurement errors. The format of certain commercial EEG devices does not match the format of the HMDs themselves. For example, the attachment point of the EEG is located where the HMD straps are located and this increases the complexity of the problem. Currently, most commercial devices and HMD add-ons collect data from the frontal part of the brain, which is more adequate to detect concentration, but not so adequate to detect more complex measurements involving immersion and enjoyment across participants. Thus, we currently recommend the use of EEG to measure focus in games and for level adjustments that are ideally associated with redundant measurements such as HR extracted from the HMDs movement. For Simulator Sickness measurement, we recommend the use of eye-tracking and in-game behavioural changes. All these techniques are non-invasive and require little additional cost.

Table 9-1 Summary of recommendations based on game types and their components.

	Race	**FPS - Competitive**	**Sandbox**	**Social**	**Casino / Puzzle**
Virtual Agent	Indifferent	Emotional	Emotional	Emotional	Indifferent
Locomotion	Continuous	Continuous	Teleport	Assisted Teleport	Teleport
Game Control	Controller	Keyboard and Mouse	Controller	Indifferent	Indifferent
View	2D View	2D View	3D View	3D View	2D View
Perspective	Toggle	1PP	Toggle	Toggle	Toggle

Section 9.3 Conclusion

In conclusion, this dissertation investigates the factors influencing enjoyment in VR games. It first presents a literature review outlining some of the current research on the topic and proposes several areas of investigation: (1) emotions in virtual characters; (2) aesthetical realism in virtual characters; (3) viewing perspective in games; (4) measuring subjective feelings in VR games; (5) preventing Simulator Sickness without disrupting enjoyment. Then, we have presented studies identifying recommendations in these areas of investigation and proposed one specific solution to preventing Simulator Sickness without disrupting enjoyment. At last, we conclude this dissertation with a set of guidelines for VR game development.

All in all, this thesis can act as a starting point for designers who are interested in creating games for VR HMDs. With the rapid advancements of the technology and the increase in interest from the general public, it is important to have guidelines to develop the most enjoyable experiences.

Section 9.4 Future Work

Several features could further affect enjoyment, immersion, and Simulator Sickness in VR games, but that could not be implemented and explored in this thesis due to time constraints.

In this work, we focused our investigation on visual stimuli. That is, in all our chapters, the topic involved how changes in appearance and presentation affect enjoyment, immersion, and Simulator Sickness. Even though this is fitting but given that the current VR HMDs focus on this kind of stimulus, future studies could evaluate how other senses affect these factors. Works on haptics, acoustics and olfactory feedback have great potential and are some of the subsequent barriers to a completely immersive system.

The conflict between senses seems to be one of the leading causes of Simulator Sickness. Thus, the development of prototypes that stimulate other senses can create more realistic and immersive experiences in VR and mitigate Simulator Sickness.

These prototypes can be done through wearables, other haptic technologies, or completely new devices to be investigated. And it could benefit teaching and industry as training tools given that academia is now grounds for research.

Furthermore, measuring user experience is challenging. Measuring devices can be expensive or incompatible among themselves. In addition, it is often unclear what they are measuring, requiring specialized interpretation. With the knowledge acquired from this research, we intend to research ways to measure user experience based on the data already collected and can be collected by existing devices, be it a Head-Mounted Display or the controllers. These studies will help develop feedback adaptable systems that are cost-effective and, in the process, incentivize an even wider adoption of the technologies.

Our study focuses on the short-term effects of these technologies. Longitudinal studies might bring new insights into the use of these techniques in the long term. Long-term adaptation might affect how a single modality of emotional expression is perceived. Moreover,because additional modalities of emotion expression might suffer from diminishing returns, their impact on a player’s affect on the long-term should be further evaluated. The long-term effects of VR sickness and their mitigation techniques are also an open field, such as how the sopite syndrome affects long-term VR users.

Finally, our investigation mainly focuses on young, healthy adults. Future works could focus on investigating the motion-based interaction with different population groups (e.g., elderly adults and users with physical disabilities). In addition, it will be useful to evaluate other playing positions in VR, like playing while lying on a bed or couch, which might be more convenient for the elderly or people who cannot leave the bed for other reasons. Findings from these studies will allow designing VR games and technologies that can cater to a broader range of users.

Bibliography

[1] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch, "The effect of avatar realism in immersive social virtual realities," *Proc. ACM Symp. Virtual Real. Softw. Technol. VRST*, vol. Part F1319, no. 2, pp. 185–190, 2017, doi: 10.1145/3139131.3139156.

[2] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik, "The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response," *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 4, pp. 1643–1652, Apr. 2018, doi: 10.1109/TVCG.2018.2794629.

[3] P. H. Huang and S. K. Wong, "Emotional virtual crowd on task completion in virtual markets," *Comput. Animat. Virtual Worlds*, vol. 29, no. 3–4, pp. 1–10, 2018, doi: 10.1002/cav.1818.

[4] J. M. Allbeck and N. I. Badler, "Toward Representing Agent Behaviors Modified by Personality and Emotion," *Work. Represent. Annot. Eval. Non-Verbal Verbal Commun. Acts to Achieve Context. Embodied Agents Auton. Agents*, vol. 2, pp. 15–19, 2002, doi: 10.1.1.19.2054.

[5] S. Kshirsagar and N. Magnenat-thalmann, "A multilayer personality model," *Proc. 2nd Int. Symp. Smart Graph. - SMARTGRAPH '02*, no. January 2002, pp. 107–115, 2002, doi: 10.1145/569005.569021.

[6] A. Egges, S. Kshirsagar, and N. Magnenat-thalmann, "A Model for Personality and Emotion Simulation," 2003.

[7] E. André, M. Klesen, P. Gebhard, S. Allen, and T. Rist, "Integrating Models of Personality and Emotions into Lifelike Characters," *Affect. Interact.*, vol. 1814, pp. 150–165, 2000, doi: 10.1007/10720296_11.

[8] J. Lasseter and R. San, "Principles of Traditional Animation Applied to 3D Computer Animation," 1987.

[9] E. Schneider, Y. Wang, and S. Yang, "Exploring the Uncanny Valley with Japanese video game characters," *3rd Digit. Games Res. Assoc. Int. Conf. "Situated Play. DiGRA 2007*, pp. 546–549, 2007.

[10] N. Afonso and R. Prada, "Agents that relate: Improving the social believability of Non-Player Characters in Role-Playing Games," *Lect. Notes Comput. Sci.*

(including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5309 LNCS, pp. 34–45, 2008, doi: 10.1007/978-3-540-89222-9-5.

[11] K. Kallinen, M. Salminen, N. Ravaja, R. Kedzior, and M. Sääksjärvi, "Presence and emotion in computer game players during 1st person vs. 3rd person playing view: Evidence from self-report, eye-tracking, and facial muscle activity data," *Proc. Presence*, pp. 187–190, 2007, [Online]. Available: http://www.temple.edu/ispr/prev_conferences/proceedings/2007/Kallinen, et al.pdf.

[12] J. J. Cummings and J. N. Bailenson, "How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence," *Media Psychol.*, vol. 19, no. 2, pp. 272–309, 2016, doi: 10.1080/15213269.2015.1015740.

[13] S. Rogers, *Level Up! The Guide to Great Video Game Design*, no. 2. 2014.

[14] E. Medina, R. Fruland, and S. Weghorst, "VIRTUSPHERE: walking in a human size VR 'hamster ball,'" *Proc. Hum. Factors Ergon. Soc.*, vol. 3, no. 27, pp. 2102–2106, 2008, doi: 10.1177/154193120805202704.

[15] A. Denisova and P. Cairns, "First-person vs. Third-person perspective in digital games: Do player preferences affect immersion?," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2015-April, pp. 145–148, 2015, doi: 10.1145/2702123.2702256.

[16] W. Xu, D. Monteiro, H. N. Liang, K. Hasan, Y. Yu, and C. Fleming, "Assessing the Effects of a Full-body Motion-based Exergame in Virtual Reality," in *ACM International Conference Proceeding Series*, 2019, pp. 1–6, doi: 10.1145/3332169.3333574.

[17] Y.-H. Peng *et al.*, "WalkingVibe: Reducing Virtual Reality Sickness and Improving Realism while Walking in VR using Unobtrusive Head-mounted Vibrotactile Feedback," pp. 1–12, 2020, doi: 10.1145/3313831.3376847.

[18] M. Gallagher, R. Dowsett, and E. R. Ferrè, "Vection in virtual reality modulates vestibular-evoked myogenic potentials," *Eur. J. Neurosci.*, vol. 50, no. 10, pp. 3557–3565, 2019, doi: 10.1111/ejn.14499.

[19] R. W. Wood, "The 'Haunted Swing' illusion," *Psychol. Rev.*, vol. 2, no. 3, pp. 277–278, 1895, doi: 10.1037/h0073333.

[20] M. Gallagher and E. R. Ferrè, "Cybersickness: A Multisensory Integration

Perspective,” *Multisens. Res.*, vol. 31, no. 7, pp. 645–674, 2018, doi: 10.1163/22134808-20181293.

[21] I. B. Adhanom, N. Navarro Griffin, P. Macneilage, and E. Folmer, “The Effect of a Foveated Field-of-view Restrictor on VR Sickness,” *Proc. - 2020 IEEE Conf. Virtual Real. 3D User Interfaces, VR 2020*, no. February, pp. 645–652, 2020, doi: 10.1109/VR46266.2020.1581314696458.

[22] J. P. Jorge and D. G. Fumanal, “Reducing Discomfort in Eye-Tracked HMDs using Defocus Blur Reducing Discomfort in Eye-Tracked HMDs using Defocus Blur,” no. September, pp. 2–3, 2017.

[23] P. Budhiraja, M. R. Miller, A. K. Modi, and D. Forsyth, “Rotation Blurring: Use of Artificial Blurring to Reduce Cybersickness in Virtual Reality First-person Shooters,” Oct. 2017, [Online]. Available: http://arxiv.org/abs/1710.02599.

[24] S. Bullard Rhodes College, N. Coomer Rhodes College Sadler Bullard Rhodes College William Clinton Rhodes College Betsy Williams-Sanders, N. Coomer, S. Bullard, W. Clinton, and B. Williams-Sanders, “Evaluating the Effects of Four VR Locomotion Methods: Joystick, Arm-Cycling, Point-Tugging, and Teleporting,” *SAP*, vol. 18, doi: 10.1145/3225153.3225175.

[25] “History Of Virtual Reality - Virtual Reality Society.” https://www.vrs.org.uk/virtual-reality/history.html (accessed Dec. 05, 2020).

[26] J.-H. K. J.-H. Kim *et al.*, “Virtual Reality History, Applications, Technology and Future,” *Digit. Outcasts*, vol. 63, no. ISlE, pp. 92–98, 2013, doi: 10.1.1.42.7849.

[27] A. Hutchison, *Back to the Holodeck-New Life for Virtual Reality?* 2007.

[28] A. Popescu, “Palmer Luckey | MIT Technology Review.” https://www.technologyreview.com/innovator/palmer-luckey/ (accessed Dec. 05, 2020).

[29] P. Cipresso, I. A. C. Giglioli, M. A. Raya, and G. Riva, “The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature,” *Front. Psychol.*, vol. 9, Nov. 2018, doi: 10.3389/fpsyg.2018.02086.

[30] “Profiles in Innovation Virtual & Augmented Reality Understanding the Race for the next computing platform,” Jan. 2016.

[31] “Virtual reality has been boosted by coronavirus – here’s how to avoid it leading us to dystopia.” https://theconversation.com/virtual-reality-has-been-boosted-by-coronavirus-heres-how-to-avoid-it-leading-us-to-dystopia-141073 (accessed Dec. 03, 2020).

[32] C. G. Christou and P. Aristidou, “Steering versus teleport locomotion for head mounted displays,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10325 LNCS, no. June 2017, pp. 431–446, 2017, doi: 10.1007/978-3-319-60928-7_37.

[33] W. Xu, H. N. Liang, Y. Zhao, D. Yu, and D. Monteiro, “DMove: Directional Motion-based Interaction for Augmented Reality Head-Mounted Displays,” in *Conference on Human Factors in Computing Systems - Proceedings*, 2019, pp. 1–14, doi: 10.1145/3290605.3300674.

[34] C. Lai and R. P. Mcmahan, “The Cognitive Load and Usability of Three Walking Metaphors for Consumer Virtual Reality,” 2020, doi: 10.1109/ISMAR50242.2020.00091.

[35] T. Weibker, A. Kunert, B. Frohlich, and A. Kulik, “Spatial Updating and Simulator Sickness during Steering and Jumping in Immersive Virtual Environments,” *25th IEEE Conf. Virtual Real. 3D User Interfaces, VR 2018 - Proc.*, pp. 97–104, 2018, doi: 10.1109/VR.2018.8446620.

[36] J. Clifton and S. Palmisano, “Effects of steering locomotion and teleporting on cybersickness and presence in HMD-based virtual reality,” *Virtual Real.*, no. 0123456789, 2019, doi: 10.1007/s10055-019-00407-8.

[37] J. Clifton and S. Palmisano, “Comfortable Locomotion in VR: Teleportation is Not a Complete Solution,” 2019, doi: 10.1145/3359996.3364722.

[38] N. N. Griffin and E. Folmer, “Out-of-body locomotion: Vectionless navigation with a continuous avatar representation,” *Proc. ACM Symp. Virtual Real. Softw. Technol. VRST*, 2019, doi: 10.1145/3359996.3364243.

[39] Y. Wang, J. R. Chardonnet, and F. Merienne, “A Semiautomatic Navigation Interface to Reduce Visually Induced Motion Sickness in Virtual Reality,” *13e Journées la realité virtuelle*, 2018.

[40] Z. Wang, H. Wei, K. Zhang, and L. Xie, “Real Walking in Place: HEX-CORE-PROTOTYPE Omnidirectional Treadmill,” pp. 382–387, 2020, doi: 10.1109/vr46266.2020.00058.

[41] S. H. Pyo, H. S. Lee, B. M. Phu, S. J. Park, and J. W. Yoon, "Development of an Fast-Omnidirectional Treadmill (F-ODT) for Immersive Locomotion Interface," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 760–766, 2018, doi: 10.1109/ICRA.2018.8460669.

[42] A. Barberis, T. Bennet, and M. Minear, "'Ready player one': Enhancing omnidirectional treadmills for use in virtual environments," *26th IEEE Conf. Virtual Real. 3D User Interfaces, VR 2019 - Proc.*, pp. 848–849, 2019, doi: 10.1109/VR.2019.8798323.

[43] D. Englmeier, F. Fan, and A. Butz, "Rock or Roll-Locomotion Techniques with a Handheld Spherical Device in Virtual Reality," 2020, doi: 10.1109/ISMAR50242.2020.00089.

[44] D. Hufnal, E. Osborne, T. Johnson, and C. Yildirim, "The impact of controller type on video game user experience in virtual reality," *2019 IEEE Games, Entertain. Media Conf. GEM 2019*, 2019, doi: 10.1109/GEM.2019.8811543.

[45] L. E. Nacke, "Wiimote vs. Controller : Electroencephalographic Measurement of Affective Gameplay Interaction," *Proc. Int. Acad. Conf. Futur. Game Des. Technol.*, pp. 159–166, 2010, doi: 10.1145/1920778.1920801.

[46] B. Sarupuriy, S. Hoermann, M. C. Whitton, and R. W. Lindeman, "Evaluating and Comparing Game-controller based Virtual Locomotion Techniques," *Int. Conf. Artif. Real. Telexistence Eurographics Symp. Virtual Environ. ICAT-EGVE 2017*, pp. 133–139, 2017, doi: 10.2312/egve.20171350.

[47] S. E. Kober, J. Kurzmann, and C. Neuper, "Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: An EEG study," *Int. J. Psychophysiol.*, vol. 83, no. 3, pp. 365–374, 2012, doi: 10.1016/j.ijpsycho.2011.12.003.

[48] A. N. Iwaniuk, J. E. Nelson, and S. M. Pellis, "Do big-brained animals play more? Comparative analyses of play and relative brain size in mammals.," *J. Comp. Psychol.*, vol. 115, no. 1, pp. 29–41, 2001, doi: 10.1037/0735-7036.115.1.29.

[49] C. Crawford, *The Art of Computer Game Design*. 1997.

[50] R. Koster, *A Theory of Fun for Game Design*, no. 2. 2013.

[51] "Virtual reality in gaming — Jasoren." https://jasoren.com/vr-in-gaming/ (accessed Dec. 03, 2020).

[52] L. Liu, R. K. F. Ip, C. Wagner, and A. Shum, "Learning effects of virtual game

worlds: An empirical investigation of immersion, enjoyment and performance," *20th Am. Conf. Inf. Syst. AMCIS 2014*, no. Kiili 2005, pp. 1–11, 2014.

[53] L. Nacke and C. A. Lindley, "Flow and immersion in first-person shooters: Measuring the player's gameplay experience," *ACM Futur. Play 2008 Int. Acad. Conf. Futur. Game Des. Technol. Futur. Play Res. Play. Share*, pp. 81–88, 2008, doi: 10.1145/1496984.1496998.

[54] B. De Schutter and V. Vanden Abeele, "Designing meaningful play within the psycho-social context of older adults," *ACM Int. Conf. Proceeding Ser.*, no. January 2010, pp. 84–93, 2010, doi: 10.1145/1823818.1823827.

[55] J. Hamari and J. Koivisto, "Measuring flow in gamification: Dispositional Flow Scale-2," *Comput. Human Behav.*, vol. 40, pp. 133–143, 2014, doi: 10.1016/j.chb.2014.07.048.

[56] R. R. Wehbe, E. D. Mekler, M. Schaekermann, E. Lank, and L. E. Nacke, "Testing incremental difficulty design in platformer games," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2017-Janua, pp. 5109–5113, 2017, doi: 10.1145/3025453.3025697.

[57] B. E. R. Garcia, M. K. Crocomo, and K. O. Andrade, "Dynamic Difficulty Adjustment in a Whac-a-Mole like Game," *Brazilian Symp. Games Digit. Entertain. SBGAMES*, vol. 2018-Novem, pp. 88–96, 2019, doi: 10.1109/SBGAMES.2018.00020.

[58] S. Wu and T. Lin, "Exploring the use of physiology in adaptive game design," *2011 Int. Conf. Consum. Electron. Commun. Networks, CECNet 2011 - Proc.*, pp. 1280–1283, 2011, doi: 10.1109/CECNET.2011.5768186.

[59] A. S. Fernandes and S. K. Feiner, "Combating VR sickness through subtle dynamic field-of-view modification," in *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, Mar. 2016, pp. 201–210, doi: 10.1109/3DUI.2016.7460053.

[60] X. Yu, D. Weng, and L. Cai, "Reduce simulator sickness by overwritten symbol in smartphone-based VR System," *Proc. - 2016 Int. Conf. Virtual Real. Vis. ICVRV 2016*, pp. 426–429, 2017, doi: 10.1109/ICVRV.2016.78.

[61] J. L. Dorado and P. A. Figueroa, "Ramps are better than stairs to reduce cybersickness in applications based on a HMD and a Gamepad," *IEEE Symp. 3D User Interfaces 2014, 3DUI 2014 - Proc.*, pp. 47–50, 2014, doi: 10.1109/3DUI.2014.6798841.

[62] J. Mütterlein and T. Hess, “Immersion, presence, interactivity: Towards a joint understanding of factors influencing virtual reality acceptance and use,” *AMCIS 2017 - Am. Conf. Inf. Syst. A Tradit. Innov.*, vol. 2017-Augus, pp. 1–10, 2017.

[63] S. Weech, S. Kenny, and M. Barnett-Cowan, “Presence and cybersickness in virtual reality are negatively related: A review,” *Front. Psychol.*, vol. 10, no. FEB, pp. 1–19, 2019, doi: 10.3389/fpsyg.2019.00158.

[64] K. Poels, Y. de Kort, and W. A. IJsselsteijn, “Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games,” pp. 1–9, 2015, [Online]. Available: https://pure.tue.nl/ws/files/21666907/Game_Experience_Questionnaire_English.pdf.

[65] J. M. Kivikangas *et al.*, “Review on psychophysiological methods in game research,” *Proc. 1st Nord. DiGRA*, no. November, 2010.

[66] M. Abujelala, A. Sharma, C. Abellanoza, and F. Makedon, “Brain-EE: Brain enjoyment evaluation using commercial EEG headband,” *ACM Int. Conf. Proceeding Ser.*, vol. 29-June-20, no. June, pp. 1–5, 2016, doi: 10.1145/2910674.2910691.

[67] J. P. Tauscher *et al.*, “Exploring neural and peripheral physiological correlates of simulator sickness,” *Comput. Animat. Virtual Worlds*, vol. 31, no. 4–5, pp. 1–11, 2020, doi: 10.1002/cav.1953.

[68] D. Shin, G. Lee, D. Shin, and D. Shin, “Mental state measurement system using EEG analysis,” *Lect. Notes Electr. Eng.*, vol. 309 LNEE, no. Future Information Technology, pp. 451–456, 2014, doi: 10.1007/978-3-642-55038-6_70.

[69] D. Monteiro, H. Liang, W. Xu, M. Brucker, V. Nanjappan, and Y. Yue, “Evaluating enjoyment, presence, and emulator sickness in VR games based on first- and third- person viewing perspectives,” *Comput. Animat. Virtual Worlds*, vol. 29, no. 3–4, p. e1830, May 2018, doi: 10.1002/cav.1830.

[70] H. G. Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulic, “Characterizing first and third-person viewpoints and their alternation for embodied interaction in virtual reality,” *PLoS One*, vol. 12, no. 12, pp. 1–19, 2017, doi: 10.1371/journal.pone.0190109.

[71] D. Monteiro, H.-N. Liang, Y. Zhao, and A. Abel, “Comparing Event Related Arousal-Valence and Focus among Different Viewing Perspectives in VR

gaming," in *Advances in Brain Inspired Cognitive Systems 9th International Conference, BICS 2018, Proceedings*, 2018, pp. 1–10, doi: 10.1007/978-3-030-00563-4_75.

[72] D. Monteiro, H.-N. Liang, and N. Baghaei, "Correlating gamers' brainwaves to their subjective feelings in virtual reality games under different viewing perspectives," in *Proceedings of 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI'18)*, 2018, pp. 1–4, doi: 10.1145/3284398.3284430.

[73] D. Monteiro, H.-N. Liang, A. Abel, N. Baghaei, and R. de C. R. V. Monteiro, "Evaluating Engagement of Virtual Reality Games based on First- and Third-Person Viewing Perspectives using EEG and Subjective Metrics," 2018.

[74] L. A. Wagner, "When Your Smartphone Is Too Smart for Your Own Good: How Social Media Alters Human Relationships," *J. Individ. Psychol.*, vol. 71, no. 2, pp. 114–121, 2015, doi: 10.1353/jip.2015.0009.

[75] C. V. Bellieni, "Meaning and importance of weeping," *New Ideas Psychol.*, vol. 47, no. July, pp. 72–76, 2017, doi: 10.1016/j.newideapsych.2017.06.003.

[76] K. P. Kuypers, "Context-dependent emotional empathy in virtual reality," *Adv. Soc. Sci. Res. J.*, vol. 5, no. 4, pp. 379–392, 2018, doi: 10.14738/assrj.54.4514.

[77] M. Scheutz, "The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots," *Robot Ethics Ethical Soc. Implic. Robot.*, no. November, pp. 205–221, 2011, [Online]. Available: http://books.google.com/books?hl=en&lr=&id=oBb-lt3l4oYC&oi=fnd&pg=PA205&dq=The+Inherent+Dangers+of+Unidirectional+Emotional+Bonds+between+Humans+and+Social+Robots&ots=yuaSs8y33n&sig=h5KrpJxIcyk5RT2vih0YpgSHqNY.

[78] M. El Jed, N. Pallamin, J. Dugdale, and B. Pavard, "Modelling character emotion in an interactive virtual environment," 2004.

[79] A. S. Horta, L. F. De Almeida, R. D. C. R. V. Monteiro, and D. V. Monteiro, "A proposal of a game for education and environmental consciousness," in *IAMOT 2015 - 24th International Association for Management of Technology Conference: Technology, Innovation and Management for Sustainable Growth, Proceedings*, 2015, pp. 1664–1676.

[80] M. Martončik, "E-Sports: Playing just for fun or playing to satisfy life goals?,"

Comput. Human Behav., vol. 48, pp. 208–211, Jul. 2015, doi: 10.1016/j.chb.2015.01.056.

[81] R. Zheng, H.-N. Liang, R. Xie, Y. Shi, F. Lu, and K. Papangelis, “BlockTower: A Multi-player Cross-platform Competitive Social Game,” *Yiyu Cai, Wouter van Joolingen Zachary Walk. (Eds.), VR, Simul. Serious Gaming Educ.*, 2018.

[82] H. N. Liang, F. Lu, Y. Shi, V. Nanjappan, and K. Papangelis, “Evaluating the effects of collaboration and competition in navigation tasks and spatial knowledge acquisition within virtual reality environments,” *Futur. Gener. Comput. Syst.*, no. March, 2018, doi: 10.1016/j.future.2018.02.029.

[83] D. Yu, K. Fan, H. Zhang, D. Monteiro, W. Xu, and H.-N. Liang, “PizzaText: Text Entry for Virtual Reality Systems Using Dual Thumbsticks,” *IEEE Trans. Vis. Comput. Graph.*, vol. PP, no. c, pp. 1–1, 2018, doi: 10.1109/TVCG.2018.2868581.

[84] M. J. P. Wolf, “Abstraction in the video game,” *Video Game Theory Read.*, pp. 47–65, 2013, doi: 10.4324/9780203700457-9.

[85] W. Jin, D. Gromala, and X. Tong, “Serious game for serious disease: Diminishing stigma of depression via game experience,” *2015 IEEE Games Entertain. Media Conf. GEM 2015*, no. January 2016, 2016, doi: 10.1109/GEM.2015.7377256.

[86] P. Gebhard and K. H. Kipp, “Are computer-generated emotions and moods plausible to humans?,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4133 LNAI, pp. 343–356, 2006, doi: 10.1007/11821830_28.

[87] R. R. McCrae and O. P. John, “An Introduction to the Five-Factor Model and Its Applications,” *J. Pers.*, vol. 60, no. 2, pp. 175–215, 1992, doi: 10.1111/j.1467-6494.1992.tb00970.x.

[88] J. A. Russell and L. F. Barrett, “Russell-Barrett-1999,” *J. Pers. Soc. Psychol.*, vol. 76, no. 5, 1999.

[89] G. S. Thirunavukkarasu, H. Abdi, and N. Mohajer, “A smart HMI for driving safety using emotion prediction of EEG signals,” in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2016, vol. 1, pp. 004148–004153, doi: 10.1109/SMC.2016.7844882.

[90] A. Popescu, J. Broekens, and M. Van Someren, “GAMYGDALA: An emotion

engine for games," *IEEE Trans. Affect. Comput.*, vol. 5, no. 1, pp. 32–44, 2014, doi: 10.1109/T-AFFC.2013.24.

[91] A. Leonardo, A. Brisson, and A. Paiva, "Intelligent Virtual Agents," in *Intelligent Virtual Agents*, vol. 5773, no. May 2016, Z. Ruttkay, M. Kipp, A. Nijholt, and H. H. Vilhjálmsson, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 523–524.

[92] T. Verhagen, J. van Nes, F. Feldberg, and W. van Dolen, "Virtual customer service agents: Using social presence and personalization to shape online service encounters," *J. Comput. Commun.*, vol. 19, no. 3, pp. 529–545, 2014, doi: 10.1111/jcc4.12066.

[93] Z. Kasap, M. Ben Moussa, P. Chaudhuri, and N. Magnenat-Thalmann, "Making them remember - Emotional virtual characters with memory," *IEEE Comput. Graph. Appl.*, vol. 29, no. 2, pp. 20–29, 2009, doi: 10.1109/MCG.2009.26.

[94] P. Ruijten, "On the Perceived Human-likeness of Virtual Health Agents : Towards a Generalized Measurement of Anthropomorphism On the Perceived Human-likeness of Virtual Health Agents : Towards a Generalized Measurement of Anthropomorphism," no. AUGUST, 2015.

[95] S. R. Fussell, S. Kiesler, L. D. Setlock, and V. Yew, "How people anthropomorphize robots," *Proc. 3rd Int. Conf. Hum. Robot Interact. - HRI '08*, no. January 2008, p. 145, 2008, doi: 10.1145/1349822.1349842.

[96] L. Frädrich, F. Nunnari, M. Staudte, and A. Heloir, "(Simulated) listener gaze in real-time spoken interaction," *Comput. Animat. Virtual Worlds*, vol. 29, no. 3–4, pp. 1–11, 2018, doi: 10.1002/cav.1831.

[97] J. Funge, X. Tu, and D. Terzopoulos, "Cognitive Modeling: Knowledge, Reasoning and Planning for Intelligent Characters," *Proc. 26th Annu. Conf. Comput. Graph. Interact. Tech. (SIGGRAPH '99)*, pp. 29–38, 1999, doi: 10.1145/311535.311538.

[98] M. Mustafa, S. Guthe, J. P. Tauscher, M. Goesele, and M. Magnor, "How human am I? EEG-based evaluation of animated virtual characters," in *Conference on Human Factors in Computing Systems - Proceedings*, 2017, vol. 2017-May, pp. 5098–5108, doi: 10.1145/3025453.3026043.

[99] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley," *IEEE Robot. Autom. Mag.*, vol. 19, no. 2, pp. 98–100, 2012, doi:

10.1109/MRA.2012.2192811.

[100] J. Bates, A. Bryan Loyall, and W. Scott Reilly, “An architecture for action, emotion, and social behavior,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 830 LNAI, pp. 55–68, 1994, doi: 10.1007/3-540-58266-5_4.

[101] D. Rousseau, “Personality in Computer Characters,” *AAAI Work. Entertain. AI / A-Life*, pp. 38–43, 1996, [Online]. Available: http://www.aaai.org/Papers/Workshops/1996/WS-96-03/WS96-03-008.pdf.

[102] B. N. Colby, A. Ortony, G. L. Clore, and A. Collins, “The Cognitive Structure of Emotions.,” *Contemp. Sociol.*, 1989, doi: 10.2307/2074241.

[103] A. Egges, S. Kshirsagar, and N. Magnenat-Thalmann, “Generic personality and emotion simulation for conversational agents,” *Comput. Animat. Virtual Worlds*, vol. 15, no. 1, pp. 1–13, 2004, doi: 10.1002/cav.3.

[104] P. Gebhard, “ALMA - A layered model of affect,” in *Proceedings of the International Conference on Autonomous Agents*, 2005, no. December, pp. 177–184, [Online]. Available: http://dl.acm.org/citation.cfm?id=1082478&CFID=539211132&CFTOKEN=92307229.

[105] H. C. Van Vugt, J. F. Hoorn, E. A. Konijn, and A. de B. Dimitriadou, “Affective affordances Improving interface character engagement through interaction,” *Int. J. ofHuman-Computer Stud.*, vol. 64, no. 9, pp. 874–888, 2006.

[106] R. S. Kennedy, N. E. Lane, S. Kevin, and M. G. Lilienthal, “The International Journal of Aviation Psychology Simulator Sickness Questionnaire : An Enhanced Method for Quantifying Simulator Sickness,” *Int. J. Aviat. Psychol.*, vol. 3, no. 3, pp. 203–220, 1993, doi: 10.1207/s15327108ijap0303.

[107] C. B. Mills and A. M. Carwile, “The good, the bad, and the Borderline: Separating teasing from bullying,” *Commun. Educ.*, vol. 58, no. 2, pp. 276–301, 2009, doi: 10.1080/03634520902783666.

[108] A. Oleszkiewicz, T. Frackowiak, A. Sorokowska, and P. Sorokowski, “Children can accurately recognize facial emotions from emoticons,” *Comput. Human Behav.*, vol. 76, pp. 372–377, 2017, doi: 10.1016/j.chb.2017.07.040.

[109] K. K. Dwyer and M. M. Davidson, “Is Public Speaking Really More Feared Than Death?,” *Commun. Res. Reports*, vol. 29, no. 2, pp. 99–107, 2012, doi:

10.1080/08824096.2012.667772.

[110] E. Carl *et al.*, “Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials,” *J. Anxiety Disord.*, vol. 61, no. July 2018, pp. 27–36, 2019, doi: 10.1016/j.janxdis.2018.08.003.

[111] P. Anderson, B. O. Rothbaum, and L. F. Hodges, “Virtual Reality Exposure in the Treatment of Social Anxiety,” *Cogn. Behav. Pract.*, vol. 10, no. 3, pp. 240–247, 2003, doi: 10.1016/S1077-7229(03)80036-6.

[112] A. Cheng, L. Yang, and E. Andersen, “Teaching Language and Culture through a Virtual Reality Game,” *Chi*, pp. 541–549, 2017, doi: 10.1145/3025453.3025857.

[113] A. Pack, A. Barrett, H. N. Liang, and D. V. Monteiro, “University EAP students’ perceptions of using a prototype virtual reality learning environment to learn writing structure,” *Int. J. Comput. Lang. Learn. Teach.*, vol. 10, no. 1, pp. 27–46, 2020, doi: 10.4018/IJCALLT.2020010103.

[114] M. Slater, D.-P. Pertaub, C. Barker, and D. M. Clark, “An Experimental Study on Fear of Public Speaking Using a Virtual Environment,” *CyberPsychology Behav.*, vol. 9, no. 5, pp. 627–633, 2008, doi: 10.1089/cpb.2006.9.627.

[115] D. Monteiro, H.-N. N. Liang, J. Wang, L. Wang, X. Wang, and Y. Yue, “Evaluating the effects of a cartoon-like character with emotions on users’ behaviour within virtual reality environments,” in *Proceedings - 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality, AIVR 2018*, Dec. 2019, pp. 229–236, doi: 10.1109/AIVR.2018.00053.

[116] D. Monteiro, H. N. Liang, A. Abel, N. Bahaei, and R. De Cassia Monteiro, “Evaluating engagement of virtual reality games based on first and third-person perspective using EEG and subjective metrics,” in *Proceedings - 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality, AIVR 2018*, 2019, pp. 53–60, doi: 10.1109/AIVR.2018.00015.

[117] D. Monteiro, H. N. Liang, W. Xu, M. Brucker, V. Nanjappan, and Y. Yue, “Evaluating enjoyment, presence, and emulator sickness in VR games based on first- and third- person viewing perspectives,” *Comput. Animat. Virtual Worlds*, vol. 29, no. 3–4, p. e1830, 2018, doi: 10.1002/cav.1830.

[118] M. Slater, D. P. Pertaub, and A. Steed, “Public Speaking in Virtual Reality: Facing an Audience of Avatars,” *IEEE Comput. Graph. Appl.*, vol. 19, no. 2,

pp. 6–9, 1999, doi: 10.1109/38.749116.

[119] H. S. Wallach, M. P. Safir, and M. Bar-Zvi, "Virtual reality cognitive behavior therapy for public speaking anxiety: A randomized clinical trial," *Behav. Modif.*, vol. 33, no. 3, pp. 314–338, 2009, doi: 10.1177/0145445509331926.

[120] S. Poeschl and N. Doering, "Virtual training for fear of public speaking - Design of an audience for immersive virtual environments," *Proc. - IEEE Virtual Real.*, no. October 2014, pp. 101–102, 2012, doi: 10.1109/VR.2012.6180902.

[121] S. Poeschl, "Virtual reality training for public speaking-A QUEST-VR framework validation," *Front. ICT*, vol. 4, no. JUN, 2017, doi: 10.3389/fict.2017.00013.

[122] S. Côté and S. Bouchard, "Virtual reality exposure for phobias: A critical review," *J. Cyber Ther. Rehabil.*, vol. 1, no. 1, pp. 75–92, 2008.

[123] S. R. Harris, R. L. Kemmerling, and M. M. North, "Brief virtual reality therapy for public speaking anxiety," *Cyberpsychology Behav.*, vol. 5, no. 6, pp. 543–550, 2002, doi: 10.1089/109493102321018187.

[124] R. A. I. Cardoş, O. A. David, and D. O. David, "Virtual reality exposure therapy in flight anxiety: A quantitative meta-analysis," *Comput. Human Behav.*, vol. 72, no. 37, pp. 371–380, 2017, doi: 10.1016/j.chb.2017.03.007.

[125] H. M. Zinzow *et al.*, "Virtual Reality and Cognitive-Behavioral Therapy for Driving Anxiety and Aggression in Veterans: A Pilot Study," *Cogn. Behav. Pract.*, vol. 25, no. 2, pp. 296–309, 2018, doi: 10.1016/j.cbpra.2017.09.002.

[126] H. Lee *et al.*, "Annotation vs. Virtual tutor: Comparative analysis on the effectiveness of visual instructions in immersive virtual reality," in *Proceedings - 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2019*, 2019, pp. 318–327, doi: 10.1109/ISMAR.2019.00030.

[127] J. Psotka, "Educational Games and Virtual Reality as Disruptive Technologies," *Educ. Technol. Soc.*, vol. 16, no. 2, pp. 69–80, 2013.

[128] Z. Merchant, E. T. Goetz, L. Cifuentes, W. Keeney-Kennicutt, and T. J. Davis, "Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis," *Comput. Educ.*, vol. 70, no. January 2014, pp. 29–40, 2014, doi: 10.1016/j.compedu.2013.07.033.

[129] W. Xu, H.-N. Liang, Z. Zhang, and N. Baghaei, "Studying the Effect of Display Type and Viewing Perspective on User Experience in Virtual Reality

Exergames," *Games Health J.*, vol. 9, no. 4, pp. 1–10, 2020, doi: 10.1089/g4h.2019.0102.

[130] L. Chittaro and F. Buttussi, "Assessing knowledge retention of an immersive serious game vs. A traditional education method in aviation safety," *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 4, pp. 529–538, 2015, doi: 10.1109/TVCG.2015.2391853.

[131] G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding immersive virtual reality to a science lab simulation causes more presence but less learning," *Learn. Instr.*, vol. 60, no. May, pp. 225–236, Dec. 2019, doi: 10.1016/j.learninstruc.2017.12.007.

[132] M. M. North, J. Hill, A. S. Aikhuele, and S. M. North, "Virtual reality training in aid of communication apprehension in classroom environments," *Int. J. Emerg. Technol. Learn.*, vol. 3, no. 2, pp. 34–37, 2008.

[133] C. O'Donnell, "Game dev tycoon," in *How to Play Video Games*, M. T. Payne and N. B. Huntemann, Eds. New York: New York University Press, 2019.

[134] M. Sicart, "FAMILY VALUES : IDEOLOGY , COMPUTER GAMES & THE Sims," *DiGRA '03 - Proc. 2003 DiGRA Int. Conf. Lev. Up*, p. 12, 2004, [Online]. Available: http://www.digra.org/wp-content/uploads/digital-library/05150.09529.pdf.

[135] J. Ranalli, "Learning English with The Sims : exploiting authentic computer simulation games for L2 learning," *Comput. Assist. Lang. Learn.*, vol. 21, no. 5, pp. 441–455, Dec. 2008, doi: 10.1080/09588220802447859.

[136] M. del C. Miranda-Duro, P. Concheiro-Moscoso, J. L. Viqueira, L. Nieto-Riveiro, N. C. Domínguez, and T. P. García, "Virtual Reality Game Analysis for People with Functional Diversity: An Inclusive Perspective," *Proceedings*, vol. 54, no. 1, p. 20, 2020, doi: 10.3390/proceedings2020054020.

[137] J. K. Argasiński, P. Wegrzyn, and P. Strojny, "Affective VR serious game for firefighter training," *CEUR Workshop Proc.*, vol. 2166, 2018.

[138] R. Shewaga, A. Uribe-Quevedo, B. Kapralos, and F. Alam, "A Comparison of Seated and Room-Scale Virtual Reality in a Serious Game for Epidural Preparation," *IEEE Trans. Emerg. Top. Comput.*, vol. 8, no. 1, pp. 218–232, 2020, doi: 10.1109/TETC.2017.2746085.

[139] D. Villani, C. Repetto, P. Cipresso, and G. Riva, "May i experience more

presence in doing the same thing in virtual reality than in reality? An answer from a simulated job interview,” *Interact. Comput.*, vol. 24, no. 4, pp. 265–272, 2012, doi: 10.1016/j.intcom.2012.04.008.

[140] A. Gruber and R. Kaplan-Rakowski, “User Experience of Public Speaking Practice in Virtual Reality,” *Cogn. Affect. Perspect. immersive Technol.*, no. January, pp. 235–249, 2020, doi: 10.4018/978-1-7998-3250-8.ch012.

[141] J. N. Hook, C. A. Smith, and D. P. Valentiner, “A short-form of the Personal Report of Confidence as a Speaker,” *Pers. Individ. Dif.*, vol. 44, no. 6, pp. 1306–1313, 2008, doi: 10.1016/j.paid.2007.11.021.

[142] S. G. Hofmann and P. M. DiBartolo, “An Instrument to Assess Self-Statements During Public Speaking: Scale Development and Preliminary Psychometric Properties,” *Behav Ther*, vol. 135, no. 2, pp. 612–615, 2000.

[143] J. H. Kim and I. Choi, “Choosing the Level of Significance: A Decision-theoretic Approach,” *Abacus*, pp. 1–45, 2019, doi: 10.1111/abac.12172.

[144] A. C. F. Marinho, A. M. de Medeiros, E. de P. Lima, J. J. Pantuza, and L. C. Teixeira, “Prevalence and factors associated with fear of public speaking,” *Codas*, vol. 31, no. 6, pp. 1–5, 2019, doi: 10.1590/2317-1782/20192018266.

[145] H. A. Lister, C. D. Piercey, and C. Joordens, “The effectiveness of 3-D video virtual reality for the treatment of fear of public speaking,” *J. Cyber Ther. Rehabil.*, vol. 3, no. 4, pp. 375–381, 2010.

[146] T. Schlösser, D. Dunning, K. L. Johnson, and J. Kruger, “How unaware are the unskilled? Empirical tests of the ‘signal extraction’ counterexplanation for the Dunning-Kruger effect in self-evaluation of performance,” *J. Econ. Psychol.*, vol. 39, pp. 85–100, 2013, doi: 10.1016/j.joep.2013.07.004.

[147] J. Kolligian and R. J. Sternberg, “Perceived Fraudulence in Young Adults: Is There an ‘Impostor Syndrome’?,” *J. Pers. Assess.*, vol. 56, no. 2, pp. 308–326, 1991, doi: 10.1207/s15327752jpa5602_10.

[148] M. Ibáñez Gallén, P. Quintana, S. Bouchard, C. Botella Arbona, and B. Serrano Zárate, “Validación de la interacción verbal con un avatar en un ambiente de realidad virtual para generar ansiedad social,” *Àgora Salut*, vol. IV, pp. 185–195, 2017, doi: 10.6035/agorasalut.2017.4.19.

[149] S. Stupar-Rutenfrans, L. E. H. Ketelaars, and M. S. Van Gisbergen, “Beat the Fear of Public Speaking: Mobile 360° Video Virtual Reality Exposure Training

in Home Environment Reduces Public Speaking Anxiety," *Cyberpsychology, Behav. Soc. Netw.*, vol. 20, no. 10, pp. 624–633, 2017, doi: 10.1089/cyber.2017.0174.

[150] A. Felnhofer, O. D. Kothgassner, T. Hetterle, L. Beutl, H. Hlavacs, and I. Kryspin-Exner, "Afraid to be there? Evaluating the relation between presence, self-reported anxiety, and heart rate in a virtual public speaking task," *Cyberpsychology, Behav. Soc. Netw.*, vol. 17, no. 5, pp. 310–316, 2014, doi: 10.1089/cyber.2013.0472.

[151] J. D. J. L. G. Ibánez and A. I. Wang, "Learning recycling from playing a kinect game," *Int. J. Game-Based Learn.*, vol. 5, no. 3, pp. 25–44, 2015, doi: 10.4018/IJGBL.2015070103.

[152] D. Sue, P. Ray, A. Talaei-Khoei, J. Jonnagaddala, and S. Vichitvanichphong, "Assessing video games to improve driving skills: A literature review and observational study," *J. Med. Internet Res.*, vol. 16, no. 8, p. e5, 2014, doi: 10.2196/games.3274.

[153] A. Hans and E. Hans, "Kinesics, Haptics and Proxemics: Aspects of Non - Verbal Communication," *IOSR J. Humanit. Soc. Sci. Ver. IV*, vol. 20, no. 2, pp. 47–52, 2015, doi: 10.9790/0837-20244752.

[154] R. Barmaki, "Improving social communication skills using kinesics feedback," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 07-12-May-, pp. 86–91, 2016, doi: 10.1145/2851581.2890378.

[155] D. Monteiro, J. Wang, H. N. Liang, N. Baghaei, and A. Abel, "SwarMotion: A 3D Point Cloud Video Recording Tool for Classification Purposes," in *Lecture Notes in Electrical Engineering*, vol. 590, Springer, Singapore, 2020, pp. 190–195.

[156] Z. Saenz-De-Urturi and B. Garcia-Zapirain Soto, "Kinect-based virtual game for the elderly that detects incorrect body postures in real time," *Sensors (Switzerland)*, vol. 16, no. 5, 2016, doi: 10.3390/s16050704.

[157] M. N. Selzer, N. F. Gazcon, and M. L. Larrea, "Effects of virtual presence and learning outcome using low-end virtual reality systems," *Displays*, vol. 59, no. May, pp. 9–15, 2019, doi: 10.1016/j.displa.2019.04.002.

[158] W. M. Hsieh *et al.*, "Virtual reality system based on Kinect for the elderly in fall prevention," *Technol. Heal. Care*, vol. 22, no. 1, pp. 27–36, 2014, doi:

10.3233/THC-130769.

[159] Y. Pisan, J. J. G. Marin, and K. F. K. Navarro, "Improving lives: using Microsoft Kinect to predict the loss of balance for elderly users under cognitive load," *Proc. 9th Australas. Conf. Interact. Entertain. Matters Life Death - IE '13*, pp. 1–4, 2013, doi: 10.1145/2513002.2513026.

[160] L. Chen, C. W. Leong, G. Feng, C. M. Lee, and S. Somasundaran, "Utilizing multimodal cues to automatically evaluate public speaking performance," *2015 Int. Conf. Affect. Comput. Intell. Interact. ACII 2015*, pp. 394–400, 2015, doi: 10.1109/ACII.2015.7344601.

[161] N. Linardopoulos, "Teaching and learning public speaking online," *MERLOT J. Online Learn. Teach.*, vol. 6, no. 1, pp. 198–209, 2010, [Online]. Available: http://jolt.merlot.org/vol6no1/linardopoulos_0310.pdf.

[162] R. C. Shih, "Blended learning using video-based blogs: Public speaking for English as a second language students," *Australas. J. Educ. Technol.*, vol. 26, no. 6, pp. 883–897, 2010.

[163] S. J. Lackey, K. A. Badillo-Urquiola, E. C. Ortiz, and I. L. Hudson, "A process for developing accurate kinesic cues in virtual environments," *24th Conf. Behav. Represent. Model. Simulation, BRiMS 2015, co-located with Int. Soc. Comput. Behav. Model. Predict. Conf. SBP 2015*, pp. 2–9, 2015.

[164] C. Groth, T. U. Braunschweig, S. Castillo, and M. Magnor, "Altering the Conveyed Facial Emotion Through Automatic Reenactment of Video Portraits," 2020.

[165] Ł. Dziuda, M. P. Biernacki, P. M. Baran, and O. E. Truszczyński, "The effects of simulated fog and motion on simulator sickness in a driving simulator and the duration of after-effects," *Appl. Ergon.*, vol. 45, no. 3, pp. 406–412, 2014, doi: 10.1016/j.apergo.2013.05.003.

[166] G. Gorisse, O. Christmann, E. A. Amato, and S. Richir, "First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users," *Front. Robot. AI*, vol. 4, no. July, pp. 1–12, 2017, doi: 10.3389/frobt.2017.00033.

[167] P. Salamin, T. Tadi, O. Blanke, F. Vexo, and D. Thalmann, "Quantifying effects of exposure to the third and first-person perspectives in virtual-reality-based training," *IEEE Trans. Learn. Technol.*, vol. 3, no. 3, pp. 272–276, 2010, doi:

10.1109/TLT.2010.13.

[168] D. Clarke, G. McGregor, B. Rubin, J. Stanford, and T. C. N. Graham, “Arcaid: Addressing situation awareness and simulator sickness in a virtual reality Pac-Man game,” *CHI Play 2016 - Proc. Annu. Symp. Comput. Interact. Play Companion*, pp. 39–45, 2016, doi: 10.1145/2968120.2968124.

[169] S. Lim and B. Reeves, “Being in the Game: Effects of Avatar Choice and Point of View on Psychophysiological Responses During Play,” *Media Psychol.*, vol. 12, no. 4, pp. 348–370, 2009, doi: 10.1080/15213260903287242.

[170] H. G. Debarba, E. Molla, B. Herbelin, and R. Boulic, “Characterizing embodied interaction in First and Third-person Perspective viewpoints,” *2015 IEEE Symp. 3D User Interfaces, 3DUI 2015 - Proc.*, pp. 67–72, 2015, doi: 10.1109/3DUI.2015.7131728.

[171] E. Kokkinara, K. Kilteni, K. J. Blom, and M. Slater, “First-person Perspective of Seated Participants over a Walking Virtual Body Leads to Illusory Agency over the Walking,” *Sci. Rep.*, vol. 6, no. July 2015, pp. 1–11, 2016, doi: 10.1038/srep28879.

[172] P. Salamin, D. Thalmann, and F. Vexo, “The benefits of third-person perspective in virtual and augmented reality,” *Proc. ACM Symp. Virtual Real. Softw. Technol. VRST*, pp. 27–30, 2006, doi: 10.1145/1180495.1180502.

[173] A. Covaci, A. H. Olivier, and F. Multon, “Visual Perspective and Feedback Guidance for VR Free-Throw Training,” *IEEE Comput. Graph. Appl.*, vol. 35, no. 5, pp. 55–65, Sep. 2015, doi: 10.1109/MCG.2015.95.

[174] A. Covaci, A. H. Olivier, and F. Multon, “Third-person view and guidance for more natural motor behaviour in immersive basketball playing,” *Proc. ACM Symp. Virtual Real. Softw. Technol. VRST*, no. December 2016, pp. 55–64, 2014, doi: 10.1145/2671015.2671023.

[175] S. Bouchard, G. Robillard, P. Renaud, and F. Bernier, “Exploring New Dimensions in the Assessment of Virtual Reality Induced Side Effects,” *J. Comput. Inf. Technol.*, vol. 1, no. 3, pp. 20–32, 2011.

[176] C. Jennett *et al.*, “Measuring and defining the experience of immersion in games,” *Int. J. Hum. Comput. Stud.*, vol. 66, no. 9, pp. 641–661, 2008, doi: 10.1016/j.ijhcs.2008.04.004.

[177] S. Iwata, “The 72nd Annual General Meeting of Shareholders Q&A,” *Nintendo*

The 72nd Annual General Meeting of Shareholders, 2012. https://www.nintendo.co.jp/ir/en/stock/meeting/120628qa/index.html (accessed Feb. 01, 2018).

[178] S. Hollensen, "The Blue Ocean that disappeared - the case of Nintendo Wii," *J. Bus. Strategy*, vol. 34, no. 5, pp. 25–35, 2013, doi: 10.1108/JBS-02-2013-0012.

[179] EEDAR, "2018 Q4 Round Up," 2019. https://www.eedar.com/post/2018-q4-round-up.

[180] Dolphin VR, "Dolphin Vr – a Gamecube and Wii Emulator With Vr Support," *Dolphin VR 5.0-250*, 2016. https://dolphinvr.wordpress.com/ (accessed Feb. 01, 2018).

[181] Nintendo of America, "Mario Kart Wii for Wii - Nintendo Game Details," *Mario Kart Wii*, 2008. https://www.nintendo.com/games/detail/O8zz_eN8oGRK9ApOgJ86zbE6zRv3pdB2 (accessed Feb. 01, 2018).

[182] IBM Analytics, "IBM SPSS Software," *Ibm*, 2016. http://www.ibm.com/analytics/us/en/technology/spss/ (accessed Feb. 01, 2018).

[183] N. Baghaei, D. Nandigam, J. Casey, A. Direito, and R. Maddison, "Diabetic Mario: Designing and Evaluating Mobile Games for Diabetes Education," *Games Health J.*, vol. 5, no. 4, pp. 270–278, 2016, doi: 10.1089/g4h.2015.0038.

[184] S. S. Ho, M. O. Lwin, J. R. H. Sng, and A. Z. H. Yee, "Escaping through exergames: Presence, enjoyment, and mood experience in predicting children's attitude toward exergames," *Comput. Human Behav.*, vol. 72, pp. 381–389, 2017, doi: 10.1016/j.chb.2017.03.001.

[185] D. Villani, C. Carissoli, S. Triberti, A. Marchetti, G. Gilli, and G. Riva, "Videogames for Emotion Regulation: A Systematic Review," *Games Health J.*, vol. 7, no. 2, p. g4h.2017.0108, 2018, doi: 10.1089/g4h.2017.0108.

[186] M. Hosťovecký and B. Babušiak, "Brain activity : beta wave analysis of 2D and 3D serious games using EEG," *JAMSI*, vol. 13, no. 2, pp. 39–53, 2017, doi: 10.1515/jamsi-2017-0008.

[187] F. Lu, D. Yu, H. N. Liang, W. Chen, K. Papangelis, and N. M. Ali, "Evaluating Engagement Level and Analytical Support of Interactive Visualizations in Virtual Reality Environments," in *Proceedings of the 2018 IEEE International*

Symposium on Mixed and Augmented Reality, ISMAR 2018, 2019, pp. 143–152, doi: 10.1109/ISMAR.2018.00050.

[188] F. Buttussi and L. Chittaro, “Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario,” *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 2, pp. 1063–1076, 2018, doi: 10.1109/TVCG.2017.2653117.

[189] I. Kosunen, M. Salminen, S. Järvelä, A. Ruonala, N. Ravaja, and G. Jacucci, “RelaWorld: Neuroadaptive and Immersive Virtual Reality Meditation System,” *Iui 2016*, pp. 208–217, 2016, doi: 10.1145/2856767.2856796.

[190] M. McGill, A. Ng, and S. Brewster, “I Am The Passenger: How Visual Motion Cues Can Influence Sickness For In-Car VR,” *Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. - CHI ’17*, no. May, pp. 5655–5668, 2017, doi: 10.1145/3025453.3026046.

[191] M. Walch *et al.*, “Evaluating VR Driving Simulation from a Player Experience Perspective,” *Proc. 2017 CHI Conf. Ext. Abstr. Hum. Factors Comput. Syst. - CHI EA ’17*, pp. 2982–2989, 2017, doi: 10.1145/3027063.3053202.

[192] A. I. Nordin, A. Denisova, and P. Cairns, “Too Many Questionnaires: Measuring Player Experience Whilst Playing Digital Games,” *Proc. Seventh York Dr. Symp. Comput. Sci. Electron.*, pp. 69–75, 2014, doi: 10.1007/s11031-006-9051-8.

[193] K. Salehzadeh Niksirat, C. Silpasuwanchai, M. Mohamed Hussien Ahmed, P. Cheng, and X. Ren, “A Framework for Interactive Mindfulness Meditation Using Attention-Regulation Process,” *Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. - CHI ’17*, pp. 2672–2684, 2017, doi: 10.1145/3025453.3025914.

[194] N. H. Liu, C. Y. Chiang, and H. C. Chu, “Recognizing the degree of human attention using EEG signals from mobile sensors,” *Sensors (Switzerland)*, vol. 13, no. 8, pp. 10273–10286, 2013, doi: 10.3390/s130810273.

[195] B. Güntekin and E. Başar, “Review of evoked and event-related delta responses in the human brain,” *Int. J. Psychophysiol.*, vol. 103, pp. 43–52, 2016, doi: 10.1016/j.ijpsycho.2015.02.001.

[196] M. Balconi and C. Lucchiari, “Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. A gamma band analysis,” *Int.*

J. Psychophysiol., vol. 67, no. 1, pp. 41–46, 2008, doi: 10.1016/j.ijpsycho.2007.10.002.

[197] T. McMahan, I. Parberry, and T. D. Parsons, "Evaluating Player Task Engagement and Arousal Using Electroencephalography," *Procedia Manuf.*, vol. 3, no. Ahfe, pp. 2303–2310, 2015, doi: 10.1016/j.promfg.2015.07.376.

[198] H. Krugman, "Brain wave measures of media involvement," *J. Advert. Res.*, vol. 11, no. 1, pp. 3–9, 1971.

[199] A. T. Pope, E. H. Bogart, and D. S. Bartolome, "Biocybernetic system evaluates indices of operator engagement in automated task," *Biol. Psychol.*, vol. 40, no. 1–2, pp. 187–195, 1995, doi: 10.1016/0301-0511(95)05116-3.

[200] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion Recognition from Brain Signals Using Hybrid Adaptive Filtering and Higher Order Crossings Analysis," *IEEE Trans. Affect. Comput.*, vol. 1, no. 2, pp. 81–97, 2010, doi: 10.1109/t-affc.2010.7.

[201] E. Schmidt, W. Kincses, and M. Schrauf, "Assessing driver's vigilance state during monotonous driving," *... Hum. Factors Driv. ...*, pp. 138–145, 2007, [Online]. Available: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Assessing+drivers?+vigilance+state+during+monotonous+driving#0.

[202] M. Hassib, S. Schneegass, P. Eiglsperger, N. Henze, A. Schmidt, and F. Alt, "EngageMeter," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 2017, pp. 5114–5119, doi: 10.1145/3025453.3025669.

[203] O. E. Krigolson, C. C. Williams, A. Norton, C. D. Hassall, and F. L. Colino, "Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research," *Front. Neurosci.*, vol. 11, no. MAR, pp. 1–10, 2017, doi: 10.3389/fnins.2017.00109.

[204] T. Mondéjar, R. Hervás, E. Johnson, C. Gutierrez, and J. M. Latorre, "Correlation between videogame mechanics and executive functions through EEG analysis," *J. Biomed. Inform.*, vol. 63, pp. 131–140, 2016, doi: 10.1016/j.jbi.2016.08.006.

[205] S. Higuchi, Y. Motohashi, Y. Liu, and A. Maeda, "Effects of playing a computer game using a bright display on presleep physiological variables, sleep latency,

slow wave sleep and REM sleep," *J. Sleep Res.*, vol. 14, no. 3, pp. 267–273, 2005, doi: 10.1111/j.1365-2869.2005.00463.x.

[206] S. E. Kober, J. Kurzmann, and C. Neuper, "Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: An EEG study," *Int. J. Psychophysiol.*, vol. 83, no. 3, pp. 365–374, 2012, doi: 10.1016/j.ijpsycho.2011.12.003.

[207] S. M. Slobounov, W. Ray, B. Johnson, E. Slobounov, and K. M. Newell, "Modulation of cortical activity in 2D versus 3D virtual reality environments: An EEG study," *Int. J. Psychophysiol.*, vol. 95, no. 3, pp. 254–260, 2015, doi: 10.1016/j.ijpsycho.2014.11.003.

[208] M. Clemente, A. Rodríguez, B. Rey, and M. Alcañiz, "Assessment of the influence of navigation control and screen size on the sense of presence in virtual reality using EEG," *Expert Syst. Appl.*, vol. 41, no. 4 PART 2, pp. 1584–1592, 2014, doi: 10.1016/j.eswa.2013.08.055.

[209] Interaxon, "Available Data - Muse Direct - Muse Developer," 2018. http://developer.choosemuse.com/tools/windows-tools/available-data-muse-direct (accessed Mar. 28, 2018).

[210] J. Ren, J. Jiang, and Y. Feng, "Activity-driven content adaptation for effective video summarization," *J. Vis. Commun. Image Represent.*, vol. 21, no. 8, pp. 930–938, 2010, doi: 10.1016/j.jvcir.2010.09.002.

[211] J. Chen, J. Ren, and J. Jiang, "Modelling of content-aware indicators for effective determination of shot boundaries in compressed MPEG videos," *Multimed. Tools Appl.*, vol. 54, no. 2, pp. 219–239, 2011, doi: 10.1007/s11042-010-0518-y.

[212] J. Ren and J. Jiang, "Hierarchical modeling and adaptive clustering for real-time summarization of rush videos," *IEEE Trans. Multimed.*, vol. 11, no. 5, pp. 906–917, Aug. 2009, doi: 10.1109/TMM.2009.2021782.

[213] W. Xu, H. N. Liang, Y. Zhao, T. Zhang, Di. Yu, and Di. Monteiro, "RingText: Dwell-free and hands-free Text Entry for Mobile Head-Mounted Displays using Head Motions," *IEEE Trans. Vis. Comput. Graph.*, vol. 25, no. 5, pp. 1991–2001, 2019, doi: 10.1109/TVCG.2019.2898736.

[214] D. Yu, K. Fan, H. Zhang, D. Monteiro, W. Xu, and H.-N. Liang, "PizzaText: Text Entry for Virtual Reality Systems Using Dual Thumbsticks," *IEEE Trans. Vis. Comput. Graph.*, vol. 24, no. 11, pp. 2927–2935, Nov. 2018, doi:

10.1109/TVCG.2018.2868581.

[215] D. C. Rachevsky, V. C. De Souza, and L. Nedel, “Visualization and interaction in immersive virtual reality games: A user evaluation study,” *Proc. - 2018 20th Symp. Virtual Augment. Reality, SVR 2018*, no. November, pp. 89–98, 2018, doi: 10.1109/SVR.2018.00024.

[216] R. McGloin and M. Krcmar, “The Impact of Controller Naturalness on Spatial Presence, Gamer Enjoyment, and Perceived Realism in a Tennis Simulation Video Game,” *Presence Teleoperators Virtual Environ.*, vol. 20, no. 4, pp. 309–324, Aug. 2011, doi: 10.1162/PRES_a_00053.

[217] S. W. Lee, J. W. An, and J. Y. Lee, “The Relationship between e-Sports Viewing Motives and Satisfaction: The Case of League of Legends,” 2014, doi: 10.15242/icehm.ed0214019.

[218] J. Rambusch, P. Jakobsson, and D. Pargman, “Exploring E-sports: A case study of gameplay in Counter-strike,” *3rd Digit. Games Res. Assoc. Int. Conf. "Situated Play. DiGRA 2007*, pp. 157–164, 2007.

[219] M. H. Amir, A. Quek, N. R. Bin Sulaiman, and J. See, “DUKE: Enhancing virtual reality based FPS game with full-body interactions,” *ACM Int. Conf. Proceeding Ser.*, no. August 2019, 2016, doi: 10.1145/3001773.3001804.

[220] J. L. Lugrin, F. Charles, M. Cavazza, M. Le Renard, J. Freeman, and J. Lessiter, “Are immersive FPS games enjoyable?,” *Proc. ACM Symp. Virtual Real. Softw. Technol. VRST*, p. 199, 2012, doi: 10.1145/2407336.2407378.

[221] J. L. Lugrin, M. Cavazza, F. Charles, M. Le Renard, J. Freeman, and J. Lessiter, “Immersive FPS games: User experience and performance,” *ImmersiveMe 2013 - Proc. 2nd Int. Work. Immersive Media Exp. Co-located with ACM Multimed. 2013*, no. October, pp. 7–12, 2013, doi: 10.1145/2512142.2512146.

[222] C. Yildirim, M. Carroll, D. Hufnal, T. Johnson, and S. Pericles, “Video Game User Experience: To VR, or Not to VR?,” *2018 IEEE Games, Entertain. Media Conf. GEM 2018*, pp. 125–131, 2018, doi: 10.1109/GEM.2018.8516542.

[223] B. Aykent, F. Merienne, C. Guillet, D. Paillot, and A. Kemeny, “Motion sickness evaluation and comparison for a static driving simulator and a dynamic driving simulator,” *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.*, vol. 228, no. 7, pp. 818–829, 2014, doi: 10.1177/0954407013516101.

[224] M. E. Mccauley, “Do Army Helicopter Training Simulators Need Motion

Bases?," *United States Army Res. Inst. Behav. Soc. Sci.*, no. February, pp. 01–52, 2006, [Online]. Available: https://www.google.com.pk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwil9em-qJnLAhWHCo4KHflYALwQFggZMAA&url=http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA444549&usg=AFQjCNGXLBE_HgX4uPrPE0GQY1yrItXXmg&sig2=ixCIE-etbmpbDqGKY4axxg&bvm=bv.

[225] J. Lee, S. Lim, J. Ahn, Y. Yi, and H. Kim, "Analysis of video image based element for motion sickness," *IS T Int. Symp. Electron. Imaging Sci. Technol.*, pp. 3–8, 2018, doi: 10.2352/ISSN.2470-1173.2018.03.ERVR-470.

[226] M. J. Cevette *et al.*, "Oculo-Vestibular Recoupling Using Galvanic Vestibular Stimulation to Mitigate Simulator Sickness," *Aviat. Space. Environ. Med.*, vol. 83, no. 6, pp. 549–555, Jun. 2012, doi: 10.3357/ASEM.3239.2012.

[227] R. J. Reed-Jones, J. G. Reed-Jones, L. M. Trick, and L. A. Vallis, "Can Galvanic Vestibular Stimulation Reduce Simulator Adaptation Syndrome?," in *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle*, Oct. 2007, pp. 534–540, doi: 10.17077/drivingassessment.1288.

[228] S. Weech, J. Moon, and N. F. Troje, "Influence of bone-conducted vibration on simulator sickness in virtual reality," *PLoS One*, vol. 13, no. 3, pp. 1–21, 2018, doi: 10.1371/journal.pone.0194137.

[229] S. Liu, L. Chan, and M. Y. Chen, "PhantomLegs : Reducing Virtual Reality Sickness using Head-Worn Haptic Devices," 2019.

[230] D. Monteiro, H. N. Liang, W. Xu, M. Brucker, V. Nanjappan, and Y. Yue, "Evaluating enjoyment, presence, and emulator sickness in VR games based on first- and third- person viewing perspectives," *Comput. Animat. Virtual Worlds*, vol. 29, no. 3–4, p. e1830, May 2018, doi: 10.1002/cav.1830.

[231] L. Pomarjanschi, M. Dorr, P. J. Bex, and E. Barth, "Simple gaze-contingent cues guide eye movements in a realistic driving simulator," *Hum. Vis. Electron. Imaging XVIII*, vol. 8651, p. 865110, 2013, doi: 10.1117/12.2002833.

[232] G.-Y. Nie, H. B.-L. Duh, Y. Liu, and Y. Wang, "Analysis on Mitigation of Visually Induced Motion Sickness by Applying Dynamical Blurring on a User's Retina," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 8, pp. 2535–2545, Aug.

2020, doi: 10.1109/TVCG.2019.2893668.

[233] R. Lou and J. R. Chardonnet, “Reducing Cybersickness by Geometry Deformation,” *26th IEEE Conf. Virtual Real. 3D User Interfaces, VR 2019 - Proc.*, pp. 1058–1059, 2019, doi: 10.1109/VR.2019.8798164.

[234] A. S. Fernandes and S. K. Feiner, “Combating VR sickness through subtle dynamic field-of-view modification,” *2016 IEEE Symp. 3D User Interfaces, 3DUI 2016 - Proc.*, pp. 201–210, 2016, doi: 10.1109/3DUI.2016.7460053.

[235] Y. Farmani and R. J. Teather, “Viewpoint snapping to reduce cybersickness in virtual reality,” *Proc. - Graph. Interface*, vol. 2018-May, pp. 159–166, Oct. 2018, doi: 10.20380/GI2018.21.

[236] P. Hu, Q. Sun, P. Didyk, L. Y. Wei, and A. E. Kaufman, “Reducing simulator sickness with perceptual camera control,” *ACM Trans. Graph.*, vol. 38, no. 6, pp. 1–12, Nov. 2019, doi: 10.1145/3355089.3356490.

[237] Y. Farmani and R. J. Teather, “Evaluating discrete viewpoint control to reduce cybersickness in virtual reality,” *Virtual Real.*, no. 0123456789, 2020, doi: 10.1007/s10055-020-00425-x.

[238] S. Davis, K. Nesbitt, and E. Nalivaiko, “A Systematic Review of Cybersickness,” in *Proceedings of the 2014 Conference on Interactive Entertainment - IE2014*, 2014, vol. 19, pp. 1–9, doi: 10.1145/2677758.2677780.

[239] D. Saredakis, A. Szpak, B. Birckhead, H. A. D. Keage, A. Rizzo, and T. Loetscher, “Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis,” *Front. Hum. Neurosci.*, vol. 14, no. March, 2020, doi: 10.3389/fnhum.2020.00096.

[240] N. Dużmańska, P. Strojny, and A. Strojny, “Can Simulator Sickness Be Avoided? A Review on Temporal Aspects of Simulator Sickness,” *Front. Psychol.*, vol. 9, no. NOV, Nov. 2018, doi: 10.3389/fpsyg.2018.02132.

[241] S. Martirosov and P. Kopecek, “Cyber sickness in virtual reality - Literature review,” *Ann. DAAAM Proc. Int. DAAAM Symp.*, no. January, pp. 718–726, 2017, doi: 10.2507/28th.daaam.proceedings.101.

[242] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, A. Liberati, and D. G. Altman, “Reprint-Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement,” 2009. Accessed: Dec. 14, 2020. [Online]. Available: http://www.annals.org/cgi/content/full/151/4/264.

[243] B. H. Thomas, D. Ciliska, M. Dobbins, and S. Micucci, "A Process for Systematically Reviewing the Literature: Providing the Research Evidence for Public Health Nursing Interventions," 2004.

[244] J. C. Valentine, T. D. Pigott, and H. R. Rothstein, "How many studies do you need? A primer on statistical power for meta-analysis," *J. Educ. Behav. Stat.*, vol. 35, no. 2, pp. 215–247, 2010, doi: 10.3102/1076998609346961.

[245] M. Isaza, J. Zhang, K. Kim, C. Mei, and R. Guo, "Mono-stereoscopic camera in a virtual reality environment: Case study in cybersickness," *2019 11th Int. Conf. Virtual Worlds Games Serious Appl. VS-Games 2019 - Proc.*, p. 1DUUMY, 2019, doi: 10.1109/VS-Games.2019.8864578.

[246] H. Buhler, S. Misztal, and J. Schild, "Reducing VR Sickness Through Peripheral Visual Effects," *25th IEEE Conf. Virtual Real. 3D User Interfaces, VR 2018 - Proc.*, pp. 517–519, 2018, doi: 10.1109/VR.2018.8446346.

[247] K. Carnegie and T. Rhee, "Reducing Visual Discomfort with HMDs Using Dynamic Depth of Field," *IEEE Comput. Graph. Appl.*, vol. 35, no. 5, pp. 34–41, 2015, doi: 10.1109/MCG.2015.98.

[248] C. Zhou, C. L. Bryan, E. Wang, N. S. Artan, and Z. Dong, "Cognitive Distraction to Improve Cybersickness in Virtual Reality Environment," *Proc. - 2019 IEEE 16th Int. Conf. Mob. Ad Hoc Smart Syst. Work. MASSW 2019*, pp. 72–76, 2019, doi: 10.1109/MASSW.2019.00021.

[249] S. Kim, S. Lee, N. Kala, J. Lee, and W. Choe, "An effective FoV restriction approach to mitigate VR sickness on mobile devices," *J. Soc. Inf. Disp.*, vol. 26, no. 6, pp. 376–384, 2018, doi: 10.1002/jsid.669.

[250] O. Basting, A. Fuhrmann, and S. M. Grünvogel, "The effectiveness of changing the field of view in a HMD on the perceived self-motion," *2017 IEEE Symp. 3D User Interfaces, 3DUI 2017 - Proc.*, pp. 225–226, 2017, doi: 10.1109/3DUI.2017.7893353.

[251] M. Al Zayer, I. B. Adhanom, P. MacNeilage, and E. Folmer, "The effect of field-of-view restriction on sex bias in VR sickness and spatial navigation performance," *Conf. Hum. Factors Comput. Syst. - Proc.*, 2019, doi: 10.1145/3290605.3300584.

[252] N. Norouzi, G. Bruder, and G. Welch, "Assessing vignetting as a means to reduce VR sickness during amplified head rotations," *Proc. - SAP 2018 ACM*

Symp. Appl. Percept., 2018, doi: 10.1145/3225153.3225162.

[253] D. Zielasko and S. Freitag, “Dynamic Field of View Reduction Related to Subjective Sickness Measures in an HMD-based Data Analysis Task,” *VR IEEE 2018 Work.*, no. March, 2018.

[254] Z. Cao, J. Jerald, and R. Kopper, “Visually-Induced Motion Sickness Reduction via Static and Dynamic Rest Frames,” *25th IEEE Conf. Virtual Real. 3D User Interfaces, VR 2018 - Proc.*, no. March, pp. 105–112, 2018, doi: 10.1109/VR.2018.8446210.

[255] C. Wienrich, C. K. Weidner, C. Schatto, D. Obremski, and J. H. Israel, “A virtual nose as a rest-frame-the impact on simulator sickness and game experience,” *2018 10th Int. Conf. Virtual Worlds Games Serious Appl. VS-Games 2018 - Proc.*, pp. 1–8, 2018, doi: 10.1109/VS-Games.2018.8493408.

[256] B. Keshavarz and H. Hecht, “Validating an Efficient Method to Quantify Motion Sickness,” *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 53, no. 4, pp. 415–426, Aug. 2011, doi: 10.1177/0018720811403736.

[257] G. Gigerenzer, “Mindless statistics,” *J. Socio. Econ.*, vol. 33, no. 5, pp. 587–606, 2004, doi: 10.1016/j.socec.2004.09.033.

[258] R. Zhang, X. Liu, Y. Guo, and S. H. B, “Image Synthesis with Aesthetics-Aware,” vol. 2, no. 2017, pp. 169–179, 2018, doi: 10.1007/978-3-030-00767-6.

[259] F. Pallavicini, A. Pepe, and M. E. Minissi, “Gaming in Virtual Reality: What Changes in Terms of Usability, Emotional Response and Sense of Presence Compared to Non-Immersive Video Games?,” *Simul. Gaming*, vol. 50, no. 2, pp. 136–159, 2019, doi: 10.1177/1046878119831420.

[260] I. Hupont, J. Gracia, L. Sanagustin, and M. A. Gracia, “How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift,” *2015 7th Int. Work. Qual. Multimed. Exp. QoMEX 2015*, no. July, 2015, doi: 10.1109/QoMEX.2015.7148110.

[261] E. Martel, A. Hassan, F. Su, A. Girouard, J. Gerroir, and K. Muldner, “Diving Head-First into Virtual Reality—Evaluating HMD Control Schemes for VR Games,” no. Fdg, 2015.

[262] W. J. Shelstad, D. C. Smith, and B. S. Chaparro, “Gaming on the rift: How virtual reality affects game user satisfaction,” *Proc. Hum. Factors Ergon. Soc.*, vol. 2017-Octob, pp. 2072–2076, 2017, doi: 10.1177/1541931213602001.

[263] E. Begue, “Panda Behaviour Tree Scripting for Unity,” 2019. http://www.pandabehaviour.com/?page_id=246.

[264] E. D. Ragan, S. Scerbo, F. Bacim, and D. A. Bowman, “Amplified Head Rotation in Virtual Reality and the Effects on 3D Search, Training Transfer, and Spatial Orientation,” *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 8, pp. 1880–1895, 2017, doi: 10.1109/TVCG.2016.2601607.

[265] S. Hillaire, A. Lécuyer, R. Cozot, and G. Casiez, “Depth-of-field blur effects for first-person navigation in virtual environments,” *IEEE Comput. Graph. Appl.*, vol. 28, no. 6, pp. 47–55, Nov. 2008, doi: 10.1109/MCG.2008.113.

[266] A. Toet, S. C. de Vries, M. L. van Emmerik, and J. E. Bos, “Cybersickness and desktop simulations: field of view effects and user experience,” *Enhanc. Synth. Vis. 2008*, vol. 6957, no. April, p. 69570P, 2008, doi: 10.1117/12.771992.

[267] J. E. Bos, S. C. de Vries, M. L. van Emmerik, and E. L. Groen, “The effect of internal and external fields of view on visually induced motion sickness,” *Appl. Ergon.*, vol. 41, no. 4, pp. 516–521, 2010, doi: 10.1016/j.apergo.2009.11.007.

Appendix 1 – Questionnaire on Affection and Realism

Please answer the following questions from 0 (the least) to.100 (the most)

1. How much do you think the agents affect you?
2. How much do you think you affect the agents?
3. How much do you think the agents affect each other?
4. How easy was it for you to fulfil the task?
5. How comfortable were you in the system?
6. How much do you feel present in this task?
7. Overall, how would you rate the realism of the virtual crowds?
8. Overall, how would you rate the realism of the conversation contents?
9. Overall, how would you rate the realism of the environment?

Appendix 2 – Simulator Sickness Questionnaire

Please indicate how you felt during the condition for each of the items. Rate from 0 (none) to 3 (severe).

1. General discomfort
2. Fatigue
3. Headache
4. Eye strain
5. Difficulty focusing
6. Salivation increasing
7. Sweating
8. Nausea
9. Difficulty concentrating
10. Fullness of the Head
11. Blurred vision
12. Dizziness with eyes open
13. Dizziness with eyes closed
14. Vertigo (vertigo is experienced as loss of orientation with respect to vertical upright)
15. Stomach awareness (stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea)
16. Burping

Appendix 3 – In-game Questionnaire

Please indicate how you felt while playing the game for each of the items. Rate from 1 (not at all) to 5 (extremely).

1. I was interested in the game's story
2. I felt successful
3. I felt bored
4. I found it impressive
5. I forgot everything around me
6. I felt frustrated
7. I found it tiresome
8. I felt irritable
9. I felt skilful
10. I felt completely absorbed
11. I felt content
12. I felt challenged
13. I felt stimulated
14. I felt good

Appendix 4 – Post Game Experience Questionnaire

1. I felt revived
2. I felt bad
3. I found it hard to get back to reality
4. I felt guilty
5. It felt like a victory
6. I found it a waste of time
7. I felt energised
8. I felt satisfied
9. I felt disoriented
10. I felt exhausted
11. I felt that I could have done more useful things
12. I felt powerful
13. I felt weary
14. I felt regret
15. I felt ashamed
16. I felt proud
17. I had a sense that I had returned from a journey

Appendix 5 – immersive Experience Questionnaire

Please indicate how you felt while playing the game for each of the items. Rate from 1 (Not at all/ Very little/ Very poor) to 5 (A lot/ Very much/ Very well).

1. At any point did you find yourself become so involved that you were unaware you were even using controls?
2. To what extent did you feel as though you were moving through the game according to you own will?
3. To what extent did you find the game challenging?
4. Were there any times during the game in which you just wanted to give up?
5. To what extent did you feel motivated while playing?
6. To what extent did you find the game easy?
7. To what extent did you feel like you were making progress towards the end of the game?
8. How well do you think you performed in the game?
9. To what extent did you feel emotionally attached to the game?
10. To what extent were you interested in seeing how the game’s events would progress?
11. How much did you want to ‘‘win’’ the game?
12. Were you in suspense about whether or not you would win or lose the game?
13. At any point did you find yourself become so involved that you wanted to speak to the game directly?
14. To what extent did you enjoy the graphics and the imagery?
15. How much would you say you enjoyed playing the game?
16. When interrupted, were you disappointed that the game was over?
17. Would you like to play the game again?

Appendix 6 – Personal Report of Confidence as a Speaker

1. My hands tremble when I try to handle objects on the platform
2. I am in constant fear of forgetting my speech
3. While preparing a speech I am in a constant state of anxiety
4. My thoughts become confused and jumbled when I speak before an audience
5. Although I talk fluently with friends, I am at a loss for words on the platform
6. The faces of my audience are blurred when I look at them
7. I feel disgusted with myself after trying to address a group of people
8. I perspire and tremble just before getting up to speak
9. My posture feels strained and unnatural
10. I am fearful and tense all the while I am speaking before a group of people
11. It is difficult for me to search my mind calmly for the right words to express my thoughts
12. I am terrified at the thought of speaking before a group of people

Appendix 7 – Self-Statements During Public Speaking

1. What do I have to lose it's worth a try
2. This is an awkward situation but I can handle it Even if things don't go well, it's no catastrophe
3. I can handle everything
4. Instead of worrying I could concentrate on what I want to say
5. I'm a loser
6. A failure in this situation would be more proof of my incapacity
7. What I say will probably sound stupid I'll probably “bomb out” anyway
8. I feel awkward and dumb; they're bound to notice

Appendix 8 – Common European Frame of Reference

Please indicate based on these descriptions the level that best describes your English knowledge

Beginner	**A1**	Can recognise and use familiar words and simple phrases for concrete purposes. Can introduce himself or someone else. Can ask and answer basic questions about home, family, surroundings, etc. Can communicate in a basic way when the other person speaks slowly and clearly, and is ready to repeat or reformulate to help communication.
Elementary	**A2**	Can understand isolated phrases and common expressions that relate to areas of high personal relevance (like personal or family information, shopping, immediate environment, work). Can communicate during easy or habitual tasks requiring a basic and direct information exchange on familiar subjects. Using simple words, can describe his or her surroundings and communicate immediate needs.
Intermediate	**B1**	Can understand the main points of clear standard speech on familiar subjects in work, school, leisure activities, etc. Can manage in most situations that come up when travelling in a region where the language is spoken. Can produce a simple and cohesive text on familiar subjects or subjects of personal interest. Can narrate an event, an experience or a dream; describe a desire or goal, and outline reasons or explanations behind a project or idea.
Upper Intermediate	**B2**	Can understand the main ideas of concrete or abstract topics in a complex text, including a technical article in the user's area of expertise. Can communicate with a degree of spontaneity and fluency during a conversation with a native speaker, in a way that is comfortable for everyone. Can speak in a clear, detailed way on a number of subjects; express an opinion on current affairs, giving the advantages and disadvantages of the various options.
Advanced	**C1**	Can understand a wide range of long and complex texts, including any subtextual or stylistic nuances. Can express him or herself freely and fluidly, without obviously fumbling for words. Can use the language effectively and fluently in a social, professional, or academic context. Can speak in a clear, organised way about complex subjects, developing a well-structured argument.
Master or Proficient	**C2**	Can effortlessly understand almost everything he or she reads or hears. Capable of a coherent summary of events or arguments from oral or written sources. Can express him or herself precisely in a spontaneous, fluent way, conveying finer shades of meaning precisely.

Appendix 9 – Presentation Text 1 Chapter 4 and Chapter 5

Dear audience, thank you very much for your attention.

Star Wars, Alien, Ad Astra, 2001: A Space Odyssey. Recent decades have seen a plethora of science fiction movies exploring the idea of space travel. Today I will be talking about the new odyssey for humankind called interstellar travel.

Interstellar travel is the travelling between stars or planetary systems. Interstellar travel is much more difficult than traveling between planets in our own solar system.

The reason for this is that the distance between stars or planetary systems is much larger than the distance between planets in the same planetary system. The distances between stars, for example, can be hundreds of thosuands of light-years.

Unfortunately, we do not currently have the advanced technology required to travel such great distances. One problem is that our spaceships do not have the speed required to cover such great distances. Another problem is that if the spaceship were to collide with dust or gas is space there would be very serious effects on both the spaceship and the passengers inside.

Yet, despite these dangers, scientists have argued the importance of discovering and traveling to earth-like planets.

There are several steps humanity can take to develop the advanced technology needed to make interstellar travel a reality. First, such a large scale project would require the cooperation and collaboration of governments from as many countries as possible
This would allow for the pooling of resources, including human talent such as scientists and engineers.

Second, clear goals for interstellar travel must be outlined. Earth-like planets must first be identified. A trajectory or path to those planets must be planned. In other words, the most important thing if for governments to agree where we are going and how we will go there. The last step is coming up with the technology that will enable humanity to complete the journey.

While interstellar travel certainly seems difficult, it may be necessary to ensure the survival of humanity. If we begin working together now, the science fiction of today will become the reality of tomorrow.

Appendix 10 – Presentation Text 2 Chapter 4 and Chapter 5

Hello ladies and gentlemen, thank you for coming to my presentation, today my topic is the Neanderthals

The Neanderthals were a group of humans that diverged from modern humans about 300,000 years ago. The Neanderthals came to their end about 30,000 years ago, at the hands of modern humans.

Their bodies were well adapted to the cold. Their bodies were short and stocky, about 20 percent heavier than modern humans of the same height.

Even though they had larger brains, it is likely that their brain structure was different from ours. They had thicker skin and a larger layer of fat.

Evidence suggest that they lived in smaller more local groups that lived on smaller ranges of land. They were not able to create very organized camps, and therefore only stayed in a camp for a short period of time.

This is visible by their less organized homes. Despite their lack of organized camps, the Neanderthals possessed many ornamental objects. This suggests the existence of culture. It is also important to note that in regards to labor there seemed to be a degree of gender equality.

There is evidence that they were able to make tools using relatively advanced techniques. There is little evidence that they exchanged raw materials with others outside their group.

The Neanderthals’ disappearance can be attributed to diverse causes. Evidence suggests that encounters with modern humans were sometimes violent.

This is evident by broken bones found in both groups. Other times seemed to be peaceful between the two kinds of humans.

There is evidence that modern humans' DNA has a percentage of Neanderthal DNA, suggesting the groups may have mixed.

In the end it seems that Neanderthals came to their end because modern humans took control over their food sources

.

Appendix 11 –Bullet Points 1 Chapter 4 and Chapter 5

- Hello

- -Many movies explore space travel -New odyssey: interstellar travel

-Interstellar travel = travelling between stars -very difficult
-large distances -take a long time

-don't have the technology -distances too great -dangerous – collision with dust or gas

-important to discover earth-like planets

-Steps to develop technology -first – cooperation of governments
-resources and human talent

-Second – clear goals -identify earth-like planets -path planned-governments need to agree-develop technology needed

-difficult but necessary for survival, -work together = make it a reality

.

Appendix 12 –Bullet Points 2 Chapter 4 and Chapter 5

- Hello

-*Another group of humans

*300,000-30,00 years ago *Died due to modern humans

- *Adapted to the cold *Short and stocky, heavier 20%

- *Larger brains, but structure similar *Thicker skin, more fat

- *small local groups *not very organized *temporary camps

- *many ornamental objects, suggests culture *gender equality in labour

- *tool making – advanced *unlikely to exchange materials with groups

- *Several causes *Violent encounters – modern humans

- *broken bones in both groups

- *Sometimes peaceful – DNA

- *Ran out of food sources due to modern humans.

Appendix 13 – Questionnaire used in Chapter 3

Pre-experiment questionnaire:

1. Participant Number (given by researcher)
2. Your age
3. Gender
    a. Female
    b. Male
    c. Non-Binary
    d. Prefer not to say
4. Are you aware and can disclose any health issues that might affect the experiment, including but not limited to colour-blindness or epilepsy?
    a. No
    b. Yes
5. Do you (or have you ever) played computer games or video games?
    a. No
    b. Yes
6. Do you (or have you ever) played games in Virtual-Reality?
    a. No
    b. Yes

Post-condition questionnaire:

7. Questionnaire on Affection and Realism
8. Did you cheat on the game?
9. If you cheated, did you feel guilty?
10. How did you feel when she was[happy, angry, sad, bored]
11. Simulator Sickness Questionnaire

Post-experiment questionnaire:

12. Order the versions by preference
13. Do you think her expressions reflect the right feelings?
14. Which version made you feel the most competitive?
15. Which of her emotions made you feel the most competitive?

Appendix 14 – Questionnaire used in Chapter 4

Pre-experiment questionnaire:

1. Participant Number (given by researcher)
2. Your age
3. Gender
    a. Female
    b. Male
    c. Non-Binary
    d. Prefer not to say
4. Are you aware and can disclose any health issues that might affect the experiment, including but not limited to colour-blindness, motor disorders or epilepsy?
    a. No
    b. Yes
5. Are you a native English speaker?
    a. No
    b. Yes
6. Based on the Common European Frame of Reference presented to you, what is the level that best describes your English skills
7. Do you play or have you ever played computer games or video games?
    a. No
    b. Yes
8. Do you have experience with Virtual-Reality?
    a. No
    b. Yes
9. Personal Report of Confidence as a Speaker

Post-conditions questionnaire:

10. Self-Statements During Public Speaking
11. How familiar were you with the topics?
12. Rate this condition from 1 (worst) to 5 (best)

Post-experiment questionnaire:

13. Rank your preparation after each version
14. Rank your nervousness after each version

15. Rank your version preference
16. How do you feel about the virtual environment, not including the audience?
17. How do you feel regarding the virtual audience?
18. Is there anything you would change in the virtual Environment?

Appendix 15 – Questionnaire used in Chapter 5

Pre-experiment questionnaire:

1. Participant Number (given by researcher)
2. Your date of birth
3. Gender
    a. Female
    b. Male
4. What is your first language?
5. How would you describe your English language ability?
    a. Beginner
    b. Elementary
    c. Intermediate
    d. Advanced
6. I feel confident when I speak English in front of others
    a. 1 (Disagree) , 2, 3, 4, 5(Agree)
7. I am shy when I have to give presentations in English
    a. 1 (Disagree) , 2, 3, 4, 5(Agree)
8. I feel positive about my ability to give presentations in English
    a. 1 (Disagree) , 2, 3, 4, 5(Agree)
9. Do you have experience with Virtual-Reality?
    a. No
    b. Yes

Post-condition questionnaire

10. Personal Report of Confidence as a Speaker

Post-experiment questionnaire:

11. I would like to continue using this VR program to practice my presentations.
12. The VR program makes me want to get better at giving presentations.
13. The VR program motivates me to practice giving presentations.
14. The VR program helps me to be more focused when giving presentations.
15. think the VR program is distracting when giving presentations.
16. It is easy to focus on presenting while using the VR program.Rank your nervousness after each version

17. Did you notice any similarities between the VR audience and the real audience? How did this affect you while giving a presentation?
18. Was giving a presentation to a VR audience different from giving a presentation to a real audience? In what way? How did this affect you while giving a presentation?

Appendix 16 – Questionnaire used in Chapter 6 and 7

Pre-experiment questionnaire:

19. Participant Number
20. Age
21. Gender
    a. Female
    b. Male
    c. Prefer not to say
22. How often do you generally play video-games?
    a. At least once a day
    b. 6-3 days a week
    c. 1 or 2 days a week
    d. Once every two weeks
    e. Once a month
    f. Almost never
23. Have you ever played VR games?
    a. Yes.
    b. No
24. Have you ever played Mario Kart?.
    a. Yes
    b. No

Post-condition questionnaire

25. Simulator Sickness Questionnaire
26. In-game questionnaire
27. Post-Game Experience Questionnaire

Post-experiment questionnaire :

28. Which version did you like the most?

Appendix 17 – Questionnaire used in Chapter 8

Pre-experiment questionnaire:

1. Participant Number
2. Age
3. Gender
    a. Female
    b. Male
    c. Prefer not to say
4. How often do you generally play videogames?
    a. At least once a day
    b. 6-3 days a week
    c. 1 or 2 days a week
    d. Once every two weeks
    e. Once a month
    f. Almost never
5. How often do you generally play First-person Shooters (FPS)?
    a. At least once a day
    b. 6-3 days a week
    c. 1 or 2 days a week
    d. Once every two weeks
    e. Once a month
    f. Almost never
6. How many times have you played VR games?
    a. 0.
    b. 1-2
    c. 2-20
    d. 20+
7. Do you ever feel sick playing FPS in the computer/video game?
    a. Yes
    b. No

Post-condition questionnaire

8. Simulator Sickness Questionnaire
9. In-game questionnaire
10. Post-Game Experience Questionnaire

Post-experiment questionnaire :

11. Which version did you like the most?