UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Characterising Embodiment in Multi-Modal Play for Virtual Reality

by

Yoan-Daniel Malinov
[BSc]
ORCiD: 0000-0003-3263-3614

A thesis for the degree of
Doctor of Philosophy

April 2025

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Doctor of Philosophy

Characterising Embodiment in Multi-Modal Play for Virtual Reality

by Yoan-Daniel Malinov

The surge in Virtual Reality's (VR) popularity has sparked a growing interest in co-located multi-modal gaming experiences to enhance inclusivity. This thesis delves into the effects of integrating a non-VR player into these experiences, focusing on how it influences embodiment, immersion, and co-presence among participants.

The research unfolds in three main studies. In the initial study, a multi-modal VR game, "StuckInSpace", serves as the testing ground for introducing a non-VR player through different mediums: a PC or a tracked Phone. The results from this study ($n = 24$) reveal notably no significant differences for the VR player, with the qualitative analysis highlighting the critical role embodiment plays in shaping the experience.

Building on these findings, a systematic literature review is conducted, leading to the development of two orthogonal models of embodiment: *Aspects* and *Levels*. These models provide various strategies to influence different aspects of embodiment in multi-modal experiences.

In the third and final study, the knowledge gained from the preceding research is applied to modify the game by incorporating some of the identified embodiment strategies. The objective is to assess the effectiveness of these strategies and explore any cross-mode embodiment effects ($n = 48$). Surprisingly, the results indicate an unexpected influence of the "Narrative" strategy used to explain the other two strategies, overpowering their intended effects.

This thesis stands as a significant contribution to the field, presenting a multi-modal VR game and unveiling two comprehensive embodiment models, along with actionable strategies for designing multi-modal games. Additionally, it sheds light on the intricate and nuanced relationship between embodiment, co-presence, and immersion within such experiences, and emphasises the importance of mixed-methods research. The findings pave the way for future research in co-located multi-modal VR experiences, underscoring the importance of thoughtful design and the potential of narrative to enhance user engagement.

Contents

List of Figures

List of Tables

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
   Yoan-Daniel Grigorov Malinov. "Characterising the Benefits of Multi-Modal Play in Virtual Reality". In: *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '20. Virtual Event, Canada: Association for Computing Machinery, 2020, pp. 10–11. ISBN: 9781450375870. DOI: 10.1145/3383668.3419955
   Yoan-Daniel Malinov et al. "StuckInSpace: Exploring the Difference Between Two Different Mediums of Play in a Multi-Modal Virtual Reality Game". In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 2021, pp. 501–510. DOI: 10.1109/VR50410.2021.00074

Signed:.. Date:..................

Acknowledgements

I am deeply grateful to my two supervisors, Dr. David Millard and Dr. Tom Blount, whose expert guidance and insightful feedback were of tremendous help throughout my PhD journey. Their support and mentorship have been invaluable in making this thesis possible.

Chapter 1

Introduction

1.1 Motivation

As Virtual Reality (VR) is becoming more popular and widespread, there is a growing amount of research that explores the VR experience and its various applications. VR is being used in education [18, 89, 21], healthcare [59, 49, 52], remote collaboration [77], and making the user feel more present in another space through virtual reality [41] (for example in a teleconference). Nevertheless, the majority of people associate VR with gaming [63].

In the realm of gaming, VR serves as a tool to unite people in a shared virtual space, bridging the physical gap between them.

Examples include multiplayer shooter *Pavlov* [88] and online chatroom *VRChat* [90], where players can connect remotely and share the same virtual space, communicating via voice-chat and body language where possible. This works well for remote collaboration, but falls short when looking at co-located experiences. In these cases, the HMD (Head-Mounted Display) user is isolated from the physical environment and anyone else that wants to participate is left either looking at a screen while waiting for their turn on the headset, or trying to communicate with the HMD player, potentially disrupting their immersion.

Of course, there are a number of local multiplayer VR games, but they are quite low in number. Examples include *Keep Talking and Nobody Explodes* [83] and *The Playroom VR* [86], where the second player is either reading a manual to help the VR player, or playing with a controller while watching on a TV screen. These games show that involving a second player, directly or indirectly, is a valid way to make VR experiences more inclusive to others, even when only one head-mounted display is available. For instance, Keep Talking and Nobody Explodes demonstrates that even a simple set of printed instructions can make for a fun and engaging experience.

This thesis focuses on **designing for co-located multi-modal VR experiences**. Given the aforementioned issues with current co-located VR experiences and the increasing number of multi-modal spaces, there is a need for more research into best practices and design strategies. Each mode, from simple AR with a phone to fully immersive VR, poses unique challenges for creating engaging and immersive experiences.

When talking about VR, the concepts of immersion, presence and embodiment are frequently discussed. In this thesis, "immersion" will be used as defined by Slater and Wilbur [80]: *"the extent to which the technology can deliver an illusion of reality to the senses of a human participant"* [80]. Research has shown that immersion is a desirable property of VR [68, 47, 75], improving both collaboration [60] and enjoyment [38]. If immersion is about the properties of a system, then "presence" is how those properties makes you feel, *"the sensation of 'being there' "*[78]. Lastly, embodiment is about experiencing external properties as if they were the properties of one's own body [42].

Embodiment, immersion, and presence are closely intertwined, with each influencing the others, yet remaining separate phenomena. Understanding these differences is crucial when designing a multi-modal experience, as each mode has specific criteria for immersion, presence, and embodiment, needing targeted design strategies to ensure an enjoyable experience. As Gugenheimer et al. state, to become "*part of the social living room environment*" [36, 35], VR will need to accommodate a wide range of experiences, from passive spectatorship to more active participation. Careful design is crucial, as reducing the immersion of the VR user due to interactions with others would undermine the purpose of the technology.

1.2 Objectives and Research Questions

The primary objective of this thesis is to further the understanding of co-located multi-modal experiences in terms of design, embodiment, immersion, and co-presence. Specifically, it aims to explore how HMD and non-HMD users can interact in such a way as to not compromise the experience, and to determine if such interaction is desirable. Learning why embodiment is important and how one can use it to improve the experience is another objective.

To make it clearer, the following three **Research Questions** have been formulated:

- **RQ1**. How does a non-HMD user's mode of play in a multi-modal VR game affect the Immersion and Co-Presence of all participants?
- **RQ2**. What are the different kinds of embodiment and what are the different strategies that are used to help achieve them?

- **RQ3**. How can design strategies impact the different modes of play in terms of Embodiment?

In this case, the term HMD user refers to Head-Mounted Display user, or in other words the person who is wearing the VR headset, while non-HMD user is everyone else (Phone or PC user for example). Multi-modal games involve multiple modalities, such as a VR game with a second player on a PC. Here, two modalities are present: the one of the person wearing the headset, and the other of the PC player looking at a screen.

“Different kinds of embodiment” refers to the different types of embodiment that the literature presents and are found after a review, while “different strategies” or “design strategy” are about the ways you can affect embodiment. For example, a strategy could be changing a small variable in a VR game to make the experience feel more embodied, or adding a new interaction in the gameplay.

With these questions in mind, the answers would help contribute to the understanding of co-located multi-modal spaces and multi-modal interactions in VR, provide important insight into the importance of embodiment, and potentially inform the design of future immersive virtual experiences that can cater to a diverse range of users and play styles, ultimately enhancing the overall efficacy and enjoyment of multi-modal VR gaming.

1.3 Challenges in Multi-Modal Experimentation

One of the significant challenges in conducting experiments involving multi-modal VR experiences is managing the mixed effects inherent in such environments. Each modality, be it VR, PC or Phone, introduces its own set of variables that can interact in complex ways. This raises an important question for future research and the design of multi-modal experiences: should factors be isolated and studied separately, or should a more “real-world” approach be taken, analysing the entire experience as a whole?

In real-world experiences, users interact with multiple modalities at the same time, and their experiences are influenced by a number of factors. Isolating variables in a controlled experiment can help with understanding specific effects but might not fully capture the nuanced interactions that can happen in the “real-world” applications. On the other hand, a more “real-world” approach that takes into account the interconnectedness of different factors and elements could offer a fuller understanding but may introduce confounding variables that are difficult to control or measure accurately.

This thesis tries to adopt a mixed-methods approach to balance these considerations. By combining quantitative and qualitative data, it attempts to capture the complexity of multi-modal experiences while also identifying specific factors that influence immersion, co-presence, and embodiment. Future research could further explore the balance between the isolated and "real-world" approaches, determining the most effective method for studying and designing multi-modal VR experiences.

1.4 Research Framework

Having established the research questions, the research framework is depicted in Figure 1.1. In order to answer each question, a number of studies and experiments were undertaken.

For the first question, an experimental study was done to understand how two different modalities of interacting with the VR player affect all participants, looking specifically at immersion and co-presence as variables. The two modes tested alongside VR were either with a novel approach where a phone that is being tracked inside the space is used and displays what it sees from the current position in VR, or a more traditional PC setup. A game called StuckInSpace was created and used for the purposes of the study, with a sample size of 24 people, divided in half for the two versions of the game.

The second question led to a Systematic Literature Review in order to more deeply understand embodiment, how it's used in the literature, and the different ways of affecting it. It led to the creation of an extended embodiment framework consisting of two orthogonal models, each consisting of sub-components, and supporting strategies for influencing said components.

This was followed by the last experiment using the same base game as the first study, StuckInSpace, but now modified according to the strategies in the embodiment framework and the results from the first experiment, with the objective of answering RQ3, evaluating the impact of the chosen design strategies on the embodiment in different modes of play (sample size of 48 people).

1.5 Publications

Publications that came from this work to date include:

- Malinov, Y.-D. (2020) *Characterising the Benefits of Multi-Modal Play in Virtual Reality.* in Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '20, page 10–11

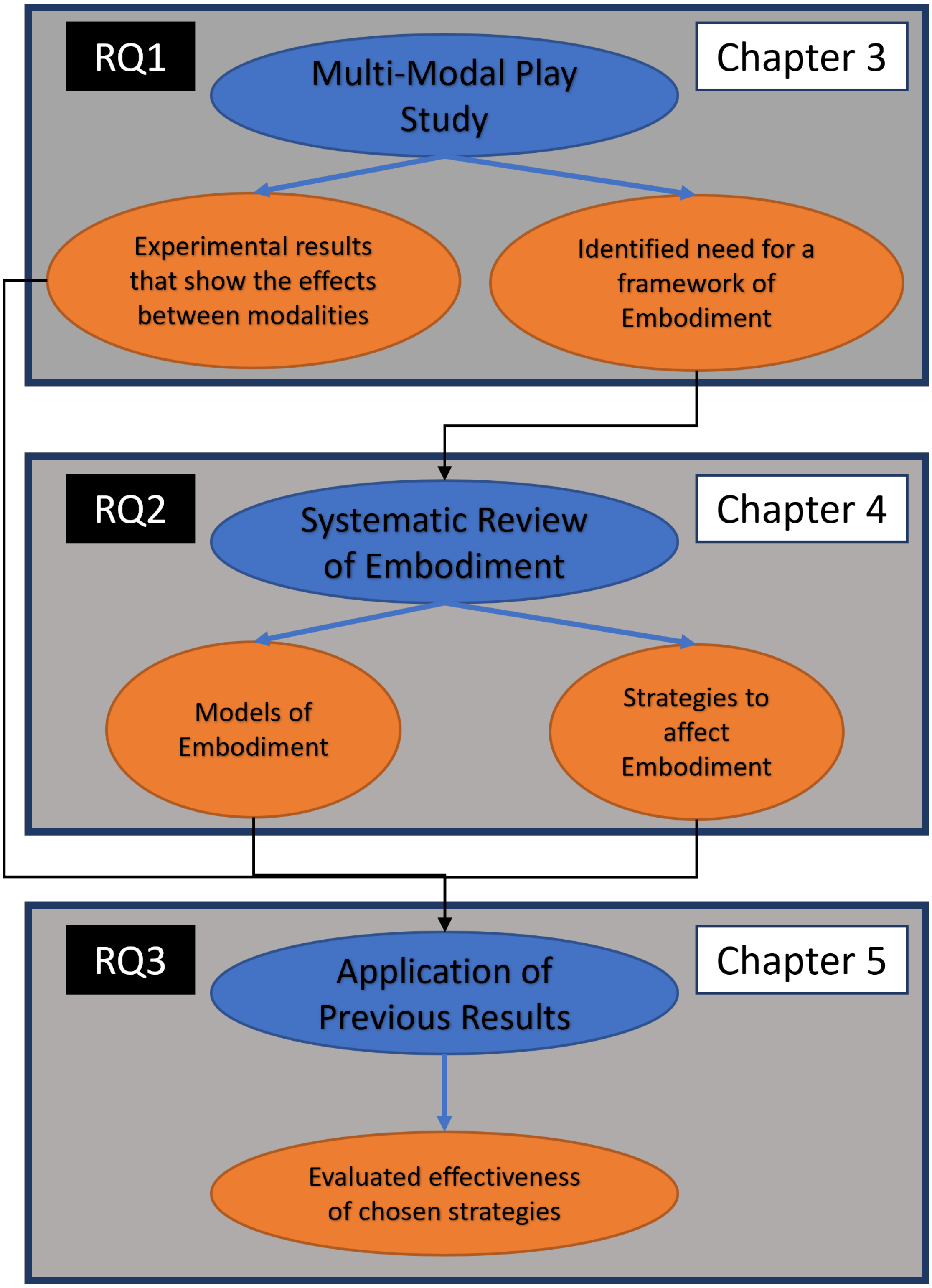

FIGURE 1.1: A chart showing the overall methodology/research framework (divided into three logical parts, each answering one of the research questions. In blue are the processes (experiments), and in orange are the outcomes from each experiment)

This was part of the Doctoral Consortium of CHI PLAY in 2020, in which a short summary of the work that was done and the future direction of my PhD was discussed, as well as a short 5 minute video that was presented at the online conference.

- Malinov, Y.-D., Millard, D. E., and Blount, T. (2021). *StuckInSpace: Exploring the Difference Between Two Different Mediums of Play in a Multi-Modal Virtual Reality Game.* in 2021 IEEE Virtual Reality and 3D User Interfaces (VR), pages 501–510

  This paper contains the first experiment results and analysis that answer **RQ1.** (discussed in Chapter 3).

1.6 Contributions

This thesis has made the following contributions:

1. The design and implementation of *Stuck in Space*, a multi-modal co-located multiplayer VR experience, using technology that is normally available to an average consumer;
2. A user study examining the differences in immersion and co-presence when the HMD player is interacting with either the PC or Phone players;
3. A framework of embodiment and the accompanying strategies used to achieve it, created after a systematic literature review;
4. An example of how the framework of embodiment can be used as a guide in the analysis and design of multi-modal experiences;
5. A subsequent user study using the framework of embodiment created to compare the experience when an embodiment strategy is present or absent, and the analysis of the results;
6. A new strategy for affecting embodiment in a multi-modal experience.

1.7 Thesis Structure

The background of this study is discussed in Chapter 2, where VR, multi-modality and the different experiential factors (co-presence, immersion, and embodiment) are examined more in-depth. Chapter 3 contains the first experiment, in which a game was created for the purposes of testing how adding a second player affects the experience. The second experiment, a systematic review, is in Chapter 4, where the resulting embodiment models and strategies are discussed. Chapter 5 is about the third experimental study, in which the results from Chapter 4 were applied to the game from Chapter 3 and explore the effectiveness of those strategies in improving the embodiment in an existing multi-modal game. The final chapter, Chapter 6, summarises the results from the three experiments, addresses the answers to the three research questions, and gives a comprehensive overview of the results of the thesis, future work, limitations, and concludes with final thoughts and recommendations about the future of multi-modal experiences.

Chapter 2

Background

2.1 Introduction to Virtual Reality (VR)

The dry definition of Virtual Reality is *"computer-generated simulation of a lifelike environment that can be interacted with in a seemingly real or physical way by a person, esp. by means of responsive hardware such as a visor with screen or gloves with sensors"*, as taken from the Oxford Dictionary [1]. In simpler terms, it's the lifelike simulation and interaction of a human with a 3D environment, normally done through a Head Mounted Display (HMD) and a set of motion-tracked controllers.

The idea of "virtual reality" itself is quite old, tracing its roots back to the 19th century and the creation of the stereoscope, which used two images side by side to create a sense of depth in the user viewing them [13]. People imagined a device that could transport you to other worlds and places, letting you experience things out of this world — science fiction author Stanley Weinbaum played with that idea in his story "Pygmalion's Spectacles" [91] back in the 1930s, a story that is frighteningly close to what we call VR today. The term "virtual reality" itself was popularised by Jaron Lanier in the 1980s during his pioneering work on early VR systems [4]. Throughout the later part of the 20th century, the path to modern VR was paved by scientists and inventors [53] who slowly built and refined the design, building into what we have today as modern HMD headsets and VR displays. Although early iterations of this technology were mostly for non-commercial use and people couldn't easily get their hands on them, there were a few unsuccessful products, such as the Nintendo Virtual Boy, which worked more like the early stereoscopes rather than actually creating an immersive experience, mostly due to hardware limitations.

This technology really became what we know today in the past decade when the first Oculus Rift prototype, DK1, was launched in 2013, followed by the consumer version,

[1] Oxford Dictionary - Virtual Reality Definition

Oculus Rift CV1, officially launched in 2016. These releases marked a significant step, putting VR in the spotlight, making it possible for people to actually be able to experience it outside of an expensive conference or inside an academic lab. Thanks to more powerful Graphics Processing Units (GPUs) that can render two images at once at high enough frame rates, displays increasing in resolution and fidelity, and better hardware that is able to track small movements and translate them to a virtual environment, VR has been propelled into the mainstream, captivating enthusiasts and developers alike.

2.2 Mixed Reality (MR) as a concept

If VR is a window into the virtual world, isolating you and immersing you within it, then Mixed Reality (MR) is the seamless fusion of the real and virtual worlds, or as per the Oxford Dictionary [2], *"A medium consisting of immersive computer-generated environments in which elements of a physical and virtual environment are combined"*. In this case, MR lets people blend a virtual environment and the real world, and aren't potentially as isolated as in a VR environment.

2.2.1 The Reality-Virtuality Continuum and Milgram et al.'s Taxonomy

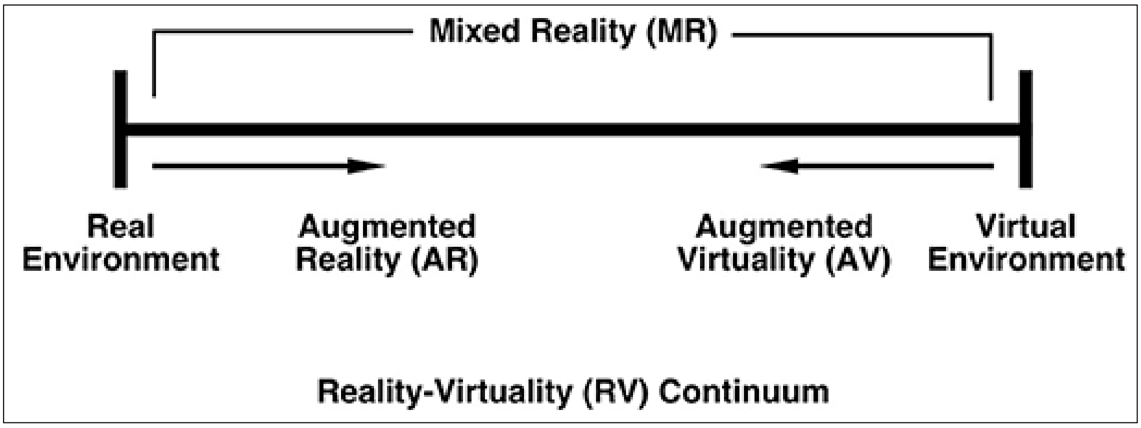

FIGURE 2.1: The Reality-Virtuality Continuum by Milgram et al. [57].

The "Reality-Virtuality continuum" was introduced by Milgram et al. [57] in 1994. This continuum allows varying levels of immersion and interactivity and covers the entire spectrum of Mixed Reality:

- **Real Environment**: The starting point on the continuum is the actual physical world.

[2] Oxford Dictionary - Mixed Reality Definition

- **Augmented Reality (AR)**: AR overlays virtual objects onto the real environment.
- **Augmented Virtuality**: Here, real-world objects are brought into the virtual space.
- **Virtual Reality (VR)**: The far end of the spectrum, where users are fully immersed in a virtual environment.

In addition to the continuum, Milgram et al. [57] proposed a three-dimensional taxonomy to classify mixed reality environments. The three dimensions are:

- **Extent of World Knowledge (EWK)**: This dimension assesses how much the system knows about the real world and its representation within the virtual environment. At one end of the spectrum is the completely unmodelled world, where nothing is known about the environment being displayed. At the other end is the completely modelled world, where the system has complete knowledge about every object and its location.
- **Reproduction Fidelity (RF)**: This dimension evaluates the quality and realism of the virtual elements being integrated into the real world. It encompasses various factors such as display hardware, signal processing, and graphic rendering techniques, all contributing to the fidelity with which the system reproduces both real and virtual objects.
- **Extent of Presence Metaphor (EPM)**: This dimension considers how immersive the experience is, reflecting the degree to which users feel present in the virtual or mixed environment. It ranges from exocentric displays, where the user looks into the virtual world from an external viewpoint, to egocentric displays, where the user feels immersed within the virtual environment.

By this definition, VR is indeed part of the Mixed Reality spectrum, even though it might not necessarily have elements of the physical world in the virtual one.

In this thesis, the term “mixed reality” will be used as a descriptive term for the combination of real and virtual as per Milgram et al., because the author views MR as more of a descriptive term rather than a set of certain technologies. The ‘mixing’ of both worlds is what is important, not the way it’s done. For simplicity, when MR is mentioned, VR is being excluded.

2.2.2 Benford et al.’s Taxonomy

Another way to classify Mixed Reality experiences is with the taxonomy and framework that Benford et al. [3] provide in their paper. Three dimensions are used to classify shared spaces:

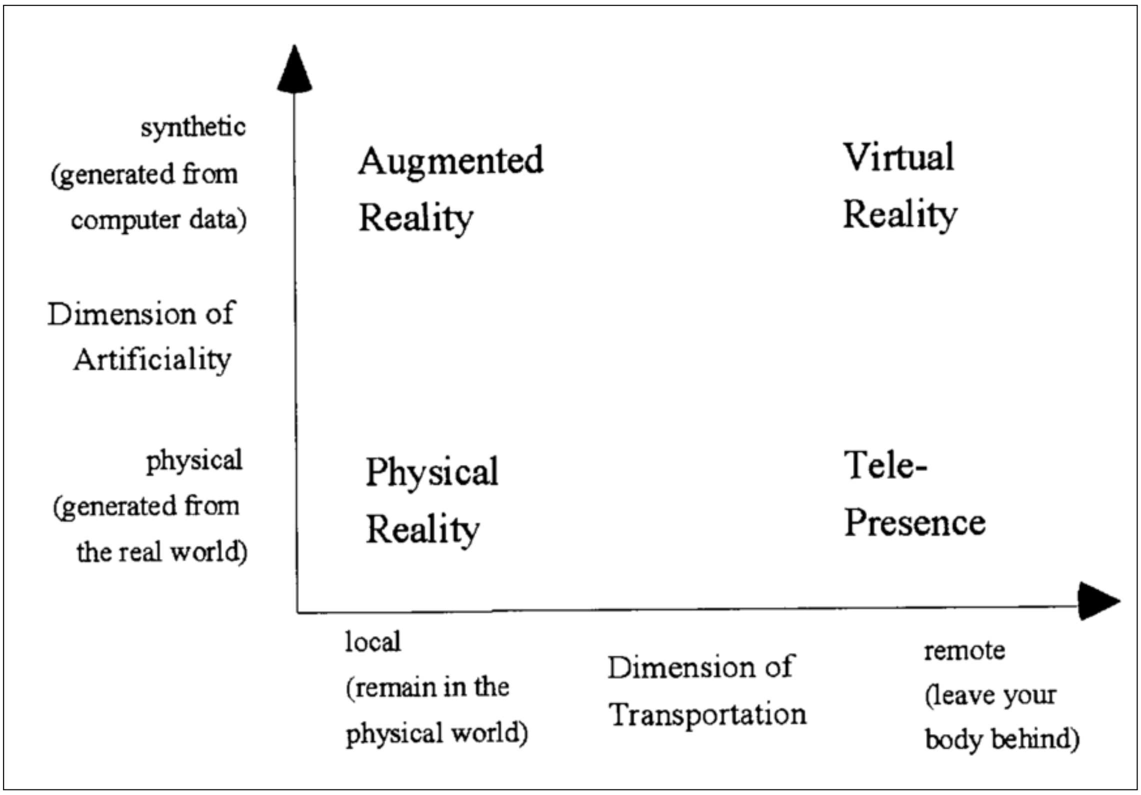

FIGURE 2.2: Two of the three dimensions of the Taxonomy by Benford et al. [3].

- **Transportation** concerns the degree to which users are transported into a new space or remain in their local space. This includes bringing elements from remote spaces into the local environment or vice versa.
- **Artificiality** addresses the degree to which the shared space is based on real-world information versus synthesized data. It reflects the balance between physical and virtual elements within the space.
- **Spatiality** involves the extent to which the shared space exhibits key spatial properties such as containment, topology, movement, and a shared frame of reference. This highlights how well the system understands the positions and locations of objects and participants within the space.

This taxonomy diverges from Milgram et al.'s taxonomy, which focuses on the extent of world knowledge, reproduction fidelity, and extent-of-presence metaphor. Benford et al.'s approach incorporates non-Cartesian spaces and emphasizes the integration of multiple distinct locations, such as topological spaces created by mediaspaces, contrasting with Milgram et al.'s focus on Cartesian spaces and augmented reality technologies.

Benford et al. illustrate the practical application of this taxonomy through the staging of a mixed-reality poetry performance. This performance involved both physical and virtual audience members interacting simultaneously, showcasing the complexities and opportunities of mixed-reality boundaries in creating shared spaces.

For mixed reality systems to work, in almost all cases, there needs to be a way of fusing the physical and virtual world, be it by the use of cameras to capture reality, or any type of other sensors that can "translate" between the two realities. There is the Microsoft "HoloLens", a mixed reality headset that combines the two worlds using "holograms". It uses a plethora of sensors to precisely track the headset position, any gestures the wearer does, as well as any voice commands that might be said. It can "ground" virtual objects in the real world, making it very immersive.

The versatile applications of MR can help reshape a multitude of sectors. In their systematic review of smart glasses and their applications, Kim and Choi show that this technology is being used and researched in many sectors such as healthcare, where it can be used for surgical practice or better visualisation of patient data [96, 71, 70]. The industrial sector benefits from MR's assistance to aid maintenance and work support [95, 44]. Similarly, in education, MR proves invaluable by helping to render complex concepts tangible, helping students understand Kirchhoff's laws [40] or conduction of heat [85]. All of these are just a few examples of the different applications of this technology in different sectors.

A keen observer might have noticed that this thesis is about Multi-Modal play, and not Mixed Reality, and might be confused about the terms. While MR is all about the blend of virtual and real, multi-modality can mean a few different things based on the context. The next section will go more in-depth about what multi-modality is in the context of a VR or MR experience.

2.3 Multi-Modality and its difference to Mixed Reality

Multi-modality is a term that can have various meanings based on context, and thus it is important to define how it is used in this thesis. In the thesis the term is used as a way to talk about the different modes of people participating in an MR experience. As an example, consider a scenario where friends gather to play games together in person: one person might be wearing a VR headset, another could be participating with their phone, and a third one might be playing on a laptop or a PC. In that case, there are three different modalities: VR, Phone, and Desktop.

In and on itself, those experiences are multi-modal, but may not necessarily qualify as mixed reality. For instance, if the VR player is completely isolated from the external world and is interacting with the others solely through the virtual world and their avatars, versus hearing the voices of the other players in the virtual world (in other words, taking the physical and transforming it into digital).

Another different definition that is used for multi-modality within MR and VR literature is the mode of interaction with the system, the interface. One can have visual, auditory, haptic, or gestural modes, and multi-modality comes into play to describe the combination of that multitude of interaction modes [2]. This definition is not used in the thesis, as it's not helpful for the work being undertaken. This research focuses on the interactions between different systems, or modes, rather than the interaction within one single system/mode.

2.4 Collaborative Experiences in Multi-Modal Spaces

Having defined multi-modality as it will be used in this thesis, the discussion can now turn to the work that has been done in general around multi-modal spaces. With multiple participants in a multi-modal setting, one can start thinking about collaboration and interaction between people in such spaces.

There's been a lot of research into this sphere, where the main focus would be collaboration on a higher level through technology. Here, the word 'groupware' is used, meaning a system that aims to assist *'communication, collaboration and coordination of a group of people'* [24]. One example of such technology is *Shared Space* [5], which is a Computer Supported Collaborative Work system where multiple participants use augmented reality to cooperate in a co-located three-dimensional space. What they found is that seeing the other person's physical body and body cues helped lower the time to complete a set of tasks when compared to a version where you only see the virtual body [5].

Collaborative Virtual Environments (CVE) are, on the other hand, completely virtual spaces where multiple users can collaborate. Goebbels and Lalioti give a number of guidelines for the use of different mediums (video, audio) in CVEs, such as designing in such a way as to decrease the mental load on each of the users, which leads to greater perception of co-presence and subsequently, greater perception of collaboration [28].

2.4.1 Co-located Asymmetric Experiences

Going back to VR, a person is very isolated from the outside world by design, wearing a headset that replaces their whole field of view with a computer-generated 3D environment. This has been something researchers have noticed and have tried to mitigate, making the experience more sociable by including non-HMD users in different ways [35, 36, 20, 98, 51]. The next paragraphs will go through some of the notable examples of such technologies.

FaceDisplay [35] do it through a number of touch displays attached to the HMD that show what the person with it sees inside, and lets bystanders interact with the HMD user through a number of Leap Motion sensors that detect hand gestures. They have created 4 different levels of interaction - just observing, through an external device (such as a smartphone), through hand gestures, and through touch. They found out that their method was able to introduce a non-HMD player and have them interact with the HMD player, although "*the former had a higher level of dominance and responsibility over the latter*" [35]. A few other papers (*FrontFace* [17] and *See What I See* [67]) also show a number of prototypes with a display attached to the front of the HMD (both papers use mobile phone VR) to help with social interaction between the HMD user and observers.

The authors of *Astaire* [98] on the other hand take another approach, where the non-HMD user has to "dance" with the HMD user in order to play the game. They have to hit notes as they come, which encourages them to be close to each other. The non-HMD player watches a screen and has the ability to see where a note will be one second in the future (but only on a two-dimensional grid), while the HMD player can see the note's position more precisely in three dimensions. A thing to note is that they do not concentrate on just the HMD player (like *ShareVR* [36]) or the non-HMD player (like *TurkDeck* [20] and *HapticTurk* [19]), and they try to make a "*well-balanced asymmetrical play experience*" [98].

CatEscape [51] is an interesting multi-modal game where there can be a maximum of 3 players in different modes: one in VR (that plays a survivor game), one that uses a tablet (a strategy game), and one that uses AR (where they play an explorer game). The AR version here is interesting as they have chosen to use a mini-projector attached to the Vive controller to use it as a "torch" and see the shadows of the VR world. One of the suggested applications is for a wider audience to be able to see "into" the virtual world without having to look into a stationary monitor, and being able to more easily interact with the players.

In the work of Gugenheimer et al., they have created *ShareVR* [36], a prototype system that uses projections onto the floor, as well as portable tracked displays to help the non-HMD player communicate and interact more easily with the HMD player. Concerning the portable display, they talk about using it as a 'window into the other world', being able to move around and peek into it easily. Before creating the prototype of *ShareVR*, they did a preliminary study to gauge the interest in such a technology, and they show that people indeed express a demand to interact with the world the HMD player is in while not necessarily using a headset themselves.

Cheng et al. tried including a second player through a different approach by using them as human actuators that would aid the HMD wearer's experience by moving physical props (*TurkDeck*) [20] or manually supporting the wearer above the ground playing a flight simulation (*Haptic Turk*) [19]. What they show is that even though the 'human actuators', as they call them, are not necessarily playing the game, they do still feel enjoyment from such a game, aiding the VR user to have a better experience and watching them have fun are main driving factors.

Lee et al. have created *RoleVR*, a multi-modal co-located system, and they have explored the difference between the different asymmetric environments [50]. They show that different kinds of interaction can increase the immersion for the players, and that depending on what you want to achieve, you might want to use one kind of interaction in place of another.

Born et al. on the other hand have researched how adding a second player either in a co-located manner, or through a remote setting, affects the performance of both participants [9]. They show that having the players close to each other actually lowers it, one of the reasons being that players have to try and avoid the other, thus limiting their ability to focus on the task, which is what Podkosova and Kaufmann also find in their paper [66]. Gómez and Fons come to the conclusion that wearing the HMD prevents social presence to increase as expected when being co-located [30]. Finding ways to mitigate that effect is something important for future research, as well as something this thesis will try to answer.

As this is a relatively new field, research for it comes out every year. As noted above, there's a lot of work being done on introducing a second player, either co-located or remote [9, 50, 35, 98], and through different means, like using different props [36, 20] or project what the HMD user is observing for others to look at [39]. Despite this progress, there are still many unexplored areas within this field. This research aims to contribute to these underexplored areas, particularly as co-located experiences become more widespread. Understanding how to design a multi-player co-located game with the goal of increasing immersion, for example, is a vital part of the future of this field. As Gugenheimer et al. say, the future of the "social living room environment" [36], Additionally, it is essential to find ways to minimise negatives associated with introducing a second player, as highlighted by Born et al. and Podkosova and Kaufmann.

With the ongoing research and the increasing relevance of this field, it is important to determine what exactly should be measured in a co-located multi-modal experiment. The next section discusses the factors chosen for measurement and the rationale behind these choices.

2.5 Experiential factors for Virtual Reality

As there are different modes that a person can be in when interacting in a multi-modal setting, different variables can affect the experience in distinct ways. For example, immersion is a very big part of a VR experience, as your vision is completely taken by the HMD, while if you're using a phone to interact, it takes a much smaller part of your vision and is less immersive. The opposite could be true, where co-presence might be stronger when using a phone as you can feel the other person's presence and see them, while in VR you are isolated from that and need a lot more cues in order to potentially feel the same level of co-presence. For this reason, this thesis explores three different factors that a multi-modal experience can have: immersion, embodiment, and co-presence.

Immersion was chosen as a property to explore because the main point of Virtual Reality is to immerse the user (a big part of research in immersion is specifically about Virtual Environments and Virtual Reality [68, 47, 75]), so if adding a second player into the multi-modal experience lowers or increases immersion, then that would be very important to know when creating such experiences in the future. This applies to both the immersion of the HMD player and the non-HMD player.

Embodiment, in contrast, is a subjective response of the user to their presence in the virtual environment. As Kilteni et al. discuss, it includes three main components: self-location (feeling of being located in the virtual body), agency (feeling of being in control of one's own virtual body), and body ownership (feeling of the virtual body belonging to oneself) [42]. It was chosen as a property because it's closely linked with immersion, being a vital part of your experience, VR or not, and affecting it in different ways has the potential to affect immersion as well.

Although immersion and embodiment are distinct, they are sometimes conflated in the literature. For example, Evans and Rzeszewski discuss the differences between immersion, embodiment and presence, and how they are sometimes treated as interchangeable terms, leading to confusion about their roles in the user experience [25].

Co-presence was chosen as a property to measure as the main goal of the thesis is exploring how adding a second player through a Phone/PC changes the experience in a multi-modal setting, and as this means it's multi-user, co-presence would be measurable and there is a possible difference between the two ways (both for the HMD and non-HMD user). Co-presence was also shown to increase enjoyment, as Gajadhar et al. show in their paper [27], which is something that an experience would want to increase, be it a game or a collaborative work system for example.

2.5.1 Immersion and Presence

The terms immersion and presence are very closely related to one another and thus require clear definitions. There are several existing definitions: for example, Slater defines immersion as a measurable property of the system [78], like field of view, screen size, etc., while Witmer and Singer define immersion as the way the person responds to the system [94]. In this thesis, the definitions of Slater for presence and immersion will be used.

Historically, the term presence came from what Minsky in 1980 called "telepresence", describing a sense of being present in a remote location through "teleoperation" technologies. In the following years this concept evolved and researchers like Jonathan Steuer and Mel Slater applied it to virtual environments. Steuer defined it as the "sense of being in an environment" generated by virtual reality systems [84], while Slater distinguished it as a psychological state of feeling "there" in a virtual environment [78].

Slater further refined the concept of presence into two key components: **Place Illusion (PI)** and **Plausibility Illusion (Psi)** [79]:

2.5.1.1 Place Illusion

Place Illusion (PI) refers to the experience of "being there" in the virtual environment. This concept aligns with the definition of "presence" as proposed by Slater [78], specifically regarding the strong illusion of being in a place, despite cognitively acknowledging one's physical absence [79]. This also has close connection to the "self-location" embodiment concept explored in the next sub-sections.

2.5.1.2 Plausibility Illusion

Plausibility Illusion (Psi) is the illusion that a scenario is occurring [79]. A central aspect of Plausibility Illusion involves the virtual environment's ability to generate events that conform to the participant's expectations of real-world behaviour [79].

Together, PI and Psi form the foundation of presence. Slater [79] argues that when both PI and Psi are achieved, the participants respond more favourably and realistically to the virtual simulations.

2.5.1.3 Break in Presence

Slater introduced the concept of “Breaks in Presence” (BIP) to describe disruptions to PI and PSi [79]. BIP happens when inconsistencies between the physical and virtual environments remind participants of the real world. For example, bumping into objects can pull participants out of the experience. Slater [79] talks about the difference between breaks in presence for PI and PSi. PI can exist in one modality (such as visual) and be absent in another (such as auditory), and thus breaks in one modality don’t necessarily transfer to another [79]. In contrast, breaks in PSi are less recoverable: it’s much harder for a participant to go back into believing a “scenario” once the realism has been broken [79]. An example of that would be if a virtual character acts oblivious to you while you are talking to them, or does something unexpected.

2.5.1.4 Narrative Embodiment

In addition to Place Illusion (PI) and Plausibility Illusion (Psi), this thesis introduces the concept of “narrative embodiment”, which refers to the participant’s immersive experience within a narrative-driven virtual environment. While closely related to Psi, “narrative embodiment” specifically focuses on the players participation into the narrative in an active role, influencing and being influenced by the unfolding story. This concept will be explored in depth later in this thesis, building upon Slater’s models of plausibility [79] and extending them to encompass narrative engagement.

Factors that impact immersion have been extensively researched in the work of Schuemie et al. [75]: for example, one of their factors is *user characteristics*, where they talk about how the type of person, their ability to focus and their experience with games affects immersion [75]. Steuer [84] mentions that suspension of disbelief plays a role in immersion [84], and Bangay and Preston find that control plays a major role in immersion, where a lack of any control makes a person less immersed [1]. It is also connected with collaboration, as Narayan et al. [60] show - increasing immersion helps with the performance of the users on a collaborative task [60].

2.5.2 Embodiment

Embodiment and being embodied in the literature have been defined in a number of different ways. De Vignemont [22] defines an object E being embodied “if and only if some properties of E are processed in the same way as the properties of one’s body” [22]. Kilteni et al. [42] also agree on that definition. That definition is vague in some sense, so Kilteni et al. [42] reviewed the literature and discovered that *embodiment* was associated with three main concepts: the sense of *self-location*, of *agency*, and of *body ownership*.

2.5.2.1 The Sense of Self-Location

Self-location is the feeling of being physically located in a certain place [S25]. There is a distinction between being self-located, which refers to one's perception of being in a body in space, and presence, which refers to the experience of being in an environment and interacting with it. Place Illusion (PI) plays a crucial role in this feeling of self-location, as it covers the sensory parts of what makes the virtual environment feel like a real place. Studies show how visual perception and perspective (first/third person) significantly influence this feeling [8, 81, 64].

2.5.2.2 The Sense of Agency

Agency refers to the feeling of being in control of one's own body, and subsequently being able to influence the environment around them. Blanke and Metzinger [8] define it as "global motor control, including the subjective experience of action, control, intention, motor selection, and the conscious experience of will" [8]. An example of agency in VR would be the ability to pick up and throw objects. Plausibility Illusion (Psi) significantly impacts the sense of agency, as the credibility of interactions and the alignment between user expectations and actual events affect how in control the user feels. Any difference between expected and actual outcomes can minimise the sense of agency, as shown by various studies [26, 7, 73].

2.5.2.3 The Sense of Body Ownership

Body ownership is described as "one's self-attribution of a body" [S25]. A famous example is the rubber hand illusion, where stimulation of a hidden real hand and a visible rubber hand at the same time can lead individuals to feel that the rubber hand is part of their body [11].

2.5.3 Co-presence

The term co-presence, as defined by Goffman and Slater et al., refers to the perception of the Other together with you, and the feeling of them perceiving you, which are two distinct elements [29, 82]. In the case of co-located multi-modal VR, co-presence would be how the HMD user perceives and feels the non-HMD user(s) around them, and vice-versa.

When talking about co-presence, the term "social presence" is sometimes used. Sallnäs explains how this is different from co-presence, defining social presence as based more on how the medium affects the user's perception, whereas co-presence concentrates more on the possible interactions between people [72]. Biocca et al. define social presence as the *"awareness of the co-presence of another person together with a sense of engagement with them"* [6], indicating that co-presence is part of social presence.

Zhao further breaks this up into four groups based on the factors "proximity" and "representation" of the other [97]. In this view "proximity" can be either physical or electronic, where one person could be close to you in physical space (in the same room) or close in the electronic space (in a teleconference where instant two-way communication can happen). "Representation" on the other hand is how the other person is simulated, either through a physical or digital simulation — a robot/virtual agent located in the same space (physical or virtual) enables the person to interact with the other.

Outside of games, co-presence has been explored extensively in the field of Computer Supported Collaborative Work (CSCW) and Collaborative Virtual Environments (CVEs) - in the work of Kraut et al. the authors show that an increase of co-presence through a shared visual space is beneficial to task performance in terms of speed and quality [46]. Another example is the aforementioned research of Goebbels and Lalioti, where they provide some guidelines for such experiences. This can be applied in fields outside of gaming, such as medical school, where a teacher and students could be in one CVE to facilitate learning, or collaboration in a teleconference, where you want to increase the sense of co-presence to increase the benefits of it.

2.6 Summary

This chapter has covered what virtual reality is, from inception to modernity, and how mixed reality is related to it by combining the real and virtual. The distinction between the terms "mixed reality" and "multi-modality" was clarified, with "multi-modality" defined in the thesis as the interaction between different modes of play (be it PC, VR, or Phone).

This led to the exploration of collaborative experiences in multi-modal spaces, providing examples of what that entails, and laying the foundation for the further exploration of co-located asymmetric experiences in the following section. Several recent examples of such games were reviewed, along with their findings.

To be able to analyse a multi-modal experience, one must also decide on variables to capture, and for this, three primary ones were chosen: Immersion, Embodiment, and Co-Presence. Each was explained in terms of importance for the overall user experience. Immersion, being a fundamental characteristic of VR, stands as a very important point due to its desirability. Embodiment, closely linked to and capable of affecting immersion, is intrinsically tied to the human experience, as we all are "embodied" in our own bodies. Lastly, co-presence is important the thesis as it is exploring co-located experiences with multiple participants, and being able to measure how each mode is being affected is pivotal to understanding the experience.

Building on top of this background research and understanding of multi-modal experiences, the following chapter is going to introduce the primary game created to test the thesis: a co-located multi-modal VR game.

Chapter 3

Studying Immersion and Co-Presence in a multi-modal VR game

In the first experiment, two different ways of involving a second player in the VR game were explored to test how they affect the Immersion and Co-presence of the participants. An important thing that was considered while choosing the two different ways was the *accessibility* to the average consumer: while some papers use projectors [36, 39], which is not something that a person has readily available, or a big open space with people propping and changing the physical scenery [20], a phone was chosen as a way to act like a "window" into the virtual world. This is because phones are widely used and almost everyone has access to one. The other chosen version was a desktop one, which uses the machine currently running the game. This was done to evaluate whether multi-modal VR experiences can be easily brought into the living room (as Gugenheimer et al. [36] found in their online survey, there is a demand for co-located experiences in VR).

To this end, *StuckInSpace* was developed as an asymmetric co-located multi-modal VR game where two players cooperate to complete various challenges. The player with the HMD is an astronaut whose ship is malfunctioning and needs the help of the second player, who plays as a drone with access to special information on how to fix the problems.

3.1 Design and Implementation

The game was explicitly made to be asymmetric, as *ShareVR* does [36], to increase the engagement and inclusion of the non-HMD user.

This is based on the findings of Cheng et al. [19], who showed engaging the second user even just by having them help the HMD player increases their enjoyment [19]. Therefore, a cooperative scenario was chosen for the game, featuring an astronaut who needs to repair their space station with the assistance of a drone.

The astronaut is played by the HMD user, who can move around in a room-scale VR environment and interact with the buttons and world. The drone is played by the second player, either through a PC or a Phone. Both versions have a first-person view of the drone as they control it and move around the space station, either by physically moving (in the case of the phone player), or by using the mouse and keyboard (for the PC player). The drone itself has access to special information that is not visible to the astronaut, such as instructions, codes, or hints.

Looking at the taxonomy by Benford et al. [3], one could put the three different modes of play in as seen in Figure 3.1:

- StuckInSpace (VR) — this one uses the most synthetic artificiality, as it completely uses computer-generated graphics, and is in the middle of the transportation axis, as there is a second player close to them, reminding them of the real world;
- StuckInSpace (PC) — here, there is less transportation than VR, as the computer screen isn't as immersive as an HMD, yet the two players are in the same room. The artificiality is still high, as it's using completely computer-generated graphics as well;
- StuckInSpace (Phone) — for the phone version, there is less artificiality based on the fact that the screen is so small and most of the real world remains in sight, even though it's still using computer-generated graphics. The transportation is still partial, but thanks to the fact the player is next to the VR player, it's more towards the local side of the axis.

The design of the multi-modal game was inspired by games like *Keep Talking and Nobody Explodes* [83], especially in how they present asymmetry of information to players. They have to communicate verbally without seeing each other's screens, creating a high-pressure situation where both players have to rely on each other's skills and trust. A similar dynamic was intended for *Stuck in Space*, where both players have different perspectives and information but have to work together towards a common goal. The game was divided into two segments: the tutorial and the main game. The tutorial is needed to help both players get used to the controls and understand what to do when the game starts. There are seven actions that players are prompted to take, as shown in Table 3.1.

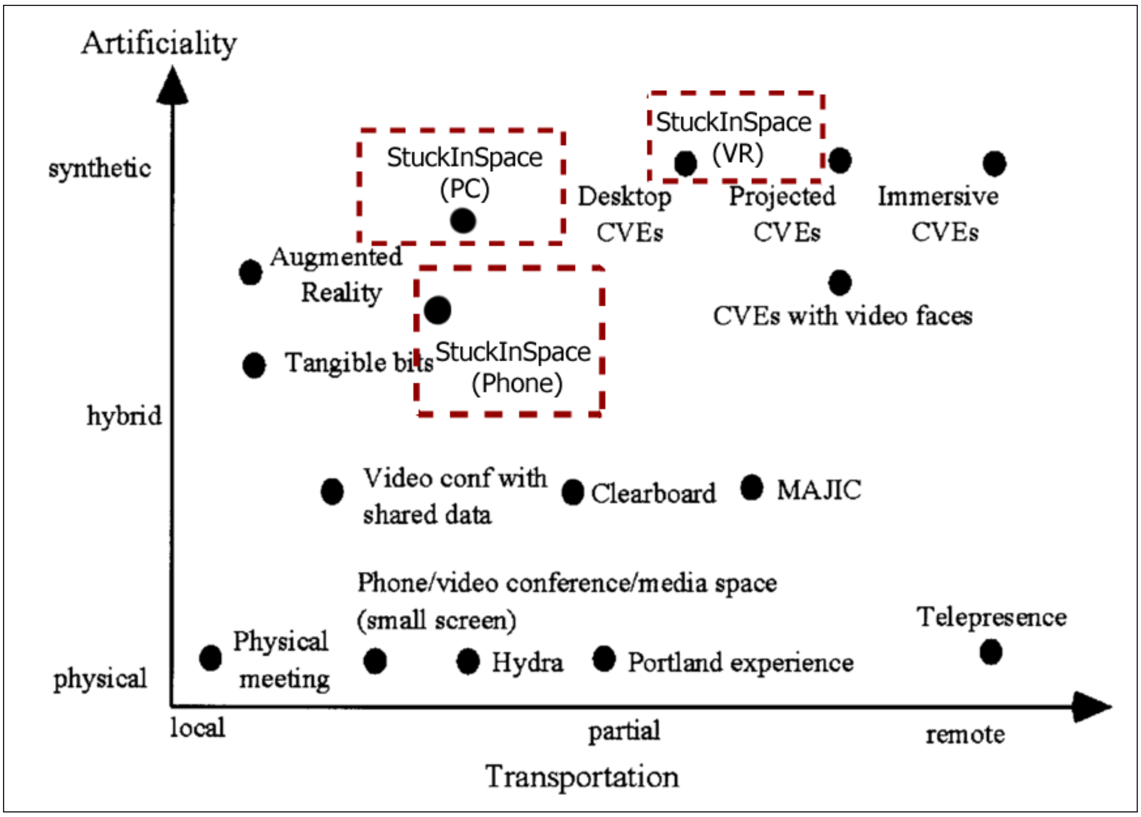

FIGURE 3.1: The three different modes of play of StuckInSpace in the taxonomy of Benford et al. [3].

- 'Button wall' refers to a section of the play-area that has twelve red buttons which must be pressed in a certain order (Figure 3.2). Only the non-HMD user can see the order in which the buttons must be pressed, meaning that the two players have to communicate to complete it. The number of buttons to be pressed in order to continue steadily increases to keep up with players' skill level;
- 'Pressure button' (Figure 3.6) is a button that only the HMD user has to press in order for a constantly rising pressure gauge to remain in the 'green zone'. When it gets too high an alarm starts sounding (only to the HMD player) so as to have them remember to press the button. In later stages the speed of it increases in order to enact a sense of urgency - as the alarm is only heard by the HMD player, the non-HMD player can keep track of it and warn the HMD player before it gets too much. It should be noted that this does not lead to any game over state if left on its own, it only forces the HMD player to move around the room more to silence the alarm;
- The 'keypad' and 'specialised buttons' (Figure 3.6) are another way of making the information available asymmetric, as only the non-HMD user can see the keypad code/the button that needs pressing in order to continue (a 4-digit

TABLE 3.1: List of actions, showing the player that does it and information exchange between the players if any.

Action	Enacted by	Information flow
1. Button wall	HMD	Other to HMD
2. Pressure button	HMD	N/A
3. Keypad	HMD	Other to HMD
4. Specialised buttons	HMD	Other to HMD
5. Light panel	HMD	N/A
6. Flashlight	Other player	N/A
7. Fix drone camera	HMD	Both ways

code/button is randomly generated and has to be communicated to the HMD player);

- The second player has a 'flashlight' on their disposal, which they can use to guide the HMD user whenever the lights go out in the game, as well as using it to aid communication;
- 'Fix drone camera' is needed when the screen on the non-HMD user's device becomes gray-scale and they are unable to tell the order of the buttons needed (see Figure 3.3). Thus, they have to communicate to the HMD user that they need to be fixed to continue the game. This is done by the HMD player getting close to the "drone" and pressing a button on their controller.

There are nine stages, with a set type and number of actions, with the solutions to these actions being randomly generated. This ensures enough difference between play sessions that players cannot just repeat their actions, while ensuring that every player experiences each type of puzzle. Players are not penalised for failing tasks, as the focus is on their interactions rather than their performance. For example, pressing an incorrect button on the 'wall' task does not lead to a failure state; instead, the player can simply try again.

To differentiate between the two possible modalities in the study, the designations HMD^{Phone} and HMD^{PC} are used. If no platform is specified, it is applicable to both groups of HMD users.

When looking at the possible types of co-presence in the system, one can see how Zhao [97] has divided these types based on proximity between the two people and whether they are present in the physical space (in person) or simulated virtually (or a mix of these) [97]. The co-presence types of the game fit into this taxonomy:

- The first type is what the HMD^{Phone} player feels when they play with the Phone user — the other is in physical proximity (Figure 3.5), but using a virtual representation of themselves (in the form of a drone) in VR;

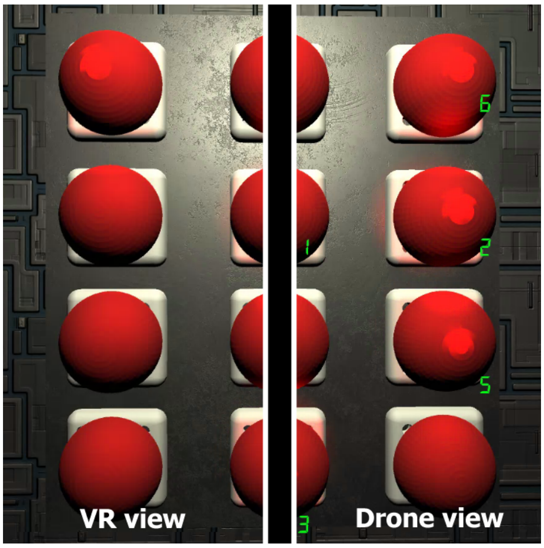

FIGURE 3.2: A comparison between what the HMD player is seeing (on the left) versus what the drone player is seeing (on the right)

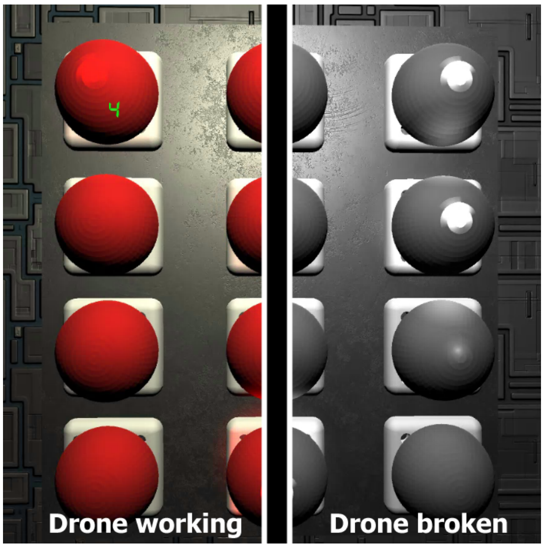

FIGURE 3.3: A comparison between what the drone sees when it's working (on the left) versus when it's broken (on the right)

- The second type is when the HMD^{PC} user plays with the PC user — in this case, they still see the same virtual representation, but this time they do not feel the physical proximity to the same degree, as the PC user is sitting away from them (Figure 3.5);
- The third type is with the Phone player — they see the HMD^{Phone} user both as a virtual representation in the game as well as seeing their physical form, and they feel the other in physical proximity;
- The last type, using the PC, proximity is less than the Phone, and the non-HMD user in that case does not see the HMD^{PC} player as much because their attention is on the screen and they have their back facing the HMD^{PC} player (although they have the ability to do so).

3.1.1 Hardware Implementation

The game was built using an HTC Vive as the base HMD (although compatibility across platforms should be easily achievable). The headset itself uses a lighthouse based tracking system: two sensors on opposing corners of the play space point at the headset and controllers to enable tracking, which provide sub-millimetre accuracy and low latency for tracking.

A way of tracking the phone screen had to be developed as to make it act as a "portal to the other world". Ultimately, using one of the controllers was decided, just as Gugenheimer et al. did, to enable tracking of the phone in the play space (Figure 3.4).

FIGURE 3.4: Phone player uses a phone in one of their hands and a controller to track their position in the other (no need for any holder as we wanted this to be as accessible as possible)

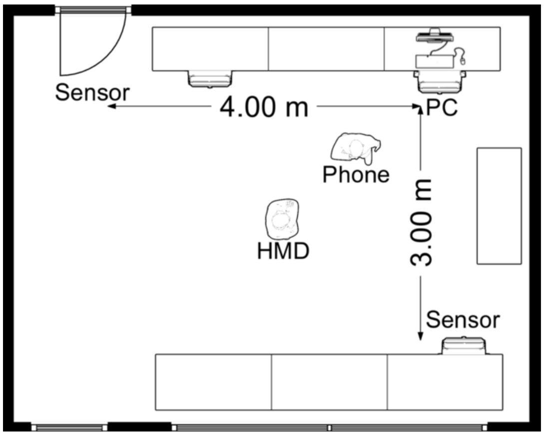

FIGURE 3.5: Floor plan of room used for the experiment that shows two of the users (for HMD-Phone) and the position of the PC (for HMD-PC)

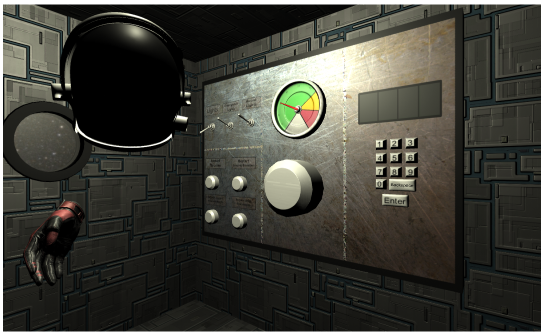

FIGURE 3.6: The view from the drone's perspective — the astronaut head can be seen (which is directly mapped to the headset position), as well as the hand which the astronaut uses

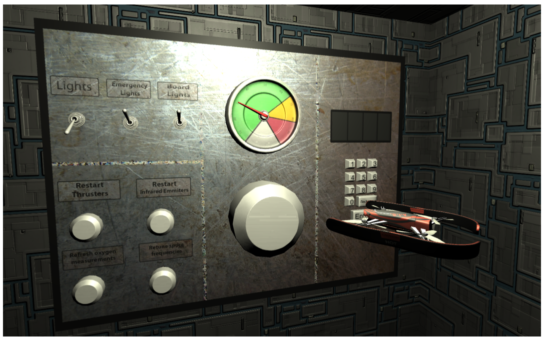

FIGURE 3.7: The view from the astronaut's perspective — the drone takes up a small amount of the FOV, and its position is directly mapped to the second controller's position

This way, the non-HMD user can not only be easily tracked, but they can also use the controller to interact with the virtual environment, giving them more presence in the virtual space, as well as solving the tracking problem with already present technology (as HMDs often come with two controllers). The game is designed to be able to run on a wide range of Android phones (the one used in the experiment being a Samsung Galaxy Note 9).

For the PC version, the same machine that is running the game was used, controlling it using a keyboard, and testing on a Dell 34″ curved monitor (3440x1440), ensuring that the monitor was positioned in such a way that the HMD user would be *behind* the PC player (Figure 3.5).

3.1.2 Software Implementation

The game was developed using Unity 2019.2.8f1, a free and popular engine that has support for multiple platforms, including VR and Android. Additionally, the official SteamVR plugin for VR integration was utilised. Networking functionality was implemented using the Bearded Man Studios "Forge Networking Remastered" framework [1], a Unity Asset Store package facilitating multiplayer development.

Two choices were available for the Android version: streaming the rendered video from the desktop to the phone, or sending just the tracking data and letting the phone render the scene. The second approach was chosen, as sending more data (a video stream in this case) would introduce higher latency, as well as being harder to implement.

A server is started on the desktop by the game , and waits for a connection from the phone through Wi-Fi (using Forge Networking to establish communication). This can also be used by people who are not playing the game but still want to see what is happening. The PC version does not initiate a server, as the game is already running on the same machine, but instead creates a different instance to allow the player to control the view.

One of the design decisions that were made for the game was using the different modalities in ways that enforce their uniqueness, as suggested by Gugenheimer et al. [36]. For example, the HMD user experiences a fully immersive VR environment that tracks their head and hand movements, while the non-HMD user uses a phone or a PC to move around the world and access information unavailable to the HMD player. For the HMD version, some general guidelines and best practices were followed when creating the experience, such as attempting to minimise motion sickness [56] by using hardware that is powerful enough to keep the recommended 90 FPS, as well as avoiding moving the player camera during the game. As Michalak [56] mentions, graphics were kept consistent and fairly realistic.

Another important guideline mentioned by Gugenheimer et al. [36] is designing the experience with the shared physical space in mind. in the case of the Phone version, people would be close to each other, and potentially bump into each other. For that reason, virtual representations of both players was added to the game. This way, the HMD user would be able to see where the drone is and be more careful. The drone was also designed to be a specific size, larger than the controller, in order to minimise risk of hitting the other player accidentally.

For the virtual representation of both players, simple 3D models sourced from the internet were used, such as an astronaut helmet and a modified drone model to fit the

[1]https://assetstore.unity.com/packages/tools/network/forge-networking-remastered-38344

theme. The default hand provided by SteamVR was used for the HMD player as it resembled an astronaut glove.

An attempt was made to distribute the power level between the two players, where both players would feel they are needed to complete the game and increase levels of cooperation. The HMD player had more control over the environment and could interact with various object and switches to activate or deactivate certain features. The non-HMD player had more access to information and instructions for solving the puzzles.

In the Phone version, the goal was for the player to hold the phone with one hand and the controller with the other (see Figure 3.4). Efforts were made to minimise lag between input and output in order for the player to feel like they are in control in real time, helping with their sense of *body ownership*, *agency* and *self-location*.

For the desktop version of the game, a keyboard-based control scheme was chosen to mirror the current market standard and 'emulate' the operation of a real drone.

This control scheme is also similar to how a real drone would be operated, with two sticks that control yaw, pitch and roll. In this case, the "WASD" keys controlled horizontal movement, while the arrow keys controlled the elevation and rotation.

A major design decision was having the non-HMD user be more of a helper role and not have as much power over the world, with the reasoning that, as Cheng et al. [19] show, this would still be enjoyable to the player.

In terms of audio, the tutorial for both players was played through the HMD headset and the Phone/PC speakers (subtitles were included to help people understand more easily if they misheard what was said). There were minor sounds cues provided to the HMD player, such as flipping switches and pressing buttons, to indicate when actions had been successfully performed. As an HTC Vive headset was used, which does not include headphones, over-ear headphones were provided to each participant due to their comfort.

3.2 Methodology

The main goal of the experiment was to investigate how a non-HMD user's mode of play affected the co-presence and immersion of the participants.

The two ways that were tested were through a Phone or a PC. The desktop version of the game is the baseline, as it is what is mostly used in the industry right now [83, 62, 92]. The Phone version served as the experimental condition, where the expectation was to see a difference in terms of immersion and co-presence.

Two main questionnaires were used in the study — the *Networked Minds Measure of Social Presence (NMMoSP)* [6, 37] and the *Igroup Presence Questionnaire (IPQ)* [74].

The *NMMoSP* assesses various factors and aspects of social presence, such as isolation/inclusion, mutual awareness and mutual understanding [6].

When creating the interview questions, those factors were taken into consideration. This questionnaire was chosen due to the fact it had already been used in research in this sphere [65, 34, 87, 10], including in a modified context to focus on a particular factor of the experience.

Concerning measuring 'immersive response', the *iGroup Presence Questionnaire (IPQ)* was chosen for its inclusion of sub-scales that allow for a more in-depth view into the participants's immersive experience. The IPQ evaluates four dimensions of presence in virtual environments:

- General Presence: the overall sense of "being there" in the virtual environment;
- Spatial Presence: the feeling of physically being in the environment, as if one could interact with it directly;
- Involvement: the degree to which a participant feels focused and engaged with the virtual environment;
- Experienced Realism: the extent to which the virtual environment feels authentic and believable.

These sub-scales allow for a nuanced analysis of immersion, as they address distinct aspects of the experience that could possibly respond differently to the varying modes of play. For instance, spatial presence could be influenced more by the field of view an HMD provides, but involvement could vary based on the chosen interactions. This is particularly relevant when comparing different modes in a multi-modal experience.

Unlike broader questionnaires, such as the 'Witmer and Singer Presence Questionnaire' [94], which has been critiqued for its subjectivity and overlap between dimensions [78], the IPQ's relatively independent sub-scales [74] made it a more suitable choice for this study. Due to its concise nature, it also made it less likely for participants to feel fatigued.

3.2.1 Study Design

A mixed design approach was chosen, where each session involved two participants, with one assigned the role of the HMD player (astronaut), and the other using either a PC or a Phone. The experiment consisted of two playthroughs, during which the participants swapped roles to experience the game from the other mode. This ensured that both participants played as the HMD player and the non-HMD player.

Each participant, therefore, contributed data from two playthroughs, allowing for the capture of their experience as both the HMD and non-HMD player.

3.2.1.1 Conditions

Participants were divided into two groups based on the platform assigned to the non-HMD player:

- HMD-PC: One participant used VR, while the other used a PC;
- HMD-Phone: One participant used VR, while the other used a Phone.

This means that there were four distinct modes of play:

- HMD^{Phone}: The HMD player with a Phone-based partner;
- Phone: The Phone player when paired with an HMD user;
- HMD^{PC}: The HMD player with a PC-based partner;
- PC: The PC player when paired with an HMD user.

3.2.1.2 Randomisation and Control

To ensure fairness and mitigate order effects:

- Participants were randomly assigned to either the **HMD-PC** or **HMD-Phone** group.
- The actions needed to complete each puzzle were randomised within for each playthrough, ensuring a consistent challenge while avoiding repetition.

By swapping roles within a session, this study design captured the experiences of both roles across the chosen platform conditions.

3.2.1.3 Measures and Data Collection

Co-presence was measured using a modified version of the *NMMoSP* [37] questionnaire, which consists of 19 items divided into six sub-dimensions: co-presence, attentional allocation, perceived message understanding, perceived affective interdependence, perceived behavioural interdependence, and perceived emotional interdependence. Eight items were selected from three sub-dimensions that were most relevant to the research question: co-presence, attentional allocation, and perceived message understanding. The items used a seven-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

Immersion was measured using a modified version of the IPQ [74] questionnaire, which consists of 14 items divided into four sub-scales: general presence, spatial presence, involvement, and experienced realism. Seven items were selected from three sub-scales that were most relevant to the research question: general presence, spatial presence, and involvement. The items used a seven-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). For both dependent variables, the aggregated score was calculated following the instructions for each questionnaire, as well as examples from the literature.

Not using the full questionnaires was a choice driven by the time constraints of the experiment. Previous research that used a subset of the scales of the NMMoSP questionnaire has demonstrated their validity in similar contexts ([65, 34, 87, 10]). Additionally, the IPQ consists of three sub scales that are fairly independent [74], suggesting that they could be used separately if the research requires it. This further supported the decision to limit the size of the questionnaires without compromising the collection of the data.

This experiment aims to investigate the potential differences in co-presence and immersion across each mode for both the HMD player and the non-HMD player as per Research Question 1.

The *Software Usability Scale (SUS)* [14] was chosen in order to provide a metric that can show the game created is usable and the findings from the experiment are valid (as opposed to being negatively influenced by a poor experience). The *SUS* is a short ten question 5-point Likert scale questionnaire that was given to each participant after every playthrough. The SUS score can range from 0 to 100, with higher scores indicating higher usability.

A semi-structured interview was conducted at the end of the whole session with the two players together. The questions of the interview were based on the *NMMoSP* and *IPQ*, as well as any observations made during the experiment, such as instances where one group of players relied heavily on the use of physical cues and movement to play the game. The analysis is therefore both quantitative and qualitative. The interview aimed to elicit the participants' subjective perceptions and experiences of co-presence, immersion and game enjoyment, as well as their preferences and feedback on the different modes of play.

The questionnaires given to each participants, as well as a list of the interview questions asked, can be found in Appendix B, while the Participation Information Sheet and the Consent Form given to participants can be found in Appendix A.

Times to complete each game were also logged in order to demonstrate that each game was played to the finish, thus making sure that all intended interactions were experienced by the participants. This also helped with the detection of any discrepancies between playthroughs or between participants, allowing for the recognition of abnormal times where people got stuck during gameplay.

3.2.2 Ethical Considerations

Prior to the start of the study, ethical approval was obtained from the University of Southampton's Ethics board, under the reference number ERGO 54456. This approval ensured that the research was conducted in accordance with the ethical standards of the university.

All data collected during the study was anonymised and securely stored on a password protected university PC behind a staff-only access-controlled area to prevent unauthorised access. The VR game was designed to minimise any discomfort or adverse effects to participants, and participants were told beforehand that they can end the experiment early in case they start feeling nausea.

3.2.3 Setup

The experiment took place in a university lab with an HTC Vive setup with over-ear headphones, 3x4 m^2 play area, running on a Windows 10 desktop computer (Intel Core i7-7700 3.60GHz, NVidia RTX 2060 Super, 16 GB RAM). Participants were University students/staff recruited using posters around campus, and divided equally into two groups at random: the HMD-PC version and the HMD-Phone version.

3.2.4 Procedure

Each session followed the same sequence:

1. **Introduction and Consent:** The two participants were welcomed, received an overview of the study, and signed the Participation Information Sheet and Consent Form (see Appendix A);

2. **First Playthrough:** One participant was randomly assigned to be the HMD player (astronaut), while the other used either a PC or a Phone. They played the game cooperatively, interacting with puzzles and tasks;
3. **Questionnaires After First Playthrough:** Both participants filled out the *NMMoSP*, *IPQ*, and *SUS* questionnaires based on their experiences in their respective roles;
4. **Role Swap and Second Playthrough:** The participants swapped roles (i.e., the HMD player became the PC/Phone player and vice versa) and repeated the game;
5. **Questionnaires After Second Playthrough:** The same set of questionnaires (*NMMoSP*, *IPQ*, and *SUS*) was administered again, capturing how each participant experienced the opposite role;
6. **Semi-Structured Interview:** After both playthroughs, a joint interview was conducted. The topics included co-presence, immersion, overall game experience, and any notable observations (see interview questions in Appendix B);
7. **Debrief:** Finally, participants were thanked, debriefed on the purpose of the study, and given an opportunity to ask further questions.

By following this procedure, each participant experienced both roles (HMD and non-HMD), ensuring that any differences in co-presence or immersion could be observed directly within-subject. Additionally, the semi-structured interview captured more nuanced insights into their subjective experiences.

3.3 Results and Analysis

In total, 5 female and 19 male participants took part in the study, with the average age of 21.46 ($SD = 1.52$). The majority of players had minimal first-hand experience with VR, as a lot of people experienced VR for the first time in the study. However, many people expressed familiarity with the game *Keep Talking and Nobody Explodes* [83], one of the main inspirations of *Stuck in Space*. Only 2 groups were made up of strangers who did not know each other before the experiment.

Players were divided into two groups of 20 participants each:

- One group used a **PC** when not wearing the HMD (HMD-PC group).
- The other group used a **Phone** when not wearing the HMD (HMD-Phone group).

TABLE 3.2: Mean scores of the SUS, together with sample size and standard deviation.

	n	Mean Score	SD
Phone	12	79.58	12.00
PC	12	81.25	12.97
HMD^{All}	24	81.35	12.55
All	48	80.89	12.27

Within each group, participants completed **two playthroughs**, switching between the HMD and non-HMD roles (PC or Phone). Half of each group experienced the HMD role first, while the other half started in the non-HMD role.

The game on average took 9.77 minutes to complete. As expected, the first playthrough's average duration (11.09 minutes) was longer than the second playthrough (8.45 minutes), as players already were familiar with the game.

The total SUS score for all 48 data points is 80.89 ($SD = 12.27$), which is considered as being in the acceptable range [15]. Following that, it can be inferred that the quality of the game is sufficient for the experiment. Table 3.2 summarises the scores for the different versions.

3.3.1 Quantitative results

The primary questionnaires measured **Immersion** (via the IPQ scale) and **Co-Presence** (via the NMMoSPQ scale). Table 3.3 presents descriptive statistics across the PC, Phone, HMD^{All}, $\text{HMD}^{\text{Phone}}$, and HMD^{PC} conditions.

Given the design, where each participant was assigned to either PC or Phone (**between-subject** factor: *DeviceGroup*) and completed both a HMD and a non-HMD role (**within-subject** factor: *Role*), a **2×2 mixed-design ANOVA** was conducted for each questionnaire measure. Prior to analysis, the standard ANOVA assumptions were checked:

- **Normality**: Shapiro–Wilk tests ($p > 0.05$) and Q-Q plots of the residuals indicated no major departures from normality;
- **Homogeneity of variance**: Levene's tests ($p > 0.05$) showed that variances between PC and Phone groups were comparable within each role;
- **Sphericity:** Since *Role* (HMD vs. non-HMD) had only two levels, sphericity was automatically satisfied.

These checks supported the use of a mixed-design ANOVA. The following subsections present the immersion and co-presence results in detail.

3.3.1.1 Immersion

A mixed-design ANOVA on the **aggregated immersion** scores (factors: *Role* = HMD vs. non-HMD; *DeviceGroup* = PC vs. Phone) revealed:

- **No significant main effect of DeviceGroup**, $F(1, 22) = 0.1762, p = 0.6788$, indicating that PC vs. Phone did not differ overall on immersion.
- A **significant main effect of Role**, $F(1, 22) = 52.08, p < .001$, showing that participants felt more immersed in VR (**HMD**) (M $\approx$ 5.96) than outside VR (**non-HMD**) (M $\approx$ 3.96).
- **No significant interaction** between DeviceGroup and Role, $F(1, 22) = 1.57, p = 0.2234$, suggesting the advantage of HMD immersion over non-HMD did *not* differ across PC vs. Phone.

Pairwise comparisons using EMMeans (Tukey-adjusted) confirmed that immersion was significantly higher in the HMD than non-HMD version within each group ($p = 0.0004$ for PC, $p < .0001$ for Phone). Specifically, the mean difference for the PC group was 1.65 points on the immersion scale, and for the Phone group it was 2.35 points, both in favour of VR. Thus, regardless of whether participants used a PC or a Phone in the non-HMD role, VR consistently had higher immersion.

3.3.1.2 Co-Presence

A similar 2×2 ANOVA was conducted for the **aggregated co-presence** scores:

- **No significant main effect of DeviceGroup**, $F(1, 22) = 1.83, p = 0.1899$.
- **No significant main effect of Role**, $F(1, 22) = 0.0056, p = 0.9410$.
- **No significant interaction**, $F(1, 22) = 0.0505, p = 0.8243$.

Pairwise comparisons within PC or Phone groups found no difference between HMD vs. non-HMD. These results suggest that neither the device nor the role significantly affect the co-presence scores.

The results are, to an extent, surprising. The expectation was that players using the Phone would have increased feeling of co-presence because the two participants are within the same physical play space, as well as increasing the immersion because of the added sensory information provided by the physical presence of the other player.

In order to explain these results, the codes and themes created from the qualitative interviews can be used.

TABLE 3.3: The mean scores with standard deviation for the two questionnaires, with each sub-category used.

	Phone	PC	HMD^{All}	HMD^{Phone}	HMD^{PC}
Aggregated Score (NMMoSPQ)	5.30, sd=0.72	5.72, sd=0.86	5.50, sd=0.79	5.32, sd=0.70	5.68, sd=0.59
Co-Presence	5.90, sd=0.86	6.46, sd=0.99	6.10, sd=0.76	5.79, sd=0.61	6.42, sd=0.79
Atentional Allocation	4.29, sd=1.64	4.83, sd=1.64	4.60, sd=1.55	4.46, sd=1.68	4.75, sd=1.45
Perceived Message Understanding	5.71, sd=0.89	5.88, sd=0.61	5.79, sd=1.03	5.70, sd=1.05	5.88, sd=1.05
Aggregated Score (IPQ)	3.71, sd=1.4	4.20, sd=1.41	5.95, sd=0.63	6.06, sd=0.58	5.85, sd=0.68
General immersion	4.33, sd=1.30	4.83, sd=1.53	6.54, sd=0.66	6.67, sd=0.65	6.42, sd=0.67
Involvement	2.79, sd=1.42	3.48, sd=1.51	5.13, sd=1.06	5.33, sd=0.73	4.94, sd=1.31
Spatial presence	4.00, sd=2.03	4.29, sd=1.67	6.19, sd=0.84	6.17, sd=1.05	6.20, sd=0.62

3.3.2 Qualitative results

Inductive coding was used, resulting in 13 codes spread over 5 themes: *Cognitive Engagement*, *Embodiment*, *Sensory Perception*, *Knowledge*, and *Agency in the Virtual World*. To confirm the reliability of the results, a second coding activity using the same code book was then undertaken by the same coder four months after the initial coding. This delay aimed to reduce potential bias by allowing the coder to approach the data with a fresh perspective. The intra-rater reliability achieved was 80% and the Cohen's Kappa was 0.64, indicating substantial agreement [48].

The 13 themes are shown in Table 3.4 with the number of positive and negative occurrences, and example quotes from one of the participants (from P1 to P24) categorised as Co-Presence (CP) or Immersion (I). Additionally, a '+' or '-' sign shows whether a comment was deemed positive or negative. The codes were shared between Co-presence and Immersion, with the exception of "*Control of the world*" and "*Interface mapping*" (unique to Immersion) and "*Observability of Avatar*" and "*Rationalizing Experience*" (unique to Co-Presence). The following subsections describe each theme.

3.3.2.1 Cognitive Engagement

This theme is about people mentally engaging in the activity. This could be through *Conversation* with the other player, or through a *Collaborative Task*:

> "I felt that someone is constantly guiding me as the drone, so (...) it wasn't lonely at all" P5 CP+

It could also be the overhead of *Maintaining a Mental Model of the Real Environment*. The mental model arises because the HMD user playing with a Phone user has to not only work in the virtual environment, but also think about the physical one:

TABLE 3.4: Observations: Themes, Sub-Themes, and Example Quotes.

Theme	Sub-theme	# of codes Co-presence +ve	Co-presence -ve	Immersion +ve	Immersion -ve	Example Quotes
Cognitive Engagement	Conversation	5	0	8	2	"We're communicating in real life so I (...) could feel his presence" (P13, CP+)
	Collaborative task	8	2	10	1	"It felt like teamwork, you're (...) both doing part of the same task" (P1, CP+)
	Maintaining a mental model of the real environment	1	0	6	6	"You have that fear of bumping into something" (P13, I-)
Embodiment	Narrative embodiment	1	0	5	0	"P10 got in the role of being the drone" (P3, I+)
	Physical embodiment	1	4	3	2	"I (...) connected the drone with the voice that I'm hearing" (P2, I+)
Sensory Perception	Observability of avatar	6	3	0	0	"I saw P21 pushing buttons, so I didn't feel alone" (P6, CP+)
	Observability of body	9	1	1	6	"I bumped you and then I was more aware that you're there" (P4, CP+)
	Observability of virtual environment	0	1	0	3	"It felt like I was just trying to look in through a tiny little window" (P7, I-)
	Observability of real environment	2	3	0	7	"With the VR, I wasn't very aware of the (real) world, I only sensed the cable" (P3, I-)
Knowledge	Prior experience	1	2	1	0	"I've played VR (...), so I'm used to it and forgot about the outside" (P6, I+)
	Rationalizing experience	3	2	0	0	"I don't feel alone, because I know there's another person" (P13, CP+)
Agency in the virtual world	Control of the world	0	0	7	0	"I feel like I have control, makes me feel like I'm in the game" (P18, I+)
	Interface mapping	0	0	2	2	"flicking the switch, and putting the numbers on the keypad really feels real" (P5, I+)

> "The drone started to run around me and I was thinking 'Oh wait, there's a person running around me as the drone, I don't want to hit them' " P9 I-

3.3.2.2 Embodiment

Embodiment describes the way in which the player perceives themselves as embodying the avatar. This encompasses *Physical Embodiment*, where, for example, players see a drone in VR, know it is a real person, but still embody or feel the drone avatar as the person:

> "In VR I kind of attached their voice to the drone, so whenever I'd hear P12's voice I'd think 'Ok, this is the drone, it's speaking to me' " P16 I+

Narrative Embodiment on the other hand is more about the players getting into the role assigned to them (role-playing):

> "You have to be more forceful in terms of proactively immersing yourself into the game if you're on the phone" P18 I+

Both *Physical* and *Narrative* embodiment are terms that came from the data, and do not fit necessarily into the three embodiment types that were discussed in the Background (Chapter 2).

3.3.2.3 Sensory Perception

This theme centres around perceiving the other through one's senses. It therefore includes observability through various senses (sight, touch, smell, etc.). It refers to both the player (as Observability of either their physical *Body* or their virtual *Avatar*) and the environment (both the *Virtual Environment*, and also the physical *Real Environment*).

3.3.2.4 Knowledge

In some cases people had either *Prior Experience* with VR or games in general, or they tried to *Rationalise the Experience*. In the first case, players either talked about never playing the game so they did not know what to do exactly, or they have played cooperative games before:

> "It feels like that I'm at home and I'm playing a video game with my phone and I'm alone in the room, and the person is playing inside the game, like a character inside the game" P18 CP-

People also came with the preconceived idea that they would be playing together with another player:

> "I know P23 is still there watching over me and stuff, so I totally feel their presence" P13 CP+

3.3.2.5 Agency in the Virtual World

This theme was unique to Immersion, relating to the participants' sense of control over the virtual environment. Non-HMD players found that having some kind of *Control of the World* makes them more immersed:

> "(...) when I was with the phone, and had to turn on the flashlight, so I had even more presence in P22's side" P4 I+

This type of agency is about the ability to influence the world in general, rather than specifically referring to interactions through an avatar [61]. It can be thought of closely related to the *agency* concept in embodiment [42].

Participants commented on how effective it is for the virtual world to mimic the real-world actions within the game (flipping a switch, pushing a button in, etc.). This was called *Interface Mapping*:

> "I forgot I had the controller in my hand, (...) because what you could see is your hand, so I was like 'It's fine, I'm just poking things' " P15 I+

3.4 Discussion

The quantitative data showed that there was no significant difference between any of the versions in terms of co-presence, and the only significant results for immersion being the levels between the HMD^{All} players and the Phone or PC players, which is an expected outcome (because VR is recognised as being more immersive). Analysing the interviews and the themes and sub-themes gives some insight into the lack of impact of the different mediums of play.

When looking at the themes, some relationships start to emerge. It can be seen how *Prior experience* and *Embodiment* are connected, as if you knew what to expect it is easier for you to embody the role physically and narratively. Some non-HMD players mentioned that they did not know what they looked like at the beginning of the game, they did not know the size and shape of the drone, which made it harder for them to embody the role. Having a **mirror in-game** might have helped them with that.

Physical embodiment and *Observability of body* also seem to be connected. An increase in physical interactions, such as bumping into the other player, appears to enhance *Physical embodiment* for the HMD^{Phone} players. This is likely due to the addition of sensory input beyond voice, which strengthens the connection between the virtual drone avatar and the real-world player. One participant described this explicitly: "I bumped you and then I was more aware that you're there" (P4, CP+). These interactions create a stronger association between the physical presence of the other player and their virtual representation, making it easier for HMD players to embody the drone. In contrast, PC players did not mention *Physical embodiment*, possibly because their interactions lacked a similar level of sensory reinforcement.

When talking about *conversation*, HMD players mention that the voice they heard (observed) either helped them with immersion, or not. In most cases, it helped them, but some people mentioned that if the voice came from within the headphones they would be more immersed, or in other words, **taking the real world and putting more of it in the virtual world** would have helped. Connecting it to the themes it can be said that taking *observability of the body* and turning it into *observability of the avatar* would make for a more immersive experience (when the other player is a mix of virtual and physical stimuli, it is more distracting as you have to keep two instances of the player in mind).

This further shows how *Maintaining a mental model of the world* and *Observability of the body* are connected. The difference between the physical and virtual world (the fact that the phone player is beside the HMD player and there is not a 1:1 mapping of the real body to the virtual avatar) increases the mental strain of keeping the model of the real world in mind. One way that this might be alleviated would be to have a **more direct mapping between the real body and the avatar** (in the case of the Phone player). Some players on the other hand got completely engrossed in the VR world, being reminded of the physical one only by bumping into the other player.

These inner-relationships between the themes and sub-themes is what can be built upon when talking about the differences in the versions of the game, both between PC and Phone, and the two HMD versions. The data shows how the Phone player's immersion is negatively affected by the *observability of the body* and *world*, whereas the PC player does not have these problems. One way to explain that is the fact that the phone is a small screen that takes a tiny amount of your field of view, while the monitor used in the experiment is much bigger and thus makes it easier to lose yourself in the virtual world. Moreover, being on the phone with an HMD player in front of you can distract you because they would be moving a lot, taking your concentration away from the screen. On the other hand, co-presence for the VR player is positively affected by this. Comments such as "I know P23 is still there watching over me and stuff, so I totally feel their presence" (P13, CP+) and "I felt that someone is constantly guiding me as the drone, so (...) it wasn't lonely at all" (P5, CP+) indicate that the physical presence of the other player enhanced co-presence. This was not as frequently observed for PC players, suggesting that the visibility of the other player's physical body in the shared space plays a role in co-presence.

Narrative embodiment seems to be easier to achieve on the Phone, as there was a lack of comments from the PC side. One way to explain this is looking at the ways that the two versions interact with the virtual world. One has to physically move to go around, while the other can mostly just sit without moving significantly. This in turn would help the Phone player 'feel' more in the role of a drone, rather than controlling a drone separately with PC controls. One person mentioned how they actually found the PC version more immersive because they did not have a fear of bumping into something, which further shows how *Maintaining a mental model of the real world* can affect immersion. Even when there was nothing to bump into on the HMD version, they still experienced the fear.

It was noted that *control of the world* increases immersion for both non-HMD players, which is similar to what Gugenheimer et al. talk about in their paper, where giving the second player more power to interact with the virtual world increases enjoyment rather than immersion [36], as well as what Bangay and Preston say about control [1].

The qualitative data suggests that the *mental model of the real world* plays a nuanced role in the HMD player's experience. While no significant difference in immersion was noted quantitatively between the HMD-Phone and HMD-PC modes, the interview responses highlighted how the physical presence of the Phone player impacted both immersion and embodiment. For example, some HMD players said negative comments due to the need to maintain this *mental model of the real world*: "The drone started to run around me and I was thinking 'Oh wait, there's a person running around me as the drone, I don't want to hit them" (P9, I-).

On the other hand, physical interactions such as bumping into the Phone player helped with physical embodiment: "In VR I kind of attached their voice to the drone, so whenever I'd hear P12's voice I'd think 'Ok, this is the drone, it's speaking to me' " (P16 I+). This connection between the physical and virtual world made the Phone player's physical presence feel aligned with the drone avatar, reminding the HMD player of reality. This makes *embodiment* much easier than on the PC, and thus can help with co-presence.

In most cases, *Conversation* leads to an increase in immersion and co-presence for the HMD player. Multiple people mentioned the fact that because they were **talking about the game world and not about something outside of it**, they felt more co-present and immersed. It should be noted that an important factor is where the voice is coming from: **within the headset/earphones or from outside**. Some people could not hear their partner's voice as well, which led to them having to concentrate more on the voice rather than on the experience. This can partly be helped by using either a headset that can pass through noise from the outside, a more open speaker design (such as the Valve Index[2]), or even having the voice of the second player be

[2]https://www.valvesoftware.com/en/index/deep-dive/ear-speakers

recorded from their Phone/PC and be transmitted to the HMD player in their headset simultaneously.

These relationships suggest that there are different features that affect (and mitigate) each other. For example, the aforementioned *Cognitive Engagement*, in which the overhead of maintaining a mental model of the physical space supports co-presence but detracts from immersion. Similarly in *Embodiment* the HMD player's sense of co-presence benefits from the Phone player being in the same physical space because their spatial proximity matches their avatar, but at the same time physical interactions (literally bumping into the other player) detract from immersion. Balancing each of these interactions between the different modes can lead to a more immersive or co-present experience.

3.5 Summary

From the resulting relationships between the themes, one could see that *embodiment* plays an important role. There is a distinction to be made though, as *embodiment*, both what are coded as *narrative* and *physical*, are not necessarily terms taken from the literature, but more so terms that were created from the underlying data of the interviews. In order to explore the relationship *embodiment* has with the other sub-themes, a solid understanding of it was needed, and of ways of influencing it. It was theorised that if the *embodiment* in one modality was improved, then the overall experience would be improved, as *embodiment* was connected to so many other sub-themes, and improving it for one person has the potential to improve it for the other as well.

Thus, in order to better understand *embodiment*, and the different techniques to influence it, a systematic literature review was decided to be conducted. The following chapter is going to go into details about it, as well as the resulting framework created from the data, and following that, trying to use the framework in order to understand the results from the first experiment better.

Chapter 4

Systematic Review of Embodiment Systems

After conducting the first experiment and analysing the results, a few observations were made. The embodiment theme, be it *physical* or *narrative*, was tightly connected with the other themes and sub-themes. One could see that embodiment played an important role in the experience in terms of Immersion and Co-Presence. Thus, it was theorised that improving the feeling of embodiment of one player in one modality would translate to an overall improved experience not only for them, but for the other player as well.

To influence and improve embodiment, a framework or system was necessary. Upon reviewing existing literature, it became apparent that there was no single agreed-upon definition of embodiment. Some researchers talk about the 3 concepts of embodiment from Kilteni et al. [31, 69, 76, 12], others use the term "Embodied Cognition", which is the theory that cognition is dependant on the body itself, its features and surroundings, and how all that combined helps cognitive processes [33, 93].

In this chapter, a systematic review will be conducted to navigate through these varying perspectives of embodiment, aiming to establish a foundation for subsequent research and exploration.

4.1 Systematic Review

In order to have a clearer picture of embodiment and the different ways of affecting it, I decided to conduct a systematic review. The review included an examination of the different types of embodiment present in the literature, as well as the methods that were used to influence it.

4.2 Methodology

The research question this chapter is going to answer it "*RQ2. What are the different kinds of embodiment and what are the different strategies that are used to help achieve them?*".

4.2.1 Protocol

To conduct the systematic review a standard process was followed, consisting of *Research Identification*, *Study and Quality Selection*, *Data Extraction*, and *Data Synthesis* [45].

4.2.1.1 Research Identification

"The Engineering Village[1]" was chosen as the target of the review, as it contains a very large number of databases (Ei Compendex and Inspec for example), and an extensive search function. It covers all major conferences and journals, and digital libraries such as the ACM DL and IEEE Xplore.

The search query that was used to get the list of papers was:

> *(embodiment OR embody OR embodied)[a] AND ((multimodality OR multi-modality OR multimodal OR multi-modal) OR (VR OR Virtual Reality) OR (AR OR Augmented Reality) OR (MR OR Mixed Reality))[b]*

This query included a number of different terms for embodiment in *a*, and a number of different definitions for each modality (*b*). What this search query then returned was a number of papers that mention embodiment and one of the modalities.

Saturation sampling was used as a way to scope the study, as the potential sample size is very large, and my aim is a classification model rather than a comprehensive overview of the popularity of different terms in the literature. Papers were inductively analysed in batches of 20, coded, and the codes synthesised into an evolving code book. This process was repeated until no new codes were present for two consecutive batches.

4.2.1.2 Study Selection and Quality Criteria

Only papers published in the last 10 years were used (2010-2020) with the intention of generating a contemporary rather than a historical list of definitions and techniques.

[1]https://www.engineeringvillage.com/

The subsequent list of papers created was then used, and any following exclusion criteria apply to the batches themselves (e.g. if the batch contained 20 papers, but 3 of them were not in English, then only 17 papers would be reviewed).

For practical reasons, it was chosen to exclude any papers that weren't written in English, as well as papers that didn't present directly or indirectly a definition of embodiment in the main text of the paper. This led to the exclusion of papers that only mentioned embodiment in passing, or did not contain sufficient detail to classify their view of embodiment.

There are some biases that could be expected from this sample in that "The Engineering Village" was used for the systematic review, which will reflect a more technology-focused perspective of embodiment (through the perspective of HCI for example), rather than possible psychological or philosophical views of embodiment. This suits this thesis' purpose, but does mean that the models developed reflect a view of embodiment as represented in interaction and design.

4.2.1.3 Data Extraction

The data extraction and primary analysis was done by the author, while the supervisors oversaw the process, and periodically reviewed the code book. The coding of each paper was undertaken using a table to record the evolving code book, with papers as rows, and embodiment, strategy, and mapping between the two as an evolving set of columns. The steps undertaken were:

- First, any papers that passed the exclusion criteria would be read through, and any mention of embodiment would be recorded;
- If the embodiment type was already in the code list, then an 'E' would be added in the cell matching the paper's row to denote an embodiment type. If it did not exist then the new code would first be added as a new column;
- Similarly, any strategies used in the paper to affect that kind of embodiment were added to the same code list as a new column, and the letter 'S' would be added in the cell matching the paper's row.
- Finally, the mapping between the types and strategies were written down as a separate column, using the combination code "Embodiment–Strategy" — and the letter 'M' was used to denote the presence of a mapping in the appropriate cell for that paper.

There were a number of instances where people talked about embodiment without necessarily talking about a specific definition, where some papers would just define embodiment using their own words. In that case, a specific category was used, 'Embodiment – Generic'. These were later reviewed in order to see if they fit any of the evolving codes. Similarly, papers would also talk about one type of embodiment, which would be similar to how another paper talks about another type of embodiment, in which case, those two codes would be combined in the code book and their descriptions updated.

In some cases, researchers would apply a certain strategy specifically to one of the three components of the embodiment described by Kilteni et al., and in other cases, they would refer to that trinity of definitions as a whole. To maintain this nuance in the data, four codes were used: one for the entire network and three for its individual components. Strategies were sometimes therefore mapped to the higher level code and in Section 4.3.5, where the mappings are presented, occurrences are shown where this mapping was moved to a more specific location where possible once the coding process was complete.

4.2.1.4 Data Synthesis

In the data synthesis stage, the research team used the data that was extracted and in an iterative brainstorming process established relationships between the concepts, and using these relationships uncovered two orthogonal models, refining them until they best fit the data. This was also an inductive process, and comparisons to existing models was not undertaken until after it was complete. However, quite early in the coding it became apparent that a lot of the codes matched with a number of the models in the literature. Consequently, existing terminology was adopted where the concepts sufficiently overlapped.

4.3 Results and Analysis

In total, there were 10 batches in the sample, consisting of 200 papers. Of these 140 were excluded based on the criteria, and 60 were included in the final analysis. This resulted in 16 codes for embodiment, 20 different strategies, and 39 mappings between embodiment and strategies (Table 4.1 shows the codes per batch, including the final two batches where no new codes were found). Three high-level themes were identified, as shown in Table 4.2, which also presents the codes from the inductive analysis, along with the definitions, and a total number of papers that contained that code.

A second coding session was undertaken 6 months after the initial coding to confirm the reliability of the results, this was based on 6 papers representing 10% of the sample. Intra-rater reliability was 85%, while Cohen's Kappa was 0.84, indicating near perfect agreement. There were three occurrences of differences, all associated with the aspects model and whether papers were coded, or strategies mapped, against 'Embodiment - generic' or one of the specific sub-types. This is caused by authors framing their work using general definitions, but then using a more specific concept in their detailed discussions, which again highlights the need for more precise terminology.

There are two main orthogonal perspectives uncovered through the systematic review, represented by the first two themes, which are *Aspects* and *Levels* of embodiment. The third theme, *Higher level combinations*, covers specific combinations of the Aspects and Levels that were found in the literature.

In addition, Table 4.3 shows the strategies found in the papers, with a brief description of what each strategy is, and references to the papers that used that strategy. In total there were 20 strategies, but some could be generalised (although the generalisations do not appear in the original data). For example, *Low-fidelity graphics* can be generalised as *Different graphical settings*. The following sections talk through the creation of the diagrams, as well as the analysis of the systematic review data in more detail.

4.3.1 Creation of the taxonomic diagrams

The codes derived from the data were synthesised into broader categories to form the basis of the two main models. The codes "Body Ownership", "Agency", "Self-Location" and "Social Embodiment" were grouped due to their thematic similarities. This grouping then led to the "Aspects" model of Embodiment. After a few iterations, a Venn diagram was chosen as a visual tool to represent how the strategies found in the review fit in the model, and how a number of them overlap (Figure 4.1). This helped to show how various strategies impact different aspects of embodiment, either at the same time or separately.

Another theme emerged from the analysis, alluding to the depth of embodiment. The codes "Embodied Avatars", "Embodied Cognition", "Embodied Emotion" and "Embodied Identity" were then grouped, and a pyramidal structure was chosen. It represents the hierarchy of embodiment depth, where each level builds upon the previous one (Figure 4.2).

Both of the models were informed by the patterns and relationships seen in the data. The Venn diagram of the "Aspects" captures the different sides of embodiment, while the Pyramid diagram captures the "progression" of embodiment.

TABLE 4.1: Table showing the number of papers used in each batch, as well as the amount of new codes (E - embodiment, S - Strategy, M - Mapping).

Batch number	Accepted papers	New codes		
		E	S	M
1	10	12	6	8
2	4	1	2	4
3	6	0	1	2
4	11	4	2	6
5	3	0	2	3
6	6	1	3	4
7	6	1	1	3
8	7	0	3	10
9	4	0	0	0
10	3	0	0	0

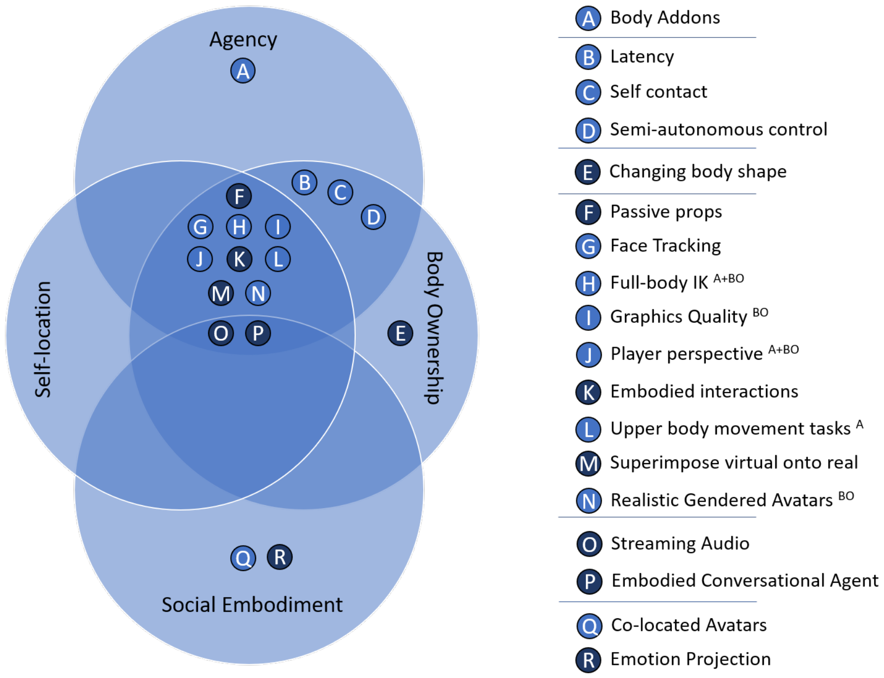

FIGURE 4.1: A Venn diagram showing how the aspects are connected, together with the strategies that affect those aspects. Dark blue means that it was specialised, superscript means that the strategy was also targeted more specifically (i.e. at Agency, Body Ownership, or both.)

4.3.2 Embodiment Aspects

There were eight embodiment codes in the Aspects theme that came up from the review, from which four codes made up the final Aspects model (shown in italics in Table 4.2), while the other four were either sub-codes or a generic code. One can see the Self-location (S-L), Agency (A) and Body ownership (BO) components of embodiment from Kilteni et al. — papers would reference that model specifically in five different ways: either through one of the components only (S-L, A or BO), all three components together, or embodiment in a generic sense (meaning in the context

TABLE 4.2: All embodiment codes that came from the review with definitions (and references), divided by each theme and ordered by number of mentions (**bold comments show similarity between different definitions**)

Theme	Embodiment code	Definition	Mentions
Aspects	Embodiment - A/BO/S-L	Using the 3 component definition [S47, S10, S39, S48, S34, S57, S49, S30, S8, S36, S12, S28]	12
	Embodiment - BO component	Body ownership component [S16, S52, S33, S43, S27, S29, S5, S8, S12, S60, S25, S14]	12
	Embodiment - Generic	No specific definition given [S54, S33, S24, S27, S40, S51, S58, S31]	8
	Embodiment - A component	Agency component [S5, S23, S1, S13, S8, S12, S25]	7
	Embodiment - S-L component	Self-location component [S27, S25]	2
	Social Embodiment	The way the user's state is when in a social context [S37]	1
	Embodied Social Virtual Reality (SVR)	Distributed and embodied face-to-face interactions (**part of Social Embodiment**) [S32]	1
	Illusion of virtual body ownership (IVBO)	Referred as body transfer illusion, agency, or embodiment **(close to Body ownership)** [S29]	1
Levels	Embodied Avatars	Perceiving the avatar as co-located and part of one's own body [S42, S16, S50, S17, S9, S3, S32, S24, S29, S4, S35, S49, S26, S41, S19]	14
	Embodied Cognition	The theory that human cognition is influenced by the body itself and the interactions with the environment [S17, S45, S54, S56, S21, S20, S46, S22, S53, S55, S38, S18]	12
	Emotion Embodiment	Embodying the emotional states of a person (making it easier to express their emotional state) [S59]	1
Higher level combinations	Tool Embodiment	A tool being used becomes part of you (you stop thinking about it when using it) [S11, S56, S2]	3
	Gender Embodiment	Embodying a gender **(close to Body ownership & Embodied Avatars)** [S52, S33]	2
	Embodied Simulations	Simulating another person's perspective [S15, S44]	2
	Animal Embodiment	Embodying an animal, non-anthropomorphic avatar **(subset of Avatar Embodiment)** [S29]	1
	Embodiment enhancement	Making an operator of a remote robot feel like they're controlling it (being co-located) [S7]	1

TABLE 4.3: All strategies from the systematic review, together with a short description and number of mentions (ordered by most to least mentioned).

Strategy	Description	Mentions
Embodied Interactions	Interactions that involve the body in a natural way [S17, S45, S21, S20, S27, S34, S6, S46, S53, S2, S55, S18]	12
Full-body IK	Using Inverse Kinematics to simulate the whole body through only a number of points (like controller/-headset position) [S16, S59, S39, S3, S48, S33, S43, S29, S5, S60, S25]	11
Gendered avatar through mirror	Having a gendered avatar that matches your own gender, as well as being as realistic as possible **(can be generalised as using a mirror to look at your avatar)** [S10, S52, S33, S30, S26, S8]	6
Face tracking	Tracking face expressions and mapping them onto the avatar [S59, S50, S9, S3, S36]	5
Upper body tasks	A series of tasks that involve the upper body to create a sense of embodiment over the avatar [S10, S23, S13, S12]	4
Latency	Introducing a delay between input and output (the translation of the motion from your physical to virtual hand is delayed) [S8, S12, S25]	3
Low-fidelity graphics	Using low-fidelity graphics to influence embodiment **(can be generalised as Different graphical settings)** [S47, S60]	2
Changing head size	Changing head size to influence embodiment **(can be generalised as changing sizes of body parts)** [S9, S41]	2
Streaming Audio	Streaming voice from the second player directly into the headset [S3, S40]	2
Body add-ons	Adding/wearing props on your body that can translate to virtual body extensions [S35, S1]	2
1st or 3rd Person Perspective	Changing different perspectives depending on what is wanted to be achieved [S8, S28]	2
Superimposing Virtual onto Real *(part of Virtual-to-Physical Transfer)*	Putting virtual models onto real life models/props (through either AR or MR glasses) [S58, S14]	2
Semi-autonomous control	A system that compensates for delay/noise of input [S12]	1
Mimicking a performer	The HMD player is trying to mimic an actor's perspective [S15]	1
Co-located avatars	A different number of co-located avatars and agents in the simulation [S32]	1
Camera on remote robot	Robot trying to mimic what a person does and sends the camera feed back [S7]	1
Embodied Conversational Agent (ECA)	Using anthropomorphic conversational agents that have the same properties as a normal human would have [S51]	1
Self contact in VR	Experiencing your physical body by touch while having the virtual body resemble the real one, which leads to being more grounded [S5]	1
Passive props *(part of Physical-to-Virtual Transfer)*	Passive props that are as close to 1:1 dimensions in the physical and virtual space [S58]	1
Different HMDs	Using different HMDs (with/without cables, lower/higher refresh rate, etc.) [S57]	1

of this model, but not specifying precisely which elements they meant). However, this model of Embodiment wasn't sufficient in explaining all of the aspects of embodiment represented in the codes, as social embodiment was missing. Social embodiment can help describe what is happening in a multi-modal context where there can be multiple people in the same space (be it physical or virtual), similar to how Dourish [23] says embodiment is more than just the person, but their surroundings and the context they're in [23]. The Aspect model therefore emerged as an extended version of Kilteni et al.'s model [S25], with Social Embodiment as an important addition.

Several of the codes were not translated into the model. Embodied Social Virtual Reality (SVR) is arguably a part of Social Embodiment, as it's defined as distributed and embodied face-to-face interactions, and clearly, the social aspect is the critical part. Finally, the Illusion of Virtual Body Ownership can be considered a sub-type of the Body Ownership component, as it's exactly about making you feel like an object is a part of one's body (e.g. the Rubber Hand Illusion).

The complete Aspects model of embodiment that was created can be seen in Figure 4.1 as a Venn diagram, where the aspects are each circle, and the strategies that are associated with those aspects are shown with letters. If the colour is dark blue, it means that the data was specialised (i.e. the strategies were originally coded against 'Embodiment - Generic' based on the review, but in analysis could be seen to more specifically target one of the aspects).

4.3.3 Embodiment Levels

If Aspects is about the breadth of embodiment then the second model that emerged from the codes, *Levels*, is about the depth. It is orthogonal to Aspects; here the codes represent an increasing quality of embodiment, that becomes more complete and sophisticated as higher levels are reached. The model is shown in Figure 4.2.

Unlike Aspects, there is no overarching model in the literature that includes all the components of the Levels model - all of them were observed independently, except the highest level (*Embodied Identity*). In particular, it brings together three different views of embodiment: *"Embodied Cognition"*, *"Embodied Emotion"* and *"Embodied Avatars"*, which are all terms from the existing literature. It is argued that these terms are related and that one can talk about these terms in an ascending way, where the more embodied one is, the higher up the pyramid they are. One can just feel that they are the avatar but not think like it, or one could be emotionally embodied in the avatar. This perspective revealed a missing capstone from the pyramid, *"Embodied Identity"*, that would be the culmination of all the three levels under it, where one both feels and thinks as the avatar and thus identifies with it.

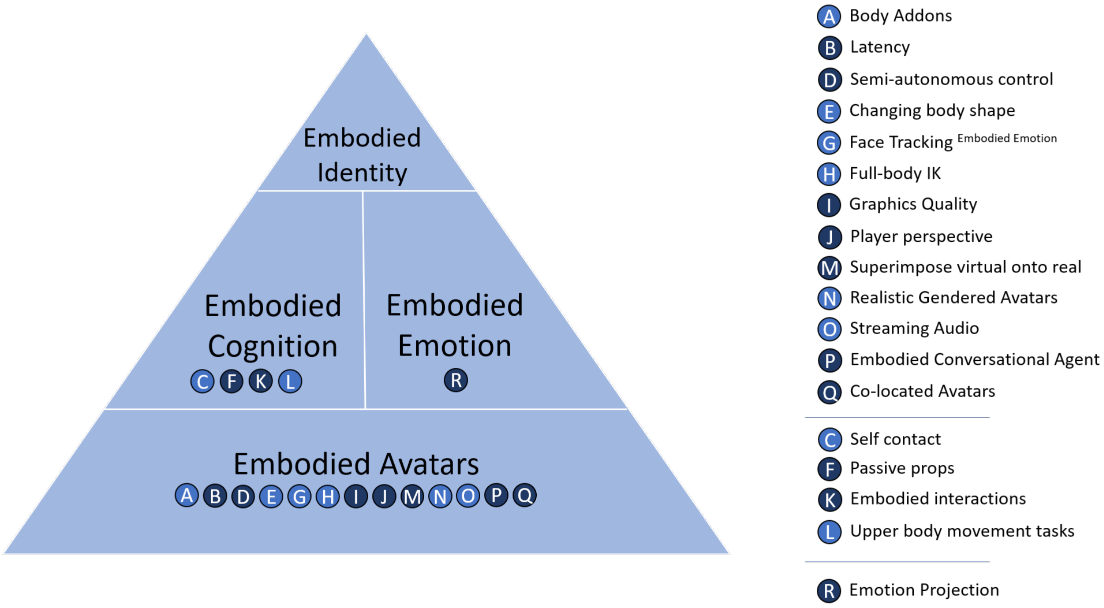

FIGURE 4.2: A pyramid diagram, showing how each level builds on top of the previous ones, together with the strategies used to affect that level from the literature. Dark blue means that the strategy wasn't mentioned in the literature for that specific level but was associated with it during the analysis. Superscript means that the strategy was also found at that specific level.

Embodied Avatars is the base level, where you feel like you have the physical extent and position of the avatar. Then, *Embodied Cognition* and *Embodied Emotion* are the next level, where you start thinking about your interactions as if you were your avatar and experiencing the avatar's emotions as your own. One of the strategies for *Embodied Emotion*, emotion projection, was originally not part of the strategies found. However, the way it was discussed in the paper by Zhao et al. [S59] was both as a type of embodiment, and as a strategy. Essentially, it involves the idea that you can experience emotional embodiment, and to achieve this state, you can use a technique called "projection", where your avatar mirrors and embodies your emotional experiences.

Embodied Identity is at the top of this model and represents the point where you begin to identify with the entity you are embodying. The levels thus reveal how 'invested' you are in the embodiment (are you just embodying the avatar on a surface level, still being yourself, or are you acting and thinking like the avatar). As an example, imagine a game where you play as a tiny mouse:

- at the first level, *Embodied Avatar*, you know you're the size and position of the mouse, you see your mouse body, and accept that you are controlling it;
- on the next level (*Embodied Cognition*), you start thinking about the world as if you were a tiny mouse, in your thinking you take into account how small you are, how fast you are, you see a hole in the wall and you think of going into it;

- adjacent to this is *Emotion Embodiment* — if you see a big cat around the corner, you would feel fear, as if you were a tiny mouse;
- finally, with *Embodied Identity* these all combine together, and you begin to think about yourself as the mouse.

The Aspects and Levels diagrams are two orthogonal ways of looking at embodiment. The Aspects emphasise the different ways you can feel embodied, while the Levels reveal the quality of that embodiment. This is a really helpful way of thinking about embodiment because it gives you a language to either deconstruct the experience of embodiment or look at it more holistically. It also provides a framework for explaining the third theme.

4.3.4 Higher Level Combinations

Some of the codes did not fit into these two models. These codes, after further analysis, were found to be representing a specific combination of both the Aspects and Levels. For example, there is *Gender Embodiment*, which is generalised as **Anthropomorphic Embodiment**, as it is about embodying a human-like avatar (examples would be race, size, or age embodiment). This would be using the Body Ownership aspect typically at the level of Embodied Cognition or Emotion. Whereas *Animal Embodiment* (or **Non-anthropomorphic Embodiment**, which is embodying a non-human-like avatar) combines Body Ownership as well as Agency (because animals have different abilities like waving a tail or flapping wings) typically at the level of Embodied Cognition and Emotion. Depending on the circumstance, Anthropomorphic Embodiment could also include Agency, contingent on whether the avatar has the same abilities as your body or not.

Perspective Simulation (*Embodied Simulations*) is about embodying a specific person or 'thing' (this is Body Ownership at the Embodied Identity level).

Enhanced Remote Embodiment (*Embodiment enhancement*) generally is about embodying an avatar that has a physical counterpart. For example, a robot that you control and has its hands controlled by yours. Interestingly, here Self-Location would play an important role because you would need to feel like you're in the same space as another physical machine. You would also need Body Ownership, and Agency, all at the Embodied Avatars level at the least.

Tool Embodiment (*Tool Embodiment*), which is when a tool you're using becomes part of you [S2], was combined together with *Technological Embodiment* in the final analysis, as Technological Embodiment is just a specialised version of tool embodiment. In this case, for Tool Embodiment to be present, you need Body Ownership, as well as

Agency, and the tool needs to become like a part of you and for you to think about solving problems using its affordances (which is Embodied Cognition).

4.3.5 Strategies

There were 20 strategies identified in the coding process, and it should be noted that not all embodiment codes had a strategy to go with them. For example, one of the aspects (Self-Location) had no strategies specifically tied to it; there were strategies that affect it together with the other 3 aspects, but not ones that target it directly. This might be explained by the fact that, in the case of VR, Self-Location is such a big part of wearing the HMD that it is a given, as one is self-located by the nature of the technology.

Virtual-to-Physical Sensory Transfer and *Physical-to-Virtual Sensory Transfer* are a "technique used in virtual reality to create the illusion of becoming a virtual avatar" [S14]. This set of strategies broadly talks about the way that you can feel grounded in the virtual world through the help of the physical world, and vice-versa. Sensory transfer supports Body Ownership, as well as Virtual-to-Physical transfer, whereas Physical-to-Virtual is more about Self-Location (you touch something in the physical world that corresponds to a virtual object, and you feel more self-located).

For the Levels diagram (Figure 4.2), each strategy contributes to the levels above it, as they build upon the previous level. Figure 4.1 shows the strategies mapped into the Venn diagram (as explained in the Embodiment Aspects subsection), and one can see that a large number fall within the Agency, Body Ownership and Self-Location circles, as a lot of the strategies were talked about more in a very general sense.

The strategies that also target more specific aspects can be seen in Figure 4.1 with a superscript next to them. For example, Upper body movement tasks was tested against the 3 components as a whole [S10] (as it is in the intersection of the Self-location, Agency and Body Ownership circles) or was looked at when talking about Agency only [S23].

All of the mappings between Embodiment and Strategy can be seen in Table 4.4, where strategies are shown on each row, and embodiments are on each column, their intersection showing whether there is a mapping between them (a paper has used that strategy for the embodiment) with the number representing how many papers have that mapping. The embodiment codes that had no strategies associated with them were left out of the table.

TABLE 4.4: Non-empty embodiment codes (divided into the Aspects, Levels and Higher Level Combinations (HLC)) and strategies to affect them, with the number of occurrences of each pair in the intersection.

	Aspects					Levels			HLC		
Strategies / Embodiments	Embodiment - A/BO/S-L	Embodiment - BO Component	Embodiment - A Component	Embodiment - Generic	Embodied Social Virtual Reality (SVR)	Embodied Avatars	Embodied Cognition	Emotion Embodiment	Tool embodiment	Embodied Simulations	Embodiment enhancement
Full-body IK	3	5	2			3					
Embodied Interactions	1			1			8		1		
Gendered avatar through mirror	3	2				1					
Face tracking	1					3		1			
1st or 3rd Person Perspective	2	1	1								
Latency		1	2								
Upper body tasks	1		2								
Changing head size						2					
Self contact in VR		1	1								
Semi-autonomous control		1	1								
Body add-ons			1			1					
Streaming audio				1		1					
Superimposing Virtual onto Real		1		1							
Low-fidelity graphics	1										
Different HMDs	1										
Embodied Conversational Agent				1							
Co-located avatars					1						
Mimicking a performer										1	
Camera on remote robot											1
Passive props				1							

4.4 Discussion

The analysis of the systematic review data uncovered two orthogonal models of embodiment (Aspects and Levels), a set of Higher Level Combinations (which were found to be specific types of embodiment that can be cross-referenced with the two models), and finally, a set of strategies that map onto the Aspects and Levels.

4.4.1 Aspects of Embodiment

The Aspects model extends Kilteni et al.'s model of embodiment to include Social Embodiment, the extent to which a person feels embodied in a social situation with another person. Along with the existing three (Self-Location, Body Ownership, and Agency) this model reflects the different sides of embodiment which can be considered independently when designing and thinking about a multi-modal experience.

As an example of how the social embodiment aspect is important, it's worth thinking of examples where it functions in isolation. Consider a phone call or a video conference call where participants don't have avatars or virtual bodies. The participants don't feel self-located, they don't have body ownership, and have no agency through an avatar; but they do feel that they are embodied in a social setting with the person that they are talking with: they are socially embodied.

4.4.2 Levels of Embodiment

The levels model emerged from the observation that Embodied Cognition lies within a broader set of other types of embodiment, and that there is a hierarchy within that set. It combines *Embodied Avatars*, *Embodied Cognition*, and *Embodied Emotion* onto a hierarchical structure, which culminates with *Embodied Identity*. This term, *Embodied Identity*, was not found in the literature and was an addition made in the analysis, representing the culmination and next logical step of all of the levels below. These four levels describe generally the 'quality' of your embodiment, or how deeply embodied you are.

They show a much richer picture of embodiment, as sometimes you don't have a strong level of embodied cognition, but you can still talk about embodying an avatar, and how you can control it like it's a part of you. Control of an avatar doesn't necessarily mean thinking about solving problems with it as you would normally think to do with your own body, but you can slowly transition between the two levels, from embodying an avatar to a deeper embodiment, where you start thinking with your avatar's body attributes. You could also be emotionally embodied within an avatar, but again, not think about the world and the problems you can solve using that body. For you to be at the top level, *Embodied Identity*, you need to have all the previous levels.

4.4.3 Higher Level Combinations

In the coding process, mentions of very specific types of embodiment were found (such as Gender Embodiment and Animal Embodiment) that could be explained by cross-referencing the two orthogonal models created. This means that some very specific types of embodiment discussed in the literature can be deconstructed only by using both models, which also suggests that some of the strategies that apply to the models could also apply to those specific types of embodiment. An example would be if you want to affect Gender embodiment, first you'd need to look at what it's made of. In this case, you can look at it as Body Ownership and Avatar embodiment at the most basic level, and thus, you can use a strategy targetting Body Ownership or Avatar embodiment to hopefully improve it.

4.4.4 Strategies

Finally, a wide set of 20 strategies were identified and mapped onto the two models, showing that they are orthogonal, and laying a basis that future researchers could use when designing an experience for multi-modal spaces. Those strategies provide an easy way for someone to affect a specific type of embodiment, depending on the context they are working in. For example, if you want to create a multi-modal VR experience where one person is in VR and the other one is interacting with them outside, you might find that you have two different strategies for the two modes that affect the same type of embodiment, or the same strategy could affect the two participants' embodiment differently.

4.5 Summary

Embodiment is an innate feeling that every human experiences in their daily life. Trying to foster that feeling can help create more immersive and enjoyable experiences in multi-modal play, but in order to understand multi-user multi-modal interactions and experiences, we need a more granular view of embodiment.

Having conducted a systematic review of the embodiment space in VR and multi-modality, two orthogonal models were put forth that emerged from the inductive coding. The first model, the "Aspects", uses terminology from a previous model of Embodiment as it was evident early on in the coding process that it is an important and useful model, but also builds on top of it, as it was insufficient in talking about multi-modal social experiences. Thus, the "Social Embodiment" aspect was added alongside the other three. The second model, the "Levels", was based on the idea of a hierarchical structure of embodiment, where the more embodied you are, the higher on the pyramid you are, culminating in "Embodied identity". The "Aspects" roughly can be viewed as the way you are embodied, or in other words, how you are embodied (only feeling self-location, or body ownership and agency, etc.), while the “Levels” is more about the depth of embodiment, where you can feel just like you’re controlling an avatar or feeling the avatar as being you.

A number of specific types of embodiment were found throughout the literature and were noted that they could be viewed as combinations of the Aspects and Levels, which then helps people more easily try to affect them through the strategies provided.

This systematic review started as a way to gain insight into what embodiment is, and how it can be affected. With two models created, the first experiment’s results can be reanalysed though them. This approach should uncover fresh perspectives and gain a deeper understanding of the overall experience. It could also potentially lead to a number of strategies that can be implemented in the game, which would both help validate the models (helping with the analysis of the experience in a different way), and see if the strategies found work in improving the embodiment. The latter is important, as one of the theories that came up from the analysis of the first experiment’s data was that targeting embodiment was a good way to influence immersion and improve the overall experience. An experiment to test this theory would need to be able to compare the multi-modal experience with and without the use of a specific strategy and show that there is a difference between them in some way. The following chapter is going to delve into the design and execution of that experiment, together with the discussion and implications of the results.

Chapter 5

Application of the Embodiment Model and Strategies

In the previous chapter, a number of different strategies were identified that can affect specific types of embodiment. These findings, combined with the results from the first experiment, led to the testing of certain strategies within "StuckInSpace", aiming to evaluate their effectiveness and independence. This chapter addresses Research Question 3 (RQ3): "How can design strategies impact the different modes of play in terms of Embodiment?".

5.1 Example application of the models

To determine the changes required, first it is necessary to explore the application of the model to the initial study, and then use this information to decide on design strategies for implementation and testing in a second experimental study.

One of the key findings from the first study was that embodiment plays a big role in the multi-modal experience. Improving embodiment in one mode would translate to improved immersion/co-presence in that mode, subsequently affecting other modes positively. In order to understand how the embodiment in the first experiment aligns with the models, participants' comments need to be examined.

5.1.1 Seeing/hearing the other and bumping into the other/the environment

One observation was how seeing/hearing the other person, and bumping into them or the environment produced conflicting responses.

Some examples of negative VR comments include difficulties in locating the drone, not hearing the drone or the second player, and bumping into objects in the real world (whether another player something else). Interestingly, some of these negatives were positives as well, depending on the context. For example, bumping the other person and seeing them was perceived positively if the real and virtual world were aligned, but negatively if they were not. Comments such as: **"I bumped you and then I was more aware that you're there"** or **"With the VR, I wasn't very aware of the (real) world, I only sensed the cable"** indicate that accidental bumping takes participants out of their experience, which is a negative for immersion, but when it's the other participant this is a positive for co-presence.

Referring to Table 4.2, this issue could be classified under *Self-Location* in the model. When participants in VR bump into something, they are reminded of the real world, which makes them less self-located in the virtual one. These observations also align with the concept of "breaks in presence" (BIP), as discussed in the background section (subsubsection 2.5.1.3). BIP occurs when mismatches between the real and virtual environments disrupt PI, reducing immersion. However, as shown in participant comments ("I bumped you and then I was more aware that you're there"), BIP can also positively impact co-presence by reinforcing shared physical space.

It can be seen that there's a problem between the players as well. The 'other' plays a significant part in some of these comments, where some people would feel less alone when they saw the avatar in VR, knowing there was a real person present (in the case of the HMD^{Phone} player). This suggests viewing it as a *social embodiment* factor as well, which can help narrow down any strategies to mitigate it as a factor.

Recognising that the problem with *seeing/hearing the other and bumping into the other/the environment* can be deconstructed into two main embodiment aspects, strategies affecting both aspects can be explored. As shown in Figure 5.1, two strategies affect Self-Location and Social Embodiment at the same time: Streaming Audio and Embodied Conversational Agents.

As some of the comments were about not hearing the other player (the drone), or not hearing them from the drone's location (for the PC player), Streaming Audio seems like a better choice to be explored. It is important to not only choose a strategy that affects the wanted types of embodiment, but also one that makes sense in the context in which it is being introduced in. This strategy could be applied by recording the audio from a microphone on the Phone or PC and then streaming it onto a headset and played in the 3D space, helping the VR player orient themselves in terms of the second player, thereby making them feel more self-located and socially embodied.

Additionally, the second strategy involves using a human-like avatar that helps with embodiment. It is important to show that a strategy can work and be adapted to fit the context of the experience. For example, using a different avatar for the second player

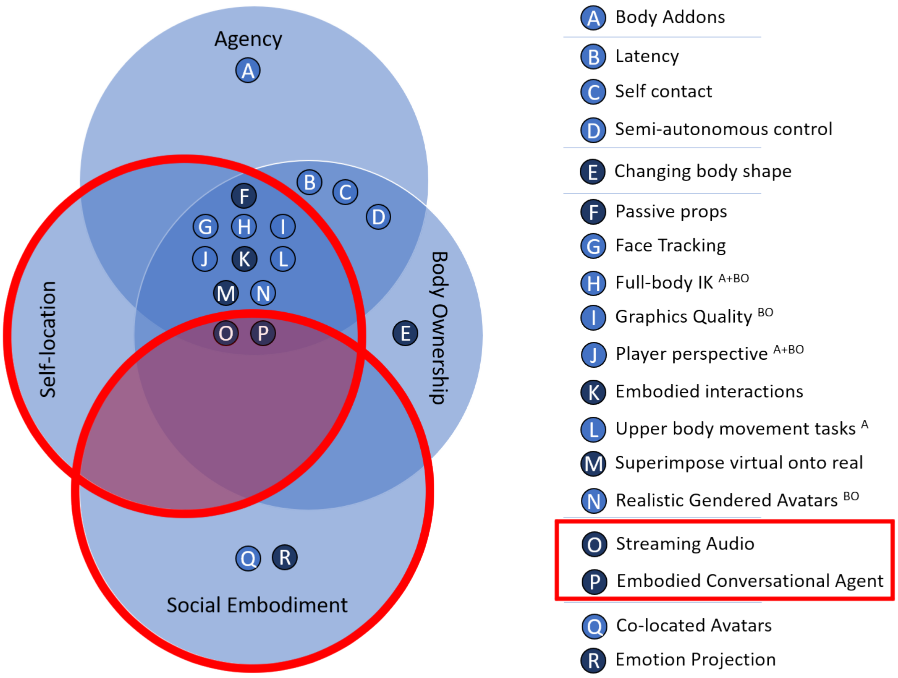

FIGURE 5.1: Highlighting the 2 aspects and the strategies in them

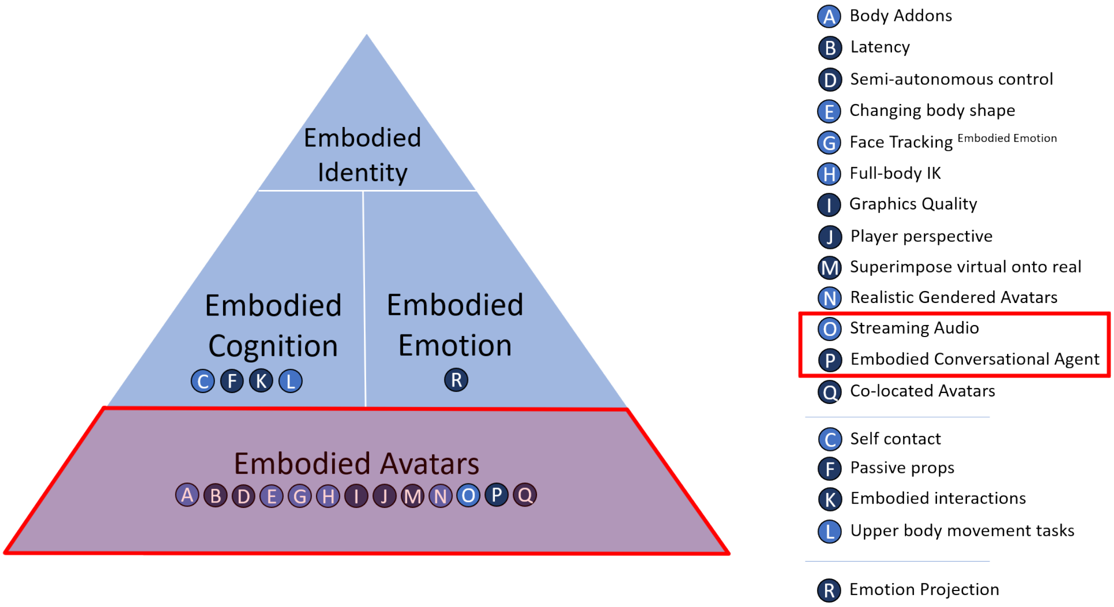

FIGURE 5.2: Highlighting the 2 strategies at the Embodied Avatars level

(instead of a drone) or incorporating a hologram of a human avatar above the drone could be considered.

Having selected those two strategies, looking at Figure 5.2, it is clear that they affect embodiment at the base level. Looking at the levels can help take the embodiment to a higher level, for example to Embodied Emotion, where emotion projection could be used to try and increase the impact of the audio and avatar strategies.

5.1.2 How the 'other' affects your own embodiment

Another interesting observation was that several comments were not necessarily about a participant's own sense of embodiment, but rather the embodiment of the other person, and how that, in turn, influenced their own sense of embodiment. This further shows how the social aspect is very important in such multi-modal contexts.

Consider this comment: **"When the other player is pushing the buttons wrongly, it makes me realise that 'wait, this is actually not the real world', because if it was an actual one they would have done it with no problem"**. This shows that one type of embodiment in one modality can affect a different type of embodiment in another. In this instance, the Agency of the VR player (ability to act) negatively affects the self-location of the PC player (who feels less self-located in the virtual world) and social embodiment (as it is talking about the other player's influence).

To improve the agency of the VR player, the Full-body IK strategy in the aspects model can be considered. It is indicated with a superscript letter A (Agency), which means that it was also used to affect only agency in one of the papers reviewed. That strategy might consist of making the VR player have more control over their body, making it easier to interact with the virtual world. One should note that having a strategy in a single aspect doesn't necessarily mean that it won't affect any other; it simply means the review didn't find any papers using that same strategy for another type of embodiment. Therefore, if there's a good reason to believe a certain strategy can help, it is worth trying.

Another way to look at how the other affects one's embodiment is by looking at the comments about the non-HMD players being distracted by the HMD player, especially on the Phone version (due to the small screen and movement around). There are numerous strategies affecting self-location, as seen in Figure 5.3. For example, the "Embodied Interactions" strategy can be used. Currently, both the Phone and PC users are not fully utilising their devices. On the Phone version, the game doesn't use the touch screen, which could be used to help the player focus on what is happening in the game using embodied interactions. Similarly, on the PC version, the mouse isn't used at all, despite being a vital part of using a computer. Implementing

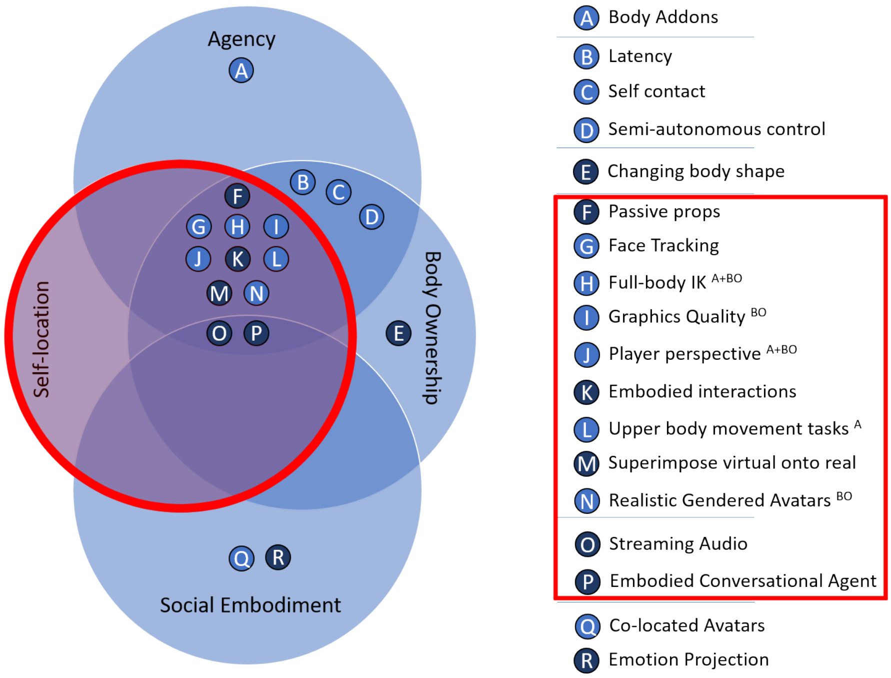

FIGURE 5.3: Highlighting the strategies in the "Self-location" Aspect

the mouse could enhance user embodiment as most people are accustomed to using the mouse and keyboard together.

5.2 Changes made

After analysis of the data and comments given by participants, several possible strategies for implementation were found. In order to decide what strategy to implement, Figure 5.1 and Figure 5.3 can be used as guides.

The following subsections discuss the rationale behind choosing the strategies and their subsequent implementation.

5.2.1 Chosen strategies

When choosing the strategies, their viability and ease of implementation was taken into account, as there was a limited amount of time to implement and test them. Based on the analysis in subsection 5.1.1 and Figure 5.1, one of the main strategies identified was **"Streaming Audio"**. This strategy involves recording the voice of the non-HMD player and playing it back through the headphones of the HMD player (and vice versa). Additionally, using the properties of the virtual 3D space to spatially position the audio would facilitate easier location of the other player. It is also relatively simple to implement, as the existing data transfer used by the game can be leveraged to include the audio stream from each of the devices.

The other main observation was how the other affects one's own sense of embodiment and the potential distraction caused by that, mainly noted by the non-HMD players. As discussed in subsection 5.1.2, "Self-location" is being affected by the HMD player, as the non-HMD player found it distracting at times. Looking at the different strategies in Figure 5.3, the "Embodied Interactions" strategy stood out as both effective and straightforward to implement. An example of such an interaction would be using the full affordances of a device, such as the touch screen on a phone, or the mouse on the computer. Familiar interactions like "Pinch to zoom in/out" on a phone or "Scroll wheel to zoom in/out" on a computer are considered embodied, as the average user instinctively understands how to perform them without thinking about it.

5.2.2 Implementation of the strategies — Streaming Audio

In order to implement the first change, 'Streaming Audio', the HMD's built-in microphone was chosen, as well as a headset with a microphone for the non-HMD player. The implementation consists of the microphones picking up the voice of the player and transmitting it through to the other player's headphones. The way that is done depends on the version:

- VR/PC version: A separate audio engine in Unity was used, making it possible to record and transmit audio to multiple outputs (the VR and PC headsets). As it was all running on the same machine, it was just a matter of recording the sound of one player and outputting it to the headset of the other;
- VR/Phone version: This version was more challenging due to running two separate versions of the game on two devices, needing communication between them. As the devices were already using the WiFi to connect and transmit the location and status of the game to the Phone player, adding an audio stream was relatively straightforward.

Although fairly simple in theory, in practice there were a few inherent problems with the system. In the VR/Phone version, there was a delay between the recording, transmitting, and replaying the audio, which was difficult to minimise due to technological limitations. This means that there was always a small delay in the audio, for example, when the VR player spoke a word that was heard in the headphones of the Phone player. This issue was mitigated by using a lower sampling rate and transmitting as little data as possible to make it faster, drawing inspiration from VoIP and similar technologies. The PC version didn't have that problem, as all the audio computation was done on a single machine with minimal delay. Another minor issue with the Phone version was that, as both players were close to each other, one microphone could pick up the speech from the other person, thus creating an echo.

The delay achieved was between 100ms and 300ms. Interestingly, this was narratively explained in the game to help people understand why it was happening.

5.2.3 Implementation of the strategies — Embodied Interactions

For the ‘Embodied Interactions’, careful consideration was made to implementing them and fitting them narratively in the game. Ultimately, a zoom-in and zoom-out function was added to both versions of the game, and a task was changed to accommodate the feature. The keypad action was identified as the most suitable task, as the current version had the non-HMD player read a code out to the HMD player, the code being on their screen as a subtitle. The decision was made to display the code on a faraway satellite outside the play area, forcing the non-HMD player to use the zoom function to read the code and communicate it to the other player.

A major modification from the first experiment was the ‘breaking’ or ‘fixing’ of the drone, where the drone would start either broken or functional, and in the second part of the game, be repaired or malfunction. This allowed for the creation of a scenario where the strategy was both absent and present within the same play session. For ‘Streaming audio’, the voice would either work or not, but the zooming task proved to be more complex to implement, as the whole task was about zooming, so having the ability to zoom broken would change the task completely.

The final decision was to have two ways to zoom: one using embodied interactions, and one using less embodied interactions. When functional, the ‘optical’ zoom could be used via fingers or mouse scroll wheel. When it malfunctioned, only ‘digital’ zoom was available, requiring the player to use an interface and have the zoom level typed on a keypad, as seen in Figure 5.4. This way provided a highly embodied zooming method together with one that detracted from the immersive experience.

5.2.3.1 Phone Holder

In order for the zoom strategy to work well on the phone, modifications to the way the phone was used together with the controller were necessary. In the first experiment, people held onto the phone and controller separately (Figure 3.4 in Chapter 3), making it impossible for someone to also use the touchscreen on the phone in order to zoom in and out. To fix the problem, a 3D-printed phone holder was designed that fits into the controller, and a phone mount attached to securely hold the phone. Images of

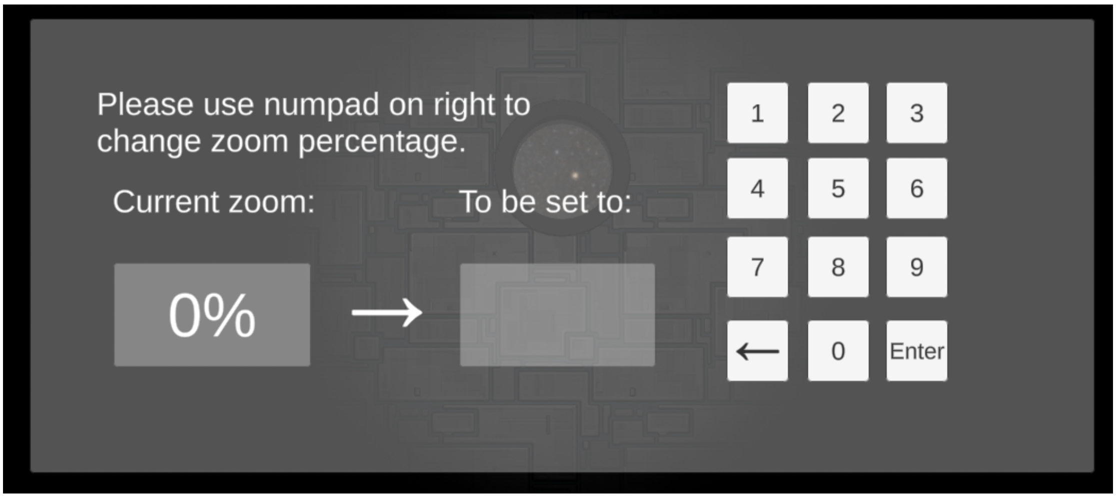

FIGURE 5.4: The zoom interface, as seen on the PC version. You have to type out the zoom percentage with your mouse, clicking on the numbers, which is supposed to be less embodied than just using the mouse wheel. The Phone version looked similar.

FIGURE 5.5: The controller, the 3D printed holder, and a phone mount

the setup can be seen in Figure 5.5 and Figure 5.6. Using this holder, people can use only one hand for the controller, and the other was free to use the touchscreen.

5.3 Methodology

The overall methodology closely resembles the first experiment's methodology, as the same game is being tested with some changes the underlying game mechanics.

Initially, a decision was required regarding which questionnaires to use and whether the same ones as the first experiment should be used. The Software Usability Scale (SUS) was the first to be discarded, as the game's usability was already shown. Since

FIGURE 5.6: The phone attached to the mount, showing a screenshot of the game

the implemented strategies were neither major or game-changing, and the main idea of the game remained the same, the other questionnaires seemed more important. The two other questionnaires, the NMMoSP and the IPQ, centred more around Immersion, Presence and Co-Presence. There was a need to find a different approach to evaluate embodiment, specifically the four different aspects within the model.

Ultimately, a version of the proposed Embodiment Questionnaire by Gonzalez-Franco and Peck was selected, as it had questions relevant to the 3 main Aspects identified by Kilteni et al. Some modifications were necessary to tailor the questions to the specific experiment [32]. For example, one of the questions *"I felt as if the virtual ___ was my ___"*, was changed by replacing the blanks with "virtual avatar (astronaut)/body" and "drone/body" respectively for the different versions of the game.

A selection of questions for Body Ownership, Agency, and Location of the body were chosen, as the other categories didn't apply to my experiment ("external appearance" or "tactile sensations" for example). As Gonzalez-Franco and Peck suggest, "some subsets of questions might not be applicable to some experiments" [32], as well as the fact that an experiment might need to include more questions.

The final part was the need for a "Social Embodiment" category, which Gonzalez-Franco and Peck did not include. Certain questions from the NMMoSP had a very strong connection to the social aspect and to embodiment, leading to their selection as the basis of this fourth aspect. Statements like *"I noticed the other player."*, *"The other player noticed me."*, *"I was easily distracted from the other player when other things were going on."* and etc. were part of the final category.

With these 4 categories, the final questionnaire (Appendix D) had 3 questions about Body Ownership, 4 questions about Agency, 2 about Self-Location, and 8 about Social Embodiment. Care was taken to avoid making the questionnaire too long, as participants would have to complete it 4 times in total. If the questionnaire was too long, tedium could have been introduced, potentially skewing the results. The questions themselves used a seven-point Likert scale identical to the first experiment, ranging from 1 (strongly disagree) to 7 (strongly agree).

Basic demographic questions, such as age group, gender, and any previous VR experience, were also included.

5.3.1 Study Design

As mentioned in precious sections, it was decided to change the strategy state in the middle of the game. This approach allowed for the questionnaire to ask about that change and the difference between the two parts of the game. Then, the roles of the players were swapped, mirroring the first experiment. Participants then played the game again and completed the same questionnaire, followed by a final semi-structured interview.

In the end, this was a mixed design study similar to the first experiment. However, within a single experiment, participants would experience the presence and absence of a strategy. As a result, the questionnaires had to be specifically tailored to include that. Each questionnaire was actually made of two identical sets of questions, the only difference being that the first set is for the start of the game, while the second set addressed the strategy change (Appendix D).

An important aspect of the game progression, which later had a surprising impact of the result, must be highlighted. It was needed to make the change of the state of strategy very clear to the players to facilitate pointing to that change without explicitly mentioning the strategy itself, potentially skewing the results (for example, the difference between *"Please answer the questions for when the zoom functionality was broken"* versus *"Please answer the questions for when the lights changed"*). At the same time, the change needed to be unnoticeable during play, so that people don't suddenly realise something experimental had changed and break their whole immersion and experience. This led to the use of an event in the middle of the game where the emergency lights went on, and everything became red, as the trigger for the strategy change. The change was explained narratively as being either magically "fixed" by the scientists on earth or "broken" because of the connection. This approach minimised the disruption and made it so people think about the whole experience when answering the questionnaire, rather than a single instance of something breaking.

The questionnaires themselves had in bold the instruction to consider the answers of the following questions where **"the emergency lights were off"** or **"the emergency lights were on"**. This was further emphasised when giving the questionnaires to make sure that participants understood the distinction.

A change between the first study and this one was the decision to record the players using a camera in the corner, which would enable analysis of any interactions they had afterwards, such as any time they collided by mistake.

5.3.2 Ethical Considerations

Ethical approval was extended from the first experiment by the University of Southampton's Ethics board, under the reference number ERGO 54456.A2. This approval ensured that the research was conducted in accordance with the ethical standards of the university.

All data collected during the study was anonymised and securely stored on a password protected university PC behind a staff-only access-controlled area to prevent unauthorised access in the same way as the first experiment. The VR game was designed to minimise any discomfort or adverse effects to participants, and participants were told beforehand that they can end the experiment early in case they start feeling nausea.

5.3.3 Setup and Procedure

In summary, a single session with two players went like this:

- The two players would play a game in one of the modalities (VR/PC or VR/Phone) and were semi-randomly chosen to either have the strategy start off being fixed or broken (calculated to ensure equal distribution by the end);
- Halfway through the game, the strategy state would flip (a broken zoom would get fixed, or a working zoom would break, etc.), signalled by the change of lights to make it more memorable to the player;
- After finishing the game, each participant would sit separately and be given a questionnaire comprised of demographic questions, and two identical questionnaires for when the emergency lights were off and on (Appendix D);
- People would swap roles, play the same game, with the only variable change being the modality (the HMD player is now the non-HMD player and vice-versa);

- The game follows the same flow, the questionnaires are identical to the first ones, but this time for the second experience (Appendix D);
- Finally, a semi-structured interview with both participants together was done, in which a number of general questions about the experience were included, as well as more focused ones on the strategy that was changed, for example, *"When the emergency lights went on, did you notice the other player taking longer/shorter?"* (Appendix D).

The same hardware and space as the first study was utilised, with the notable exception for new headphones for the non-HMD player in the Streaming Audio version, as they needed headphones to be able to hear and talk with the other player virtually, as well as the phone mount holder for the Phone player.

5.4 Results and Analysis

In total, 48 people participated in the study, with 36 male and 12 female participants, with the average age being 24.04 years. The majority of players had minimal first-hand experience using VR (34 respondents said they played VR never, rarely, or a few times an year), while 9 people used VR either monthly or weekly. Five people hadn't answered the question. On average, there were a lot more students who tried the experiment than university staff.

From the 48 people, 24 participated in the Embodied Interaction strategy, and the rest 24 participated in the Streaming Audio strategy. Each participant would fill in 4 questionnaires in total: 2 for the first playthrough, with and without the strategy, and 2 for the second playthrough, with and without the strategy. This resulted in a total of 192 different questionnaire results, divided in a number of ways. As in the first study, the categories include HMD^{All}, HMD^{PC}, HMD^{Phone}, as well as the Phone and PC categories. However, the results are now further divided based on whether the strategy was "Present" (P) or "Absent" (A), and the order in which it went through (either started Present and then was removed or started Absent and got back in the middle of the game).

For one group of participants, 2 people participated, which generated 4 VR data points (2 from the first person, and 2 from the second), and 4 non-HMD data points. In the following subsection, the unexpected quantitative results and the qualitative results that may help explain these findings will be explored.

5.4.1 Quantitative Results

Scores for both experimental strategies are presented in Table 5.1 and Table 5.5. Each table displays the results of the four aspects and their aggregated scores, categorised by the state of the game for which the questionnaire was conducted. ’P’ denotes Present, ’A’ denotes Absent, with 1 and 2 being the phases of the game (start and middle, respectively). For example, P1 represents the results where the strategy was Present at the start, while A2 indicates it was Absent in the second part. The Present and Absent states are also included separately to highlight any general differences. The number of data points is indicated in parentheses in the Mode column.

The aim of the quantitative analysis is to assess whether the strategy impacted the modes (positively or negatively) and if any effects were transferred to another mode (e.g., whether the Phone player feeling more embodied influenced the VR player’s sense of embodiment). The following subsections provide a detailed analysis of the two strategies and their results.

5.4.1.1 Embodied Interactions

A linear mixed-effects model analysis was initially performed. The results, as shown in Table 5.2, indicate Mode being the most significant factor affecting scores, which is expected and is in line with the first experiment’s results. The four-way interaction between Mode:Playthrough:Part:Strategy is also significant, indicating that the combination of these factors influence the score in a complex manner.

Following the results from the ANOVA, a subsequent EMMeans post-hoc test was done on Mode, the only significant effect, the detailed results of which are shown in Table 5.3 and Table 5.4. Table 5.4 does show a number of significant differences between the modes, mainly between the PC and VR versions.

The analysis did reveal a significant difference between the HMD and non-HMD (PC and Phone) modes (Table 5.2), which aligns with expectations, as VR is inherently more immersive and embodied than non-VR mediums.

Unexpectedly, no significant difference was found between the Phone and VR versions across most conditions, as depicted in Figure 5.7 and further detailed in Table 5.4. This contrasts with the first experiment, which showed a significant difference. This suggests that the current setup has enhanced the embodiment in the Phone mode to be more comparable to the VR version, as can be more easily seen in the scatter plot (Figure 5.7).

TABLE 5.1: Mean Scores and Standard Deviations for **Embodied Interaction** Strategy by Mode and State. The total number of data points for each mode is: HMD = 24, PC = 12, Phone = 12. Within each mode, data points are divided equally between the states P1, A1, P2, and A2.

Mode	State	Body Ownership M (SD)	Self-Location M (SD)	Agency M (SD)	Social M (SD)	Aggregated M (SD)
HMD (24)	P1 (12)	6.22 (0.67)	6.50 (0.48)	5.96 (0.81)	5.94 (0.75)	6.08 (0.51)
	A1 (12)	5.64 (1.21)	5.96 (0.66)	5.33 (1.02)	5.18 (0.97)	5.41 (0.78)
	P2 (12)	5.53 (1.43)	5.63 (1.17)	5.44 (0.96)	4.81 (0.90)	5.17 (0.83)
	A2 (12)	6.00 (1.05)	6.71 (0.40)	5.85 (1.12)	5.48 (1.04)	5.83 (0.64)
	P (24)	5.88 (1.15)	6.06 (0.98)	5.70 (0.90)	5.38 (0.99)	5.63 (0.82)
	A (24)	5.82 (1.12)	6.33 (0.65)	5.59 (1.08)	5.33 (0.99)	5.62 (0.73)
PC (12)	P1 (6)	4.00 (1.28)	5.08 (1.24)	4.33 (1.41)	5.42 (1.07)	4.94 (0.99)
	A1 (6)	3.89 (1.46)	3.33 (1.37)	3.75 (0.94)	5.35 (1.06)	4.50 (0.74)
	P2 (6)	4.00 (1.40)	2.75 (1.21)	3.67 (1.28)	5.04 (1.00)	4.26 (0.86)
	A2 (6)	4.61 (0.83)	5.00 (1.14)	4.50 (1.13)	4.99 (0.94)	4.84 (0.83)
	P (12)	4.00 (1.28)	3.92 (1.69)	4.00 (1.33)	5.23 (1.01)	4.60 (0.95)
	A (12)	4.25 (1.19)	4.17 (1.48)	4.13 (1.06)	5.17 (0.97)	4.67 (0.77)
Phone (12)	P1 (6)	5.72 (1.06)	5.08 (1.53)	6.00 (0.84)	6.06 (0.61)	5.83 (0.58)
	A1 (6)	4.89 (1.03)	4.83 (1.47)	5.71 (0.98)	5.06 (1.33)	5.10 (0.97)
	P2 (6)	5.00 (1.14)	5.42 (1.53)	5.92 (0.75)	5.29 (1.13)	5.37 (0.95)
	A2 (6)	5.50 (1.35)	4.92 (1.88)	6.13 (0.75)	5.72 (0.65)	5.62 (0.76)
	P (12)	5.36 (1.11)	5.25 (1.47)	5.96 (0.76)	5.67 (0.96)	5.60 (0.79)
	A (12)	5.19 (1.18)	4.88 (1.61)	5.92 (0.86)	5.39 (1.06)	5.36 (0.88)

TABLE 5.2: Type III ANOVA table with Satterthwaite's method (Embodied Interactions). Significant codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

Effect	Sum Sq	Mean Sq	F value	Pr(>F)
Mode	9.86	4.93	31.54	$<$ .001***
Playthrough	0.02	0.02	0.15	.699
Part	0.36	0.36	2.32	.134
Strategy	0.07	0.07	0.45	.506
Mode:Playthrough	0.10	0.05	0.31	.737
Mode:Part	0.30	0.15	0.96	.388
Playthrough:Part	0.01	0.01	0.09	.771
Mode:Strategy	0.33	0.17	1.06	.354
Playthrough:Strategy	0.08	0.08	0.50	.484
Part:Strategy	0.57	0.57	3.64	.071
Mode:Playthrough:Part	0.63	0.31	2.00	.146
Mode:Playthrough:Strategy	0.02	0.01	0.05	.947
Mode:Part:Strategy	0.29	0.14	0.93	.402
Playthrough:Part:Strategy	0.06	0.06	0.38	.543
Mode:Playthrough:Part:Strategy	1.25	0.63	4.01	.024*

TABLE 5.3: EMMeans for Mode by Playthrough, Part, and Strategy for Embodied Interaction.

Playthrough	Part	Strategy	Mode	Emmean	SE
First	Middle	Absent	PC	5.21	0.376
First	Middle	Absent	Phone	5.51	0.376
First	Middle	Absent	VR	5.62	0.322
Second	Middle	Absent	PC	4.87	0.376
Second	Middle	Absent	Phone	5.33	0.376
Second	Middle	Absent	VR	6.04	0.322
First	Start	Absent	PC	4.32	0.376
First	Start	Absent	Phone	4.68	0.376
First	Start	Absent	VR	5.45	0.322
Second	Start	Absent	PC	4.82	0.376
Second	Start	Absent	Phone	5.38	0.376
Second	Start	Absent	VR	5.38	0.322
First	Middle	Present	PC	4.39	0.376
First	Middle	Present	Phone	4.93	0.376
First	Middle	Present	VR	5.16	0.322
Second	Middle	Present	PC	4.26	0.376
Second	Middle	Present	Phone	5.67	0.376
Second	Middle	Present	VR	5.19	0.322
First	Start	Present	PC	5.12	0.376
First	Start	Present	Phone	5.78	0.376
First	Start	Present	VR	6.10	0.322
Second	Start	Present	PC	5.16	0.376
Second	Start	Present	Phone	5.48	0.376
Second	Start	Present	VR	6.06	0.322

TABLE 5.4: Post-hoc comparisons between Modes using the Kenward-Roger degrees-of-freedom method (EMMeans) for the Embodied Interaction Strategy. Significant codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

Playthrough	Part	Strategy	Contrast	Estimate	SE	t-ratio	p-value
First	Middle	Absent	PC - Phone	-0.30	0.388	-0.768	0.724
First	Middle	Absent	PC - VR	-0.42	0.300	-1.386	0.355
First	Middle	Absent	Phone - VR	-0.12	0.300	-0.391	0.919
Second	Middle	Absent	PC - Phone	-0.46	0.388	-1.188	0.465
Second	Middle	Absent	PC - VR	-1.17	0.300	-3.899	0.001***
Second	Middle	Absent	Phone - VR	-0.71	0.300	-2.360	0.056
First	Start	Absent	PC - Phone	-0.36	0.388	-0.926	0.627
First	Start	Absent	PC - VR	-1.13	0.300	-3.775	0.001***
First	Start	Absent	Phone - VR	-0.77	0.300	-2.576	0.034*
Second	Start	Absent	PC - Phone	-0.56	0.388	-1.447	0.324
Second	Start	Absent	PC - VR	-0.55	0.300	-1.841	0.166
Second	Start	Absent	Phone - VR	0.01	0.300	0.033	0.999
First	Middle	Present	PC - Phone	-0.54	0.388	-1.391	0.353
First	Middle	Present	PC - VR	-0.77	0.300	-2.554	0.035*
First	Middle	Present	Phone - VR	-0.23	0.300	-0.753	0.733
Second	Middle	Present	PC - Phone	-1.41	0.388	-3.630	0.002**
Second	Middle	Present	PC - VR	-0.93	0.300	-3.093	0.009**
Second	Middle	Present	Phone - VR	0.48	0.300	1.610	0.250
First	Start	Present	PC - Phone	-0.66	0.388	-1.698	0.215
First	Start	Present	PC - VR	-0.98	0.300	-3.279	0.005**
First	Start	Present	Phone - VR	-0.32	0.300	-1.079	0.531
Second	Start	Present	PC - Phone	-0.32	0.388	-0.830	0.686
Second	Start	Present	PC - VR	-0.90	0.300	-3.002	0.011*
Second	Start	Present	Phone - VR	-0.58	0.300	-1.927	0.141

These quantitative findings highlight the nuanced impacts of different modes under various conditions. The qualitative results will be essential for further explaining these patterns and understanding the underlying reasons for the observed differences.

5.4.1.2 Streaming Audio

A linear mixed-effects model analysis was also conducted for the Streaming Audio strategy. The ANOVA results (Table 5.6) indicated significant main effects for Mode and Playthrough, but no significant interactions involving Mode, Playthrough, Part, or Strategy together. This suggests that while the mode and playthrough influenced the scores, their interaction effects were not significant.

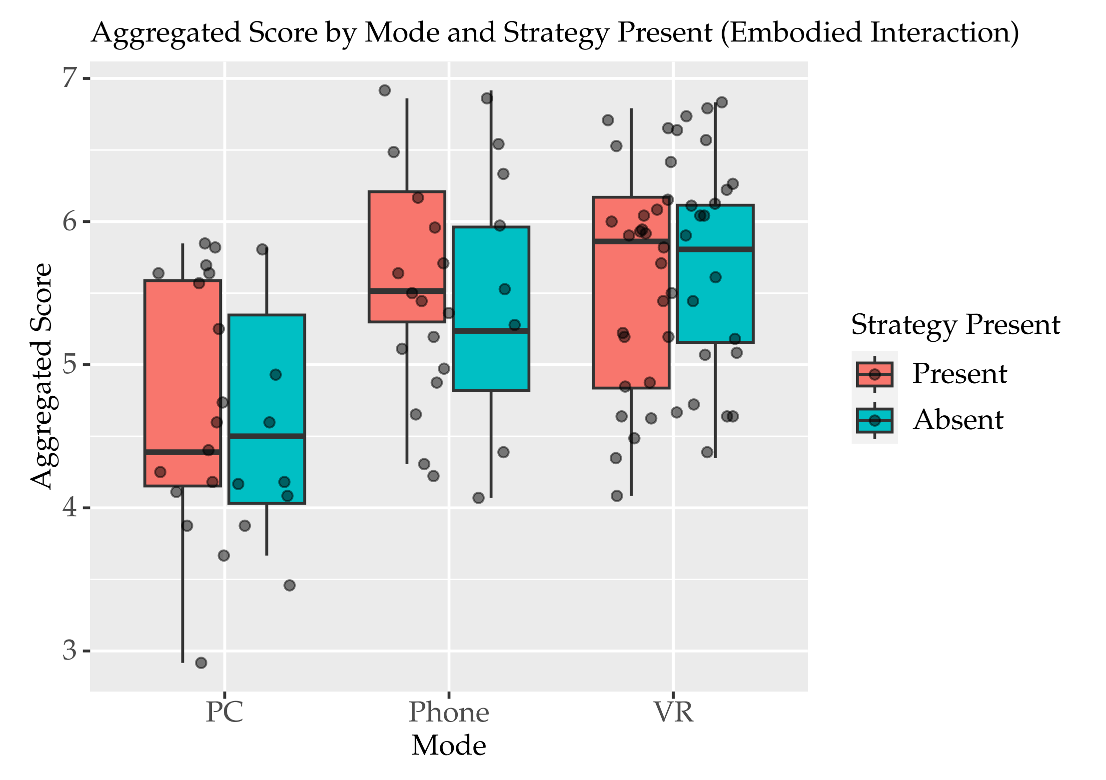

FIGURE 5.7: A scatter plot of the total embodiment scores of each of the modes (each dot represents one questionnaire).

Post-hoc analysis using EMMeans revealed significant differences between the VR and Phone modes under certain conditions, specifically considering the Playthrough and Strategy being present or not, as highlighted in Table 5.7 and Table 5.8. For instance, during the 'First' playthrough with the 'Start' part with Strategy 'Absent', the contrast between Phone and VR modes showed a significant difference ($p = 0.0361$). Similarly, during the 'First' playthrough with the 'Middle' part with the Strategy 'Present', the contrast between Phone and VR modes was also significant ($p = 0.005$). Those two results are connected because they are with the same modes (Phone-VR) and same playthrough (First), being both parts of the same game: the first one starting without the strategy, and in the middle of the same game, the strategy presented. This is inconclusive, as it can't be said with certainty why the results are as is. The expectation would have been for the part where the strategy is absent for there to be a difference, and when the strategy was presented for there to stop being a significant difference (which could indicate either the Phone rising in embodiment or the VR lowering).

The scatter plot (Figure 5.8) illustrates the distribution of the aggregated scores, clarifying the differences between the modes. Notably, while significant differences were found between the VR and Phone modes, there were no significant differences between the PC and Phone modes in most conditions.

Despite the lack of significant differences within the modes themselves across all conditions, these results underscore the complexity of the interplay between different modes and strategies in influencing players' experiences. The qualitative results will again provide essential context for these findings and offer a deeper understanding of the players' perspectives.

TABLE 5.5: Mean Scores and Standard Deviations for **Streaming Audio** Strategy by Mode and State. The total number of data points for each mode is: HMD = 24, PC = 12, Phone = 12. Within each mode, data points are divided equally between the states P1, A1, P2, and A2.

Mode	State	Body Ownership M (SD)	Self-Location M (SD)	Agency M (SD)	Social M (SD)	Aggregated M (SD)
HMD (24)	P1 (12)	5.83 (1.27)	6.04 (1.08)	5.48 (0.87)	4.77 (1.08)	5.28 (0.81)
	A1 (12)	5.31 (1.13)	5.58 (1.38)	5.48 (0.64)	5.54 (0.79)	5.50 (0.54)
	P2 (12)	5.25 (1.61)	5.75 (1.39)	5.52 (0.79)	5.61 (0.92)	5.56 (0.76)
	A2 (12)	6.06 (0.92)	6.04 (1.10)	5.40 (0.80)	4.69 (0.87)	5.26 (0.69)
	P (24)	5.54 (1.45)	5.90 (1.22)	5.50 (0.81)	5.19 (1.07)	5.42 (0.78)
	A (24)	5.68 (1.08)	5.81 (1.24)	5.44 (0.71)	5.11 (0.92)	5.38 (0.62)
PC (12)	P1 (6)	4.67 (1.14)	4.25 (1.94)	4.63 (1.29)	4.72 (0.46)	4.62 (0.76)
	A1 (6)	4.56 (0.40)	5.25 (1.08)	4.83 (0.89)	5.92 (0.71)	5.40 (0.43)
	P2 (6)	4.89 (0.34)	5.83 (0.61)	5.29 (0.49)	5.71 (0.70)	5.52 (0.46)
	A2 (6)	5.11 (0.86)	4.42 (2.18)	4.58 (1.44)	4.75 (0.81)	4.73 (0.88)
	P (12)	4.78 (0.81)	5.04 (1.60)	4.96 (0.99)	5.22 (0.76)	5.07 (0.76)
	A (12)	4.83 (0.70)	4.83 (1.70)	4.71 (1.15)	5.33 (0.95)	5.06 (0.75)
Phone (12)	P1 (6)	5.06 (1.12)	4.67 (1.33)	4.79 (0.43)	5.76 (0.41)	5.30 (0.43)
	A1 (6)	4.56 (1.26)	3.33 (1.78)	5.00 (1.34)	5.26 (1.23)	4.78 (0.98)
	P2 (6)	4.56 (1.26)	3.42 (2.13)	4.75 (0.99)	4.76 (0.87)	4.50 (0.96)
	A2 (6)	5.17 (0.94)	4.50 (1.61)	4.63 (0.63)	4.76 (0.51)	4.76 (0.42)
	P (12)	4.81 (1.17)	4.04 (1.81)	4.77 (0.73)	5.26 (0.83)	4.90 (0.82)
	A (12)	4.86 (1.11)	3.92 (1.73)	4.81 (1.02)	5.01 (0.93)	4.77 (0.72)

TABLE 5.6: Type III ANOVA table with Satterthwaite's method (Streaming Audio) Significant codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

Effect	Sum Sq	Mean Sq	F value	Pr(>F)
Mode	4.97	2.49	14.23	< .001***
Playthrough	1.81	1.81	10.38	.005**
Part	0.17	0.17	1.00	.322
Strategy	0.08	0.08	0.43	.514
Mode:Playthrough	0.12	0.06	0.34	.716
Mode:Part	0.99	0.50	2.84	.068
Playthrough:Part	0.00	0.00	0.00	.983
Mode:Strategy	0.05	0.02	0.14	.868
Playthrough:Strategy	0.04	0.04	0.24	.630
Part:Strategy	0.18	0.18	1.02	.325
Mode:Playthrough:Part	0.14	0.07	0.40	.672
Mode:Playthrough:Strategy	0.05	0.02	0.14	.870
Mode:Part:Strategy	0.17	0.09	0.49	.616
Playthrough:Part:Strategy	0.07	0.07	0.42	.525
Mode:Playthrough:Part:Strategy	0.81	0.41	2.32	.107

TABLE 5.7: EMMeans for Mode by Playthrough, Part, and Strategy for Streaming Audio.

Playthrough	Part	Strategy	Mode	Emmean	SE
First	Middle	Absent	PC	4.48	0.329
First	Middle	Absent	Phone	4.29	0.329
First	Middle	Absent	VR	4.79	0.261
Second	Middle	Absent	PC	5.34	0.329
Second	Middle	Absent	Phone	4.87	0.329
Second	Middle	Absent	VR	5.73	0.261
First	Start	Absent	PC	4.79	0.329
First	Start	Absent	Phone	4.66	0.329
First	Start	Absent	VR	5.46	0.261
Second	Start	Absent	PC	5.50	0.329
Second	Start	Absent	Phone	5.42	0.329
Second	Start	Absent	VR	5.54	0.261
First	Middle	Present	PC	5.03	0.329
First	Middle	Present	Phone	4.31	0.329
First	Middle	Present	VR	5.35	0.261
Second	Middle	Present	PC	5.50	0.329
Second	Middle	Present	Phone	5.21	0.329
Second	Middle	Present	VR	5.77	0.261
First	Start	Present	PC	4.26	0.329
First	Start	Present	Phone	4.83	0.329
First	Start	Present	VR	4.78	0.261
Second	Start	Present	PC	5.35	0.329
Second	Start	Present	Phone	5.40	0.329
Second	Start	Present	VR	5.78	0.261

TABLE 5.8: Post-hoc comparisons using EMMeans (Tukey-adjusted) for Streaming Audio Strategy. Significant codes: ** $p < .01$, * $p < .05$.

Playthrough	Part	Strategy	Contrast	Estimate	SE	t-ratio	p-value
First	Middle	Absent	PC - Phone	0.20	0.403	0.488	0.877
First	Middle	Absent	PC - VR	-0.31	0.314	-0.984	0.590
First	Middle	Absent	Phone - VR	-0.51	0.314	-1.609	0.250
Second	Middle	Absent	PC - Phone	0.47	0.403	1.177	0.471
Second	Middle	Absent	PC - VR	-0.38	0.314	-1.220	0.447
Second	Middle	Absent	Phone - VR	-0.86	0.314	-2.729	0.023*
First	Start	Absent	PC - Phone	0.13	0.403	0.314	0.947
First	Start	Absent	PC - VR	-0.67	0.314	-2.141	0.091
First	Start	Absent	Phone - VR	-0.80	0.314	-2.544	0.036*
Second	Start	Absent	PC - Phone	0.08	0.403	0.195	0.979
Second	Start	Absent	PC - VR	-0.04	0.314	-0.140	0.989
Second	Start	Absent	Phone - VR	-0.12	0.314	-0.390	0.920
First	Middle	Present	PC - Phone	0.71	0.403	1.774	0.187
First	Middle	Present	PC - VR	-0.32	0.314	-1.021	0.567
First	Middle	Present	Phone - VR	-1.04	0.314	-3.295	0.005**
Second	Middle	Present	PC - Phone	0.30	0.403	0.735	0.744
Second	Middle	Present	PC - VR	-0.27	0.314	-0.847	0.675
Second	Middle	Present	Phone - VR	-0.56	0.314	-1.790	0.182
First	Start	Present	PC - Phone	-0.57	0.403	-1.420	0.337
First	Start	Present	PC - VR	-0.52	0.314	-1.647	0.235
First	Start	Present	Phone - VR	0.05	0.314	0.174	0.984
Second	Start	Present	PC - Phone	-0.05	0.403	-0.122	0.992
Second	Start	Present	PC - VR	-0.43	0.314	-1.367	0.365
Second	Start	Present	Phone - VR	-0.38	0.314	-1.211	0.452

5.4.2 Qualitative Results

As in the first experiment, inductive coding was used, resulting in 29 codes spread across 5 themes: *Audio*, *Differences Physical-Virtual*, *Game Design*, *Social and Contextual Factors*, and *Tracking Issues*. All the themes and codes can be seen in Table 5.10, with the Theme, Sub-theme, and an Example quote from the sub-theme. Separate tables for codes that came from the two different strategies were not created, as there was some overlap between the codes between the strategies. While some overlap exists between codes from the first experiment and this one, previous codes were not used due to the inductive nature of the coding process, starting with an empty codebook.

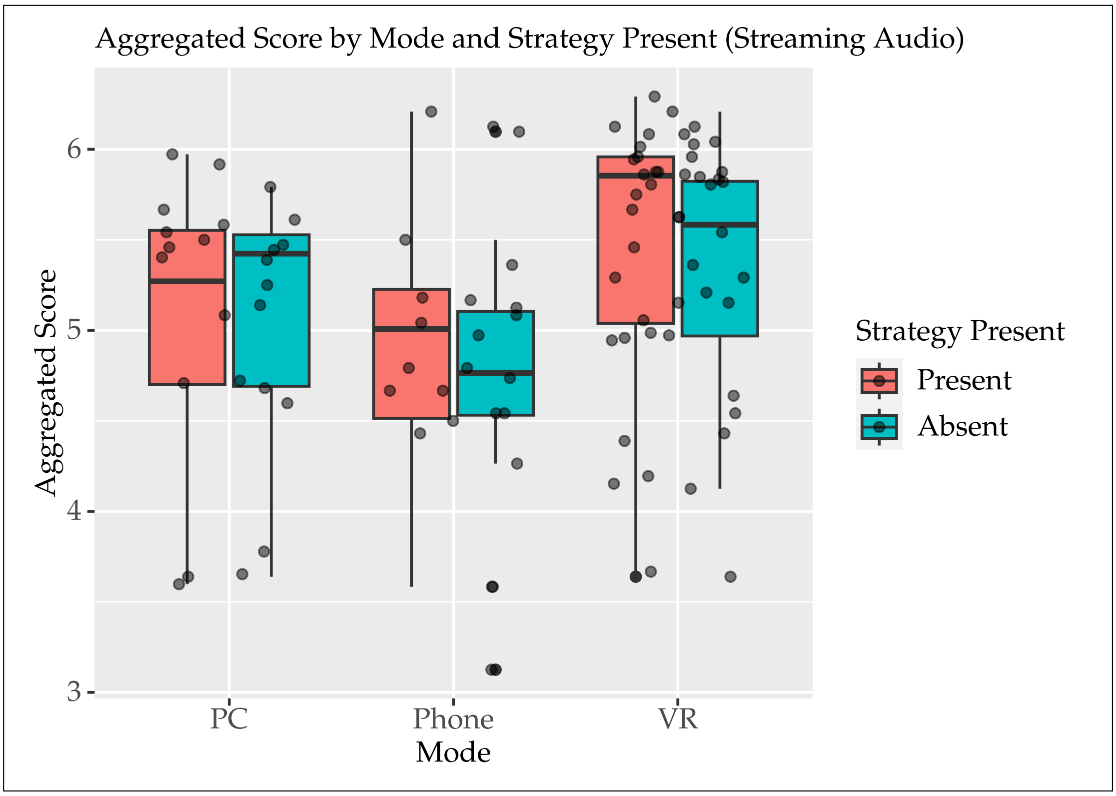

FIGURE 5.8: A scatter plot of the total audio streaming scores of each mode (each dot represents one questionnaire).

Additionally, a set of 4 sub-themes were created to keep track of the condition that applied to each quote, indicating if the person who was quoted was talking about their first VR playthrough with the Embodied Interactions strategy, or the PC drone player on their second playthrough with the Streaming Audio strategy, etc. These sub-themes aren't included in the theme table but will be discussed in the subsequent analysis.

TABLE 5.9: The modes that were coded for keeping track of the state of the game for the quote

Mode		Occurrences
Condition	Absent-Present	156
	Present-Absent	213
Device	VR	245
	PC	77
	Phone	75
Experiment	Audio	185
	Zoom	183
Turn	First	231
	Second	168

The following subsections will describe the themes.

TABLE 5.10: Observations: Themes, Sub-Themes, and Example Quotes and Player ID that said it, with the number of occurrences in parentheses in the sub-theme.

Theme	Sub-theme (#)	Example Quotes (Player ID)
Audio	Outside Audio Interference (42)	"(...) it was not unpleasant, but I was more aware that there was a person" (P16)
	Headphone Immersion (98)	"I thought it was much better to hear them in the game, because it was really quiet otherwise, because the headsets were quiet" (P31)
	Voice Quality Bad (22)	"I felt that maybe that was a problem with the connection or something, because P22's voice was very broken at the start" (P21)
	Audio Latency (33)	"It kind of felt like an echo really" (P21)
	Game Audio Interferes with Outside (32)	"It kind of drowns them out a little bit. [talking about the alarm]" (P23)
	Spatial Audio (9)	It felt a lot more natural with the spatial sounds, that when P25 spoke to me through the microphone, I was able to identify instead of just looking around" (P26)
Differences Physical-Virtual	Being Affected by the Other Player (24)	"And then you kept like moving me around" (P1)
	Affecting the Other Player (31)	"The way I was moving was just inadvertently interacting with the drone, where I wasn't intentionally trying to push you away from the side" (P2)
	Physical-Virtual Mismatch (24)	"I felt as if I was just kind of moving around 3D space and sometimes getting smacked around." (P25)
	Physical Proximity Awareness (23)	"After I noticed that the drone was the actual player I was a bit aware of it after that." (P16)
Game Design	Embodied Interaction Transfer Effect (4)	"So when I was in the VR, I think at some point I did realize that he was telling me the numbers faster." (P16)
	Intuitive Controls (33)	"Yeah, for me, the E is second nature, the keyboard, so just going there is easy." (P4)
	Unintuitive Controls (20)	"It's unintuitive [the broken zoom], compared to the scrolling, which is something that everyone has done at this point, probably." (P7)
	Narrative Enhancing Game Experience (63)	"It felt more realistic to me because I got their audio right in my ear, just like how in a space scenario, you'd hear it directly in your ear." (P27)

TABLE 5.10: Observations: Themes, Sub-Themes, and Example Quotes and Player ID that said it, with the number of occurrences in parentheses in the sub-theme.

Theme	Sub-theme (#)	Example Quotes (Player ID)
	Switch of Embodied Interaction Disrupts Game Flow (5)	"you're not thinking about the game, you're thinking about the mechanics of what button do I press" (P12)
	Unfamiliarity with own Avatar (20)	"On my first play through, I just sort of imagined myself as like a little floating camera" (P1)
	Unnatural Interactions (22)	"And when you're pressing the buttons it wasn't natural because I wouldn't press buttons like that" (P11)
	Urgency Diminishing Game Engagement (5)	"So when it came to the keypad, I sort of wanted to scoot over, press the button, completely forgetting that the other player was there, and then, whoops" (P43)
	Urgency Enhancing Game Engagement (38)	"because it happens right around the point where the pressure meter starts going faster as well, that kind of distracts the player enough to let the drone player figure out how to do the zoom quickly" (P10)
	Virtual Space Size Enhances Immersion (5)	"I've got cramped space to work with and it immersed you a little bit" (P1)
Social and Contextual Factors	Familiarity with Games (25)	"Well, it's a bit biased because I've used the WASD as a player, so it felt normal." (P6)
	No Communication Enhances Game (2)	"At least for me it was immersive right from the get-go, because we weren't really talking." (P25)
	Observable Avatar of the Other Player (33)	"It happened because when she tried to enter the codes, she literally blocked me." (P6)
	Second Playthrough Knowledge (15)	"But I think it helped because I had been in the VR, so I knew what it looked like." (P4)
	Not Moving within the Space (16)	"Didn't feel like it was a necessity, so you didn't feel like exploring." (P9)
	Misunderstanding Hinders Game Progress (4)	"So at the beginning, I just felt a little confused and didn't know how to." (P44)
	Switching Controls Disrupts Game Experience (4)	"I think the scroll is easier, but then, because I was used to the E, so I couldn't switch." (P5)

TABLE 5.10: Observations: Themes, Sub-Themes, and Example Quotes and Player ID that said it, with the number of occurrences in parentheses in the sub-theme.

Theme	Sub-theme (#)	Example Quotes (Player ID)
Tracking Issues	Accuracy of Tracking (25)	“The middle numbers are harder [to press].” (P27)
	VR Controller Loss of Tracking (29)	“It randomly put me in some stars.” (P34)

5.4.2.1 Audio

This theme is about anything audio-related, from *Audio Latency* to *Game Audio Interferes with Outside*. Numerous negative comments were noted, especially in the *Audio Latency* code, where 7 out of 24 people mentioned it:

> “It kind of felt like an echo really.” P21

> “When I was in VR it’s slightly delayed, which is kind of weird hearing them because you can kind of hear them in real life and you can kind of hear a delayed version of it.” P23

Even with the negative comments, there were 5 out of 24 people who said it actually made the experience more enjoyable or immersive:

> “ It made it harder, that was part of the fun, right? I wouldn’t say it took me out, it was just more fun, because it became more stressful.” P46

> “ When the feedback started happening, it did make it more immersive, because there was a slight sense of panic.” P45

This shows how the framing of the problem can make a person more receptive to it, and even frame is as an “expected” occurrence.

Headphone Immersion was mentioned by 26 out of 48 people, indicating how the headphones they wore helped them feel more immersed in the experience by isolating the sounds from outside, even at the detriment of not hearing the other person clearly:

> “When I had the VR, I liked the headphones. I wouldn’t have taken them off, even if I couldn’t hear her. It didn’t matter, it was better as an experience.” P6

That code is related to *Outside Audio Interference* (audio from outside the game interfered with the experience) and *Game Audio Interferes with Outside*, as 11 people had difficulty hearing the other person when there were alarms in the game:

> " [when talking about the alarm] I feel like you had to try harder to listen which kind of takes a little bit away from the game." P22

5.4.2.2 Differences Physical-Virtual

This is very similar to the *Maintaining a Mental Model of the Real Environment*, and *Sensory Perception* from the first experiment. There are 4 sub-themes identified, 2 of which are opposites: *Affecting the Other* and *Being Affected by the Other*:

> "I'm like, why am I moving? I was getting really frustrated." P1

Here, the VR player was unintentionally pushing the PC drone player around, and this led to the frustration of the PC player.

Physical Proximity Awareness is a code similar to *Maintaining a Mental Model of the Real World* from the first experiment. It refers to players' awareness of the physical presence of the other, either through physical touch or through the virtual representation of the physical:

> " I initially felt kind of weird knowing that there was someone else in the play space, so I didn't want to move." P14

Physical-Virtual Mismatch was also identified, where the VR player experienced difficulties pressing a virtual button because the controller wasn't exactly in the same place and the button's behaviour, leading to a mismatch between the physical expectation of how a button works and the virtual implementation.

5.4.2.3 Game Design

For this theme, everything about the Game Design was covered, such as the *Intuitiveness* or *Unintuitiveness* of the controls, the *Switch of Embodied Interaction Disrupting the Game*, or the *Narrative Enhancing the Game Experience*:

> "I didn't feel out of it because of that, it felt like 'Okay, that's an obstacle introduced for the game' and it seemed natural in the context." P9

Several opposing codes were noted, such as *Intuitiveness* vs. *Unintuitiveness* of the controls, and the fact that *Urgency* either diminishes or enhances the game experience:

> "It was just a bit like, I was so in the moment, I didn't really have time to think about it." P34

This quote's context is for when the player is talking about an *Unnatural Interaction* they experienced in the context of *Urgency*. This combination of one negative interaction being overshadowed by a positive one is noteworthy and will be explored further in the discussion.

5.4.2.4 Social and Contextual Factors

This theme encompasses factors such as the user's *Familiarity with Games* and the self-imposed rule of *Not Communicating* through talking:

> "And the communication was done through the drone, like how the drone was moving, the flashlight and everything. So that was very immersive because I couldn't hear P26 here because he wasn't talking at all." P25

This observation highlights the narrative of the game and the way it's presented, where players are depicted as strained on a space station with a scientist from Earth assisting them virtually.

Other factors include the decision to *Not Move within the Space* of the game, particularly among Phone players, as many (16 people) noticed that they don't need to move a lot for them to do the tasks:

> 'Didn't feel like it was a necessity, so you didn't feel like exploring. Like if it said you can explore a bit, maybe I would." P9

This quote further illustrates how the way you present the narrative of the game, and the affordances each player has, greatly can affect the experience.

5.4.2.5 Tracking Issues

This theme addresses instances where people complained about losing tracking, specifically for the VR or Phone player, as they used the VR tracking in the game. Sometimes, it's about the *Accuracy of the Tracking*, where it either wasn't as fast and responsive, or some drift was introduced. These problems were mainly because there is a second player in the same space, making it harder for the lighthouse base stations to see the controllers and headset and track them accurately:

> "I felt like the phone was a little bit more immersive just because it was a bit more accurate for me" P11

Interestingly, this quote confirms some of the results from the Embodied Interactions strategy, where there wasn't a significant difference in embodiment between VR and Phone. This suggests that the Phone version benefited from the Embodied Interaction strategy.

The other code, *VR Controller Loss of Tracking*, is a more specialised version of the *Accuracy of Tracking*, where it addresses instances where the controller lost tracking:

> "..but being the drone operator, I found it more like creating dead zones in the play space that I just couldn't enter." P14

This code is connected to *Not Move within the Space*, as some people noticed that if they remained in a single place, they minimised the chance of problems with the tracking (16 out of 24 people). However, these two codes were kept separate because *Tracking Issues* were not a small part of the experience for some people, and the difference between the controller completely losing tracking for a second or two and the accuracy of the tracking not being perfect is not small.

5.4.3 Video Analysis

The videos of the playthroughs were analysed and used to help guide the qualitative analysis. The videos documented instances where people crashed/collided into one another, and included observations based on this data. As this applies only to the VR/Phone version of the game, the analysis is mostly centred around those 24 data points.

The Phone players were divided into 3 different groups: those who mostly stayed in the same place, those who moved occasionally, and those who explored the whole space and moved a lot. The number of collisions and near-collisions between VR and Phone players was also noted.

Out of 24 Phone players, 12 stayed in their starting position throughout most of the game, 4 moved around occasionally and explored the space, and the rest 8 didn't stay in a single place, constantly moving around the VR player and helping them. On average, people would get very close and almost touch each other around 2-3 times (during the *Drone fixing* task) and crash into one another around 1-2 times during a single playthrough.

5.5 Discussion

This discussion section is derived from the analysis of both the quantitative and qualitative data in order to try and understand the complexities of the interactions between modes. The quantitative analysis revealed somewhat unexpected results, such as the lack of significant differences within modes, while the qualitative interviews provided insights into participants' experiences and perceptions. By combining the two analyses, this discussion aims to shed light on how narrative and other contextual factors could influence player embodiment, offering a fuller understanding of the experimental outcomes.

The questionnaire results show no significant difference in the embodiment within the mode from both strategies, and a number of significant differences between the different modes.

Referring back to the first experiment, there is one surprising result: in the version with the Embodied Interaction, there was a significant difference between the PC and Phone, while in the first experiment, there wasn't one. It is important to note that different questionnaires were used, so the results can't be directly compared, although the questions were similar. This finding can be considered an observation rather than a definitive conclusion.

Focusing on the results of the ANOVA (Table 5.2 and Table 5.6), they were unexpected. It was hypothesised that embodiment in one mode would be detectable withing that mode and potentially in the other, which isn't shown where looking at the effects between Mode:Strategy. The inconclusive results suggest that other effects may be influencing the experience, or the questionnaire wasn't detailed enough to capture the difference. This is further showcased by the final effect of the Embodied Interaction analysis, where Mode:Playthrough:Part:Strategy have a significant difference, meaning there is a complex relationship between them. In contrast, the interviews revealed a different perspective.

5.5.1 Narrative plays a *major* part in the experience

One of the key strategies used to minimise disruption during strategy changes was to use the narrative as a driver for that change. What was not initially realised is that this, in itself, constitutes a strategy, although not one found in the systematic review.

The narrative was so prominent in participants' discussion about the experience that it was evident as a significant factor. By narrative, it is not only meant "game" narrative, but the whole "meta-narrative" as well. Participants are in a lab setting, doing an experiment, knowing that things might break, and ready to be more forgiving towards any "mistakes" that happen. This raises several important questions for experiments like this in the future:

> Question 1: When designing an experiment like this, should efforts be made to isolate and make the experience as bare bones as possible, so as to minimise the random variables?

> Question 2: Or is the act of isolating it mostly futile, as real-world experiences are rarely as clear-cut as this, with multiple factors mixing and affecting one another?

These questions are crucial for understanding such experiences and the design of studies around them. A significant difference might have been observed if the change had been more noticeable and isolated. For example, if the whole experience involved the zooming task with an embodied interaction for a few minutes, and then completely changing the scene and disabling the embodied interaction. Alternatively, the questionnaire could be more explicit about it, asking specifically about the zooming rather than the change as a whole.

Returning to the interview results with this context, 21 participants mentioned how they felt that the "breaking" of the zoom or the audio was part of the story, so they perceived it as just another event within the game:

> "It did feel more natural, but I think by the context in which you do not have that anymore, like it's presented as a malfunction, you now have to use the options menu." P9

> "But then at the same time, I can see that situation in real life, you would have to do lots of calculations in your head. So I feel like thinking about whether I need to type in 50 or 100, it makes sense in a way." P34

These quotes show how participants felt that the way that the change of the strategy state occurred "made sense", and thus it didn't really take them out of the experience, or change it in any way.

> "These things happen. It happens. Stuff happens. You just get used to it, especially with some of the new games." P43

> "Again, I think just expecting everything to be a bit terrible, because that's the whole point." P46

These two quotes highlight on the expectations of the participants, which, due to the framing of the experiment in a university lab, helped them accept and be less distracted by such occurrences. It underlines another big theme observed in the data:

5.5.2 A problem comes with a "solution"

When participants discussed a problem, there was often a "solution" that alleviated said problem. For example:

> "[When talking about being pushed by the other player unintentionally] And I thought it was bad. I thought it was just like the controller was meant to be like not as responsive because you were in space." P1

Here, the problem is *Being Affected by the Other Player* (including both physical and virtual contact), or being pushed by them. The solution was found in the narrative, once again.

> "[When asked about the Phone drone going out of bounds and if that felt unnatural] Not necessarily, because for some reason I thought that this was an opening for the drone to go out from." P16

In this instance, the design of the space station, which is part of the narrative of being in space, helped the player to not perceive *Unnatural Interactions*.

> "Personally, I would say it did feel more immersive, although there was some sort of uneasiness hearing your own voice being said twice. But I guess that's just part of the parcel." P43

The problem identified here is *Audio Latency*, but the solution is the *Narrative* of the game and setting, and the expectation of the player. This specific player had a lot of VR experience, so they anticipated many of the small issues that might arise, connecting *Narrative* to *Familiarity with Games*.

> "So instead of thinking about the red lights, you're like, yeah, we're on a mission. Thinking about the environment, you get the code instead. So I don't think red lights are an abstraction, they're more of an environment that you have to get out of. It's a natural response that red actually means threat." P36

This quote leads to the next point observed in the data:

5.5.3 Narrative is intertwined with a lot of other codes

The quote itself illustrates how *Urgency Enhanced* the game experience, due in part to the narrative of it being in space, being "alone", and having to act quickly to save the ship.

> "Yeah, the alarm is quite loud and I tried to keep it off so I could hear better. But I think it's just part of the game, it makes you focus more because it simulates a real scenario." P37

When faced with *Game Audio Interfering with the Outside*, this player was helped by the narrative explanation that it's a part of the game, of the experience, and that it actually simulates a real-life scenario.

> "[When asked if Physical-Virtual Mismatch affected them] It was in the game, yeah, because realistically also some button might not work, you know, so it didn't take the attention out for me." P41

Here, the problems are *Physical-Virtual Mismatch* and *Unnatural Interactions*, but the solution is found in the *Narrative*, as similar real-life situations might also have buttons breaking and not working properly. These instances showcase how the *Narrative* significantly impacts the game experience, suggesting a new strategy to be added to the model.

5.5.4 *Narrative* as a strategy to enhance embodiment

Having the *Narrative* play such an important role in an experience can greatly help with the design, as problems or difficulties with the system can be framed in a narrative way. This is not always possible, of course, but it definitely has a place in the *Embodiment Aspects* and *Levels*. While it was an unintentional strategy, it is evident that narrative is crucial, and future studies should explore its impact on embodiment and immersion in multi-modal experiences. Here it is important to mention Slater's [79] "Plausibility Illusion", as it pertains to a similar phenomenon of a person's ability to believe that a certain scenario is happening. This is similar to "narrative" embodiment, as called in the thesis, but differs to an extent in the level of embodiment and participation in the narrative.

This is also a form of "Seamful Design" [16], a way to design experiences where the "seam", or the problems that can happen in the system (*Loss of Tracking*, *Unnatural Interactions*, etc.), are not hidden, but are part of the experience. The difference between framing a game as a "computer simulation of a potential emergency in a space station" and as a "relaxing puzzle game" is significant. Integrating the "seams" in a "computer simulation" would be much easier than in a simple puzzle game.

> "[When asked about losing tracking as a drone] Yeah, if the game had been, I don't know, set on a farm, and I was moving crates around, then that would de-immerse me. But because the whole point was communication difficulties, I'm inferring, it was just part of the experience." P46

This quote illustrates the importance of narrative and setting. The chosen narrative fits the theme of "technical difficulties", enhancing the experience in such a way that some effects might have been mitigated. This can be an example of "seamful" design: audio problems that, instead of trying to fix, were used advantageously, making the problem into a "solution".

Returning back to the two questions posed at the start of this discussion section, it is evident that without the narrative, or with a narrative-gameplay mismatch, the experience would have differed significantly. The effect of changing an embodied interaction would likely have been more noticeable without the narrative to fall back on when problems arose.

A strong example of the *Narrative*'s power, as well as the *Intuitiveness of the Controls*, is seen in this conversation between the two players during the interview:

> "I was just clicking the..." P8
> "Oh, you were clicking the numbers on the screen?" P7
> "Yeah." P8
> "Oh, I think I just typed it in if you type on the keyboard keypad" P7
> "I don't think you can type." P8
> "Did I not? Oh my god, I don't even remember what I did. I was that immersed, clearly. I thought I typed it in." P7

This player was so immersed, due to the *Intuitive Controls* and *Narrative*, that they completely missed the fact that their zoom function broke, and now they had to use the mouse to click on each number separately on a screen UI (Figure 5.4). This demonstrates how *Narrative* can overshadow changes in embodied interactions. It was expected that such change would be noticed and reflected in the questionnaire results, potentially lowering the embodiment level.

Despite the emphasis on *Narrative* being a big part of the experience, there were participants for whom that wasn't as helpful, where they did actually notice the change and it did affect them in some way.

5.5.5 The surprising Phone results for the Embodied Interactions strategy

As shown in the tables (Table 5.2, Table 5.4) and scatter plot for Embodied Interactions (Figure 5.7), the Phone version has similar scores to the VR version and better than the PC version. This was the expected outcome from the first experiment, making these findings particularly interesting. Having done two different strategies, one can compare the results and see that there is a noticeable difference in embodiment when using the Embodied Interaction strategy, as opposed to the Streaming Audio one (Figure 5.7 and Figure 5.8).

The significant role of *Audio Latency* in the Streaming Audio strategy on the Phone version supports the effectiveness of the Embodied Interactions strategy. A t-test between the two different strategies' Phone versions (just a simple total, so both states where the strategy was present and not) showed a significant difference, with a p-value of $\mathbf{p = 0.0217}$. This further helps reinforce the fact that strategy choice impacts each mode differently.

Although an experiment where people played the game without any strategies present throughout the whole game was not conducted (as the first experiment), a comparison was made between the Phone with the Embodied Interaction present, and the Phone with the Streaming Audio not working (representing the normal game). The t-test between these conditions yielded a p-value of $p = 0.0667$. Further, comparing the zoom interaction overall, regardless of the strategy being present or not, resulted in a p-value of $\mathbf{p = 0.0342}$, indicating a difference.

The implications of this can be threefold: the Streaming Audio strategy was not executed as effectively due to the *Audio Latency* and other audio issues; the inclusion of a new task that uses the phone touchscreen together with the phone holder increased overall embodiment; or the narrative explanation for the "broken" zoom mitigated its impact on the sense of embodiment; or a combination of those three options.

5.5.6 Other minor observations

Several minor observations from the interviews are worth noting. Four instances of *Embodiment Interaction Transfer Effect* were observed, as hypothesised in the first experiment, where embodiment from one player is noticed by the other and in turn, helps them:

> " [When asked about whether they could tell the other player was taking a longer time] I definitely could tell." P34

> "So when I was in the VR, I think at some point I did realise that he was telling me the numbers faster." P16

Additionally, participants with VR experience (*Familiarity with Games*) took on a "mentor" role to help the other player when they were confused or unsure what to do, echoing one of the themes that is also in the previous experiment, *Prior Experience*. This was made as an observation from the actual playthrough rather than the interviews. As in the previous experiment, the fact that the drone is introduced as a "helper" further supported this dynamic.

Video analysis revealed a trend where the more immersed in the task the players were (and the VR player specifically), the more they crashed, which players quickly recovered from and continued playing without disruptions. There was an instance of the VR player being so immersed and concentrated on the task at hand, that they constantly kept crashing into the Phone player (when fixing the drone) and just continued to play unaffected. When asked in the interview, the player said that they were just too engrossed in the game and hitting the other player didn't detract from the experience, as it felt like a "real drone" was there. This ties back to the *Narrative* of the drone around them and the way that the drone player can interact with them through said narrative.

5.6 Summary

This second experiment was conducted with the aim of finding how a strategy and its absence affects embodiment within the mode, hoping that can also transfer to the other person in the best-case scenario. Research Question 3 asked **"How can design strategies impact the different modes of play in terms of Embodiment?"**. Initially, the quantitative results appeared unexpected and not useful, as the anticipated difference within the mode, with and without the strategy, was not observed. However, further examination, supported by the qualitative data, led to several interesting and helpful observations, including the identification of a potential new embodiment strategy that can be added to the Aspects and Levels models:

> Using the *Narrative* as a way to embody people (similar to the first experiment's *Narrative embodiment*), as it played a *significant* role in the players' experiences. This can be framed as using "seamful" design principles to accentuate rather than hide the imperfections of an

> experience. Instead of concealing problems with the narrative, these issues can be incorporated and utilised within the story.

Narrative was found to be linked with many other codes and played a major part in some of the "solutions" of the problems people mentioned. Examples include participants not hearing the other player well but rationalising it within the story, or finding certain interactions unnatural but reconciling them through the narrative context.

A recurring theme was the presence of a "problem" with the experience and a convenient "solution", often found through the *Narrative*, such as being pushed by another player but accepting it due to the space setting.

One of the major findings was that the results from the Embodied Interactions version on the Phone versus PC had a significant difference in embodiment ($\mathbf{p < 0.001}$), contrasting with the results of the first experiment. Although different questionnaires were used, the observed difference was noteworthy. Using the fact that the Streaming Audio version without the strategy is the exact same game as the first experiment, the Embodied Interactions pinch-to-zoom Phone version and the non-Streaming Audio Phone version could be compared. This found a significant difference with a p-value of $\mathbf{p = 0.0342}$. Additionally, comparing the two Phone versions with the strategies both on and off showed a p-value of $\mathbf{p = 0.0217}$.

These findings suggest that the Embodied Interaction strategy on the Phone did have a positive effect, indicating that the use of the touch screen on the phone, whether with a more or less embodied interaction, was beneficial rather than the embodied interaction alone.

To summarise the answer to RQ3, it was found that different strategies impact the modes differently (as seen with the *Embodied Interaction* strategy between the PC and Phone), but no inter-mode effects (VR-PC and VR-Phone) were demonstrated. This might be due to the overpowering influence of the *Narrative* strategy implemented alongside the other strategies being tested.

Based on these results, two major questions arise for future research in this multi-modal field: Should the testing experience be made as dry and isolated as possible to identify differences between variables? Or is it more beneficial to analyse a more comprehensive experience, as real-world applications are not as controlled as lab environments?

These questions do not have easy answers. However, from the experiment results, it is suggested that isolating variables is extremely challenging and may not reflect the final user experience. Therefore, it may be more useful to analyse and dissect a more comprehensive experience. Understanding this "full" experience through the combined use of quantitative and qualitative analysis is crucial for grasping the complex interactions within multi-modal VR experiences and their impact on user embodiment.

In the next chapter, the thesis will be concluded with a summary of the entire study, going back to the research questions answering them, listing contributions, the limitations and future work, and providing final remarks.

Chapter 6

Conclusions

In this final chapter, the thesis will be summarised, outlining the key activities undertaken and how the research questions have been answered. The contributions made by this work will be discussed, followed by a review of the limitations and future work, concluding with final remarks.

6.1 Summary

At the start of this thesis, three main research questions were posited. Subsequently, two experiments and a systematic literature review were conducted to explore these questions, providing interesting results and guiding the direction of subsequent experiments.

The overarching theme of the thesis is multi-modality in VR: how important is it, how does one design for such experiences, and what factors should people keep in mind when creating such experiences. To address these questions, a summary of the work conducted is provided.

The research started with an exploratory experiment to examine the current multi-modal research in the space. A multi-modal VR game was created for the purposes of the research and an experiment was conducted to determine whether the introduction of a second non-HMD player in a VR game helps or detracts from immersion and co-presence. The results showed only a significant difference between the non-HMD and HMD modes in term of immersion, but not co-presence, as well as no difference between the two non-HMD modes used (PC and Phone). The interviews helped interpret the results after inductive coding, showing how “embodiment”, as it was called, was very important in the experience, and how it was intertwined with a number of factors. For example, the interviews suggested how increasing embodiment using a specific strategy could positively impact immersion and co-presence.

These qualitative results from the interview highlighted the importance of embodiment in a multi-modal experience and indicated that a more nuanced understanding was needed to influence it meaningfully. In the background section, embodiment was explored as an experiential factor, noting the lack of an easy way to discuss multi-modal games and the experience of embodiment within them.

This led to a systematic literature review of embodiment and the different strategies to affecting it. The review, consisting of 60 papers after exclusion criteria, helped shed light on the use of the term "embodiment" in the literature, and resulted in the creation of two orthogonal models of embodiment, the *Aspects* and *Levels* of embodiment. These models, along with a set of strategies to affect different types of embodiment, were developed to assist in designing experiences with these aspects in mind.

The final experiment utilised the results from the first experiment and the systematic review to modify the game and test whether changing embodiment using one strategy affects the overall embodiment of the players and whether cross-transfer effects exist between modes. While different types of embodiment were measured, and differences observed, the quantitative results were inconclusive regarding cross-mode effects. However, qualitative interview results explained why this was the case: the *narrative* strategy used inadvertently overshadowed the actual strategy being tested.

In the next section, the three research questions will be addressed, demonstrating how the research has answered them.

6.2 Research questions

The following subsections will begin by addressing RQ2 first, as it provides context for the other research questions.

6.2.1 RQ2 — Embodiment and its Different Definitions

"What are the different kinds of embodiment and what are the different strategies that are used to help achieve them?"

Background research revealed that the literature on embodiment was very diverse, and potentially confusing, as there were a number of different definitions from different fields, including psychology and computer science. This diversity extended to terms like immersion and presence, which further complicated the terminology.

In the first experiment, themes were named as embodiment, such as *Narrative* and *Physical* Embodiment, without initially referencing the literature. While similar definitions existed for the latter, *Narrative Embodiment* was not found in the literature, the closest phenomena being *Plausibility Illusion*, talked about by Slater [79].

This prompted a systematic review in order to better understand embodiment and how it can be affected in different modes. Saturation sampling was used to scope the study, and "The Engineering Village" database was used to gather a number of relevant papers. The review covered the past 10 years to provide a modern perspective on embodiment..

From this review, 16 different codes for embodiment, with 20 different strategies that affect embodiment, and 39 mappings that connected a strategy to a type of embodiment were identified. The following analysis of the codes and data helped develop two orthogonal models of embodiment: the *"Embodiment Aspects"* and *"Embodiment Levels"*. They build on top of current literature, as a number of researchers used similar definitions of embodiment in their papers. The *"Aspects"* model is made out of 4 different sub-types, 3 of which are from Kilteni et al. (Self-Location, Agency, and Body Ownership), and one that was noted from the literature, Social Embodiment. The fourth one was added because the social aspect is important when talking about multi-modal multi-user experiences. As an example, one could be socially embodied just through a phone call, or in a completely virtual 3D environment with other people inside. This is an important addition, as it would be hard to talk about those experiences without the social embodiment aspect.

Opposite of that stands the *"Levels"* model. If the *"Aspects"* are *how* you are embodied in a specific instance, then the *"Levels"* are how *much* you are embodied. It consists of 4 levels, at the bottom being the Embodied Avatars (basic control of an avatar). The next 2 levels stand together: Embodied Cognition and Embodied Emotion (deeper levels of embodiment involving thought and feeling like the embodied avatar). At the top is Embodied Identity (the highest level, where identity is tied to the avatar, and one thinks, acts and feels like the avatar).

Together with those two models came a set of strategies that affect the different embodiment types, which in turn could be used in the first experiment to change it and measure the difference. Using a set model like this could help in future research, and thinking about it in terms of *"Aspects"* and *"Levels"* of embodiment could help with insight into a multi-modal experience.

Going back to the research question, this systematic review came back with two different models of embodiment, showing the different types of embodiment found in the literature, and a set of strategies used to affect said embodiment. With so many different definitions of embodiment in the literature and used by researchers, a model like this could be used to, if nothing else, help people understand the interactions within a multi-modal experience better.

For example, it can be useful when researchers aren't sure what is needed to change in order to make an experience more embodied in general, or could be used for more specific types of embodiment. To enhance social embodiment in a co-located experience, one might use the *Aspects* model and employ strategies like *Emotion Projection*. To elevate the experience to a more embodied level, the *Levels* model could guide the use of strategies like *Embodied Interactions*.

With RQ2 answered, the focus will now shift to RQ1 and how the mode of play in a multi-modal experience affects each participant in terms of immersion and co-presence, two other important variables in a multi-modal setting.

6.2.2 RQ1 — the Different Modalities' Effect on Immersion and Co-presence

"How does a non-HMD user's mode of play in a multi-modal VR game affect the Immersion and Co-Presence of all participants?"

To investigate this question, a multi-modal co-op VR game was created, incorporating two non-HMD modes of interaction and play: through a PC (baseline), and through a novel mode using a tracked Phone. The phone uses one of the controllers to precisely track its position within the virtual space, creating an effect like a "window" into the virtual world.

The game, titled 'StuckInSpace' , was designed with the VR player being in space as an astronaut, and the second player being a "drone controller" on Earth, assisting the VR player. This setup enabled the study to test for significant differences between the PC and Phone versions of the game. Several questionnaires were used to measure the usability, immersion, and co-presence, helping later to discern any significant difference between the modes.

The hypothesis was that due to the Phone player's movement and active role around the VR player, there would be a difference in the co-presence results between the VR player paired with a PC versus the one paired with a Phone player. The quantitative results of the experiment were non-significant, except for the cases between VR and non-VR modes (as expected). This prompted an examination of the qualitative interview data for further insights.

The qualitative interviews were instrumental in explaining the lack of significant quantitative differences. One of the codes in the interview, *embodiment*, had played a major role in the participants' experiences. Participants reported different "problems" in the various modalities, which were intertwined. For example, people seemed to have a problem with the physical and virtual world not matching completely, thus they had to keep a *mental model of the real world*, which in turn put a strain on their ability to feel immersed in the game.

The themes and sub-themes coded from the interviews revealed interconnected aspects, such as *mental model of the real world* and *observability of the body*. These connections suggested that modifying one part of the experience (e.g., making the body of the non-VR player easier to see, or more realistic) could affect another aspect of the experience.

Several recommendations were made based on the results and analysis, and of potential design decisions researchers might want to keep in mind when designing multi-modal experiences. A key recommendation highlighted that a single event could have a positive impact in one mode but a negative impact in another. For example, *physical embodiment* on the VR player was positively affected by the presence of the Phone player, up until the point of accidental contact, in which their immersion was briefly interrupted. A potential solution could involve designing the game to minimise direct contact or adjust the size of the Phone player's avatar.

The balance between co-presence and immersion in a multi-modal setting, such as the game used in the study, is delicate, as there were multiple factors of varying significance. This in turn makes it harder to isolate the specifics of the experience that should be changed to increase immersion or co-presence accordingly.

In addressing the research question, it has been shown that the non-HMD player's mode does change the experience of the HMD player, sometimes in a positive way (e.g., when *conversing* about the game world), or sometimes in a negative way (e.g., physical contact between the VR and Phone player). Immersion and co-presence, sometimes at odds with each other, were affected by different factors depending on the mode.

Co-located multi-modal experiences can bring a lot of new and interesting interactions between people. Many participants showed interest in more multi-modal experiences particularly in response to the common problem of VR headset owners hosting friends who merely stand, watch, and wait for their turn with the headset. There was enthusiasm for using smartphones as a means of interacting with the virtual world.

This aligns with the notion from Gugenheimer et al. that becoming a *"part of the social living room environment"* [36] would be a vital part of VR becoming more popular. Not everyone wants to be completely immersed, and having more options for participation can only be a good thing. This research has demonstrated the feasibility of implementing and deploying a co-located multi-modal system using a game ending and creativity. The challenge lies in the actual design of the experience, for which the embodiment models created in this research can provide valuable guidance. This discussion leads into the third and final research question.

6.2.3 RQ3 — the Impact of Embodiment Strategies on a Multi-Modal Experience

"How can design strategies impact the different modes of play in terms of Embodiment?"

Having created a multi-modal VR game in my first experiment together with two models of embodiment in the second study, the knowledge gained from the systematic review was applied to modify StuckInSpace to test the hypothesis: Does changing embodiment affect the modalities as predicted, and what strategies can be used in order to achieve that?

After analysing the data from the first experiment, two different strategies were selected for testing: *Streaming Audio* and *Embodied Interactions*. The selection was informed by examining participant feedback from the first experiment, contextualised within the embodiment models, and aimed at finding a strategy that was both achievable and appropriate for the desired embodiment changed.

Streaming Audio was chosen to address participants' complaints about the difficulty of hearing the other player or confusion about the position of the drone. Positional audio streaming from the microphone of the non-HMD player to the HMD player was implemented to alleviate that. The implementation was challenging, especially for the Phone version, due to inherent latency introduced by having voice data transmitted over WiFi. This latency, together with the fact that the two participants are in the same room close to each other, meant that a slight delay could be heard and potentially could be a detriment to the experience.

Embodied Interactions on the other hand was an easier strategy to implement. An embodied interaction is an interaction that can be performed effortlessly, such as pinch-to-zoom on the Phone and using the mouse scroll wheel to zoom on the PC.

The strategies were implemented to switch on/off in the middle of the game, in order to allow the analysis of their effects. The choice to make the change through a narrative device, where the drone "broke" or got "fixed" mid-game, unintentionally introduced a third strategy.

The analysis of the quantitative data didn't reveal the expected differences between the strategy present and absent within the mode. The qualitative data later on helped in understanding why exactly that might have happened. Notably, the Phone version had similar levels of embodiment as the VR version in the *Embodied Interactions* strategy, contrasting with the *Streaming Audio* strategy, which had the expected lower levels of embodiment between the Phone and VR versions.

This surprising discovery led to the theory that the data reflected a combination of the new zoom task on the *Embodied Interactions* Phone version (increasing embodiment), and the reduced embodiment from audio latency issues experienced by the phone in the *Streaming Audio* version.

The interviews and coding process further cemented that theory, revealing that the *Narrative* played a major role in participants' experience and feeling of embodiment. This narrative extended beyond the game itself to the context surrounding the game, such as being in a lab, being observed, and interacting with strangers. The *Narrative* had helped narrow the gap between the on and off condition, in some cases so much, that the participants hadn't noticed it had happened, or just accepted it without thinking about it.

Using the *Narrative* emerged as a new strategy that can be integrated in the two models created, offering a design tool for multi-modal experiences. This aligns with the concept of "seamful" design [16], where, rather than hiding the imperfections and problems of the experience, they come to the front, and can be used to the advantage of the designer. For example, increased audio latency could be explained within the story, making it an accepted part of the experience. This also ties to the notion of "Plausability Illusion" that Slater [79] talks about.

This does pose an important question for future research and design of multi-modal experiences: Should factors be isolated to be studied separately, or do you go for a more "real-world" approach, where you have the whole experience and try to analyse it realistically?

Answering the third research question, it has been shown that the choice of design strategy significantly affects embodiment. This choice in itself has to be made carefully, considering all the available options. The unintended use of *Narrative* as a device to increase embodiment overshadowed the intended strategies, complicating the quantitative analysis and obscuring the significant differences. While the question was not fully answered, as significant cross-mode effects could not be confirmed, the results provided insights into why this might have been the case.

6.2.4 Discussion

This section provides a short discussion on the implications of the results from the research questions, and how they relate to some of the literature reviewed.

In the first experiment, two *"embodiment"* codes were created, which were not necessarily connected with the definitions in the literature. The second user study showed how *"narrative embodiment"*, as it was themed, is a crucial part of a multi-modal game. The lack of studies around it highlighted the need for further research into the use of *narrative* as a strategy to affect embodiment.

The two VR experiments showed the importance of good and thoughtful design, especially in multi-modality experiences. Small changes in a feature could lead to major changes to the embodiment and enjoyment of the participants, far beyond what a standard flat-screen game might experience. This is in part thanks to the immersive and novel nature of VR, paired with the co-located nature of the multi-modal game.

Literature on co-located multi-modal games confirmed several findings from the experiments. For example, Cheng et al. [19] talk about the fact that even though the non-HMD users aren't as immersed into a virtual environment as the VR player, they still experience increased enjoyment when they help the VR player in different tasks. This was echoed in both experiments, where the interviews showed that many participants enjoyed the helper role, which made the game more enjoyable. Another observation was that the more experience a person has with games, and VR in specific, the more they expressed enjoyment with helping the other player, possibly due to their understanding of the VR experience and desire to enchance it for others.

One of the key observations of the thesis is the significant impact of design choices on multi-modal games, which can affect the overall experience in different ways. This also aligns with the literature, such as the discussion of Lee et al. [50] on the importance of design in affecting immersion. A number of papers [9, 66] also support the observation that co-located play can hinder some aspects of the experience while enhancing others. Physical touch, when expected and matched with the virtual world, seemed to help players feel more immersed and embodied. The problem was that a lot of the times it wasn't expected. There are a few potential solutions where one can try minimising unwanted and accidental contact, and increase desired physical contact.

Gómez and Fons [30] point out how the mere fact of wearing the HMD prevents social presence from increasing in a co-located experience, which is in part due to how "isolating" it is. This research does show that there are different ways of introducing a second co-located player in a different modality, and that each way brings different challenges and considerations. The importance of the *narrative* as a strategy could be useful when talking about co-presence, as quite a few participants expressed how the story helped them feel like there was an actual person close to them, trying to help.

The next section summarised the contributions made by this thesis.

6.3 Contributions

This thesis has contributed to the academic community in the following ways:

1. The design and implementation of a multi-modal multiplayer VR game, *StuckInSpace*, which was used as a basis of the experiments throughout the

thesis. This game helped explore multi-modality and the way that design is of paramount importance for such experiences. Multiple modalities interact with each other in many different ways, prompting designers to consider their decisions carefully;

2. A study that looked into the differences between immersion and co-presence throughout the modes of play (VR, PC and Phone), which put into light some questions about what embodiment is, and how it is connected to many parts of the experience. This study also showed the importance of mixed methods research, where both quantitative and qualitative data-gathering methods were used. Each one of them helped to gather a more comprehensive understanding of the whole experience;

3. A framework of embodiment consisting of two orthogonal models: the *Aspects* and *Levels* of embodiment. This framework, developed from a systematic review, includes a number of different strategies that can be used to affect said elements of embodiment. This can be used to help design co-located multi-modal experiences, and informed the final experiment;

4. An example use of the embodiment model in the analysis of the results of the first experiment, showing how a designer could use this model for deeper insight into a game, and then subsequently implement some strategies to potentially improve embodiment in a modality;

5. A user study examining the differences between two separate embodiment strategies present and absent, and the subsequent analysis and discussion of the surprising results; and

6. The identification of a potential new embodiment strategy from analysis of the final experiment: using the *Narrative* to help embody people, connected to the idea of "seamful" design and "Plausibility Illusion".

These contributions collectively advance the understanding of multi-modal VR experiences and provide practical insights and tools for future research and design in this field.

Contributions number 1 and 2 have been published in two conferences:

- Malinov, Y.-D. (2020) *Characterising the Benefits of Multi-Modal Play in Virtual Reality.* in Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '20, page 10–11

- Malinov, Y.-D., Millard, D. E., and Blount, T. (2021). *StuckInSpace: Exploring the Difference Between Two Different Mediums of Play in a Multi-Modal Virtual Reality Game.* in 2021 IEEE Virtual Reality and 3D User Interfaces (VR), pages 501–510

6.4 Limitations and Future Work

While this thesis has contributed and helped shed light on the design and intricacies of multi-modal experiences, several limitations provide future studies with more avenues for investigation. One limitation would be the sample size of people throughout the experiments, as that tended to be on the lower side. With a larger sample size it could possibly have been found that there is a bigger difference. Although it was a limitation on the quantitative analysis, it wasn't really a problem for the qualitative, which gave several interesting observations. Future studies could benefit from a larger pool of people, helping with the ability to make more generalisable statements.

Regarding the systematic review, only more recent papers were considered, and saturation sampling was used. This approach might have led to the exclusion of some older and more popular papers that talk about embodiment, and potentially missing it in the final model. A potential review of the whole space could offer a more comprehensive understanding of embodiment in multi-modal experiences. Additionally, exploration of literature beyond HCI and Computer Science can also provide novel perspectives on embodiment.

Another major limitation of the technology itself was the *Audio Streaming* on the Phone version, as that was as much of a software and implementation limitation as it was a hardware one. Achieving ultra-low latency with audio communication is a challenging problem. Future research could potentially look into how the new strategy found in the analysis, use of *narrative*, can be implemented in such a way as to make people feel fewer disruptions caused by the latency, or make it a part of the story itself. This presents an interesting venue for future research.

The last experiment showed the importance of *narrative* for designing embodied multi-modal experiences. Future research could focus on using narrative to enhance embodiment and to "hide" the problems in the experience in a way by making them an actual story element in it ("seamful" design, and making use of the "Plausibility Illusion"). Potential avenues for this could explore the different ways that story and setting can affect multi-modal experiences and the different types of people that are more or less affected by this (as some people could embody the narrative stronger than others in the experiment conducted).

This thesis can be used as a stepping stone into more research in co-located multi-modal spaces. In particular, it would be fascinating to investigate how different narratives and settings can influence various aspects of immersion, co-presence, and embodiment. Additionally, considering the diverse audience that engages in virtual reality, exploring the individual differences and preferences in narrative engagement and embodiment could uncover tailored design strategies.

Another interesting direction for future research is evaluating the effectiveness of each strategy or how strongly each strategy affects the experience. It is possible that a strategy can work really well in isolation, but when introduced into the full experience, it quickly loses strength because of cross effects and interactions between strategies.

6.5 Final Remarks

In reflecting upon this thesis, my hope is that the work will prove valuable to those studying embodiment, immersion, and co-presence in VR and multi-modal games and experiences. I hope that the lessons I learned from the experimental design can serve as a guide to potential future researchers and designers. Moreover, I trust that my key findings will help others create multi-modal experiences that are immersive and enjoyable, using the insights provided to create environments that captivate and embody people within a narrative.

The game I created, StuckInSpace, can be used not only as a template for future mixed-reality multi-modal games but also as a potential example for others attempting to create such experiences.

In conclusion, with the foundation laid by this research, the future holds promising opportunities to push the boundaries of co-located multi-modal spaces, enabling a more captivating, inclusive, and transformative virtual reality experience for a wide range of users.

Bibliography

[1] S Bangay and L Preston. "An investigation into factors influencing immersion in interactive virtual reality environments." In: *Studies in health technology and informatics* 58 (1998), p. 43. DOI: 10.3233/978-1-60750-902-8-43. https://doi.org/10.3233/978-1-60750-902-8-43.

[2] Mafkereseb Kassahun Bekele and Erik Champion. "A Comparison of Immersive Realities and Interaction Methods: Cultural Learning in Virtual Heritage". In: *Frontiers in Robotics and AI* 6 (2019). ISSN: 2296-9144. DOI: 10.3389/frobt.2019.00091. https://www.frontiersin.org/articles/10.3389/frobt.2019.00091.

[3] Steve Benford et al. "Understanding and constructing shared spaces with mixed-reality boundaries". In: *ACM Trans. Comput.-Hum. Interact.* 5.3 (Sept. 1998), pp. 185–223. ISSN: 1073-0516. DOI: 10.1145/292834.292836. https://doi.org/10.1145/292834.292836.

[4] Mehmet Ilker Berkman. "History of Virtual Reality". In: *Encyclopedia of Computer Graphics and Games*. Springer, 2024, pp. 873–881. DOI: 10.1007/978-3-031-23161-2_169.

[5] Mark Billinghurst, Suzanne Weghorst, and T Furness. "Shared space: An augmented reality approach for computer supported collaborative work". In: *Virtual Reality* 3.1 (1998), pp. 25–36. https://doi.org/10.1007/BF01409795.

[6] Frank Biocca, Chad Harms, and Jenn Gregg. "The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity". In: *4th annual international workshop on presence, Philadelphia, PA*. 2001, pp. 1–9. http://hypermedia468.pbworks.com/w/file/fetch/80689061/networked%5C%20minds%5C%20measure%5C%20presence.pdf.

[7] Sarah-Jayne Blakemore, Daniel M Wolpert, and Christopher D Frith. "Abnormalities in the awareness of action". In: *Trends in cognitive sciences* 6.6 (2002), pp. 237–242. https://doi.org/10.1098/rstb.2000.0734.

[8] Olaf Blanke and Thomas Metzinger. "Full-body illusions and minimal phenomenal selfhood". In: *Trends in cognitive sciences* 13.1 (2009), pp. 7–13. DOI: https://doi.org/10.1016/j.tics.2008.10.003. https://www.sciencedirect.com/science/article/pii/S1364661308002507.

[9] F. Born, P. Sykownik, and M. Masuch. "Co-Located vs. Remote Gameplay: The Role of Physical Co-Presence in Multiplayer Room-Scale VR". In: *2019 IEEE Conference on Games (CoG)*. 2019, pp. 1–8. DOI: 10.1109/CIG.2019.8848001. https://doi.org/10.1109/CIG.2019.8848001.

[10] C. W. Borst, N. G. Lipari, and J. W. Woodworth. "Teacher-Guided Educational VR: Assessment of Live and Prerecorded Teachers Guiding Virtual Field Trips". In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 467–474. DOI: 10.1109/VR.2018.8448286. https://doi.org/10.1109/VR.2018.8448286.

[11] Matthew Botvinick and Jonathan Cohen. "Rubber hands 'feel'touch that eyes see". In: *Nature* 391.6669 (1998), pp. 756–756. DOI: 10.1038/35784. https://doi.org/10.1038/35784.

[12] Sidney Bovet et al. "The Critical Role of Self-Contact for Embodiment in Virtual Reality". In: *IEEE Transactions on Visualization and Computer Graphics* 24.4 (2018), pp. 1428–1436. DOI: 10.1109/TVCG.2018.2794658. https://doi.org/10.1109/TVCG.2018.2794658.

[13] D. Brewster. *The Stereoscope; Its History, Theory and Construction, with Its Application to the Fine and Useful Arts and to Education, Etc*. Victorian popular culture. John Murray, 1856. https://books.google.co.uk/books?id=VI85AAAAcAAJ.

[14] John Brooke et al. "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7. DOI: 10.1201/9781498710411-35. https://doi.org/10.1201/9781498710411-35.

[15] John Brooke. "SUS: a retrospective". In: *Journal of usability studies* 8.2 (2013), pp. 29–40. DOI: 10.5555/2817912.2817913. https://doi.org/10.5555/2817912.2817913.

[16] M. Chalmers, I. MacColl, and M. Bell. "Seamful design: showing the seams in wearable computing". In: *2003 IEE Eurowearable*. 2003, pp. 11–16. DOI: 10.1049/ic:20030140. https://doi.org/10.1049/ic:20030140.

[17] Liwei Chan and Kouta Minamizawa. "FrontFace: Facilitating Communication between HMD Users and Outsiders Using Front-Facing-Screen HMDs". In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '17. Vienna, Austria: Association for Computing Machinery, 2017. ISBN: 9781450350754. DOI: 10.1145/3098279.3098548. https://doi.org/10.1145/3098279.3098548.

[18] Chih-Hung Chen et al. "A Desktop Virtual Reality Earth Motion System in Astronomy Education." In: *Educational Technology & Society* 10 (July 2007), pp. 289–304. https://www.jstor.org/stable/jeductechsoci.10.3.289?seq=1.

[19] Lung-Pan Cheng et al. "Haptic Turk: A Motion Platform Based on People". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, pp. 3463–3472. ISBN: 9781450324731. DOI: 10.1145/2556288.2557101. https://doi.org/10.1145/2556288.2557101.

[20] Lung-Pan Cheng et al. "TurkDeck: Physical Virtual Reality Based on People". In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. UIST '15. Charlotte, NC, USA: Association for Computing Machinery, 2015, pp. 417–426. ISBN: 9781450337793. DOI: 10.1145/2807442.2807463. https://doi.org/10.1145/2807442.2807463.

[21] Igor Curcio, Anna Dipace, and Anita Norlund. "Virtual realities and education". In: *Research on Education and Media* 8 (Dec. 2016). DOI: 10.1515/rem-2016-0019. https://doi.org/10.1515/rem-2016-0019.

[22] Frédérique De Vignemont. "Embodiment, ownership and disownership". In: *Consciousness and cognition* 20.1 (2011), pp. 82–93. DOI: https://doi.org/10.1016/j.concog.2010.09.004. https://www.sciencedirect.com/science/article/pii/S1053810010001704.

[23] Paul Dourish. "Embodied interaction: Exploring the foundations of a new approach to HCI". In: *Work* (1999), pp. 1–16. https://www.dourish.com/publications/misc/embodied.pdf.

[24] Clarence A Ellis, Simon J Gibbs, and Gail Rein. "Groupware: some issues and experiences". In: *Communications of the ACM* 34.1 (1991), pp. 39–58. DOI: 10.1145/99977.99987. https://doi.org/10.1145/99977.99987.

[25] Leighton Evans and Michal Rzeszewski. "Hermeneutic Relations in VR: Immersion, Embodiment, Presence and HCI in VR Gaming". In: *HCI in Games*. Ed. by Xiaowen Fang. Cham: Springer International Publishing, 2020, pp. 23–38. ISBN: 978-3-030-50164-8.

[26] Nicolas Franck et al. "Defective recognition of one's own actions in patients with schizophrenia". In: *American Journal of Psychiatry* 158.3 (2001), pp. 454–459. DOI: 10.1176/appi.ajp.158.3.454. https://doi.org/10.1176/appi.ajp.158.3.454.

[27] Brian Gajadhar, Yvonne De Kort, and Wijnand Ijsselsteijn. "Shared Fun Is Doubled Fun: Player Enjoyment as a Function of Social Setting". In: Oct. 2008, pp. 106–117. DOI: 10.1007/978-3-540-88322-7_11.

[28] Gernot Goebbels and Vali Lalioti. "Co-Presence and Co-Working in Distributed Collaborative Virtual Environments". In: *Proceedings of the 1st International Conference on Computer Graphics, Virtual Reality and Visualisation*. AFRIGRAPH '01. Camps Bay, Cape Town, South Africa: Association for Computing Machinery, 2001, pp. 109–114. ISBN: 1581134460. DOI: 10.1145/513867.513891. https://doi.org/10.1145/513867.513891.

[29] Erving Goffman. *Behavior in public places*. Simon and Schuster, 2008.

[30] Maureira Marcello A. Gómez and Verbeek Fons. "The Impact of Co-Located Play on Social Presence and Game Experience in a VR Game". In: *DiGRA/FDG*. Dundee, Scotland: Digital Games Research Association and Society for the Advancement of the Science of Digital Games, Aug. 2016.

[31] Mar Gonzalez-Franco, Parastoo Abtahi, and Anthony Steed. "Individual differences in embodied distance estimation in virtual reality". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2019, pp. 941–943. DOI: 10.1109/VR.2019.8798348. https://doi.org/10.1109/VR.2019.8798348.

[32] Mar Gonzalez-Franco and Tabitha C Peck. "Avatar embodiment. towards a standardized questionnaire". In: *Frontiers in Robotics and AI* 5 (2018), p. 74. DOI: 10.3389/frobt.2018.00074. https://doi.org/10.3389/frobt.2018.00074.

[33] Denis Gračanin. "Immersion Versus Embodiment: Embodied Cognition for Immersive Analytics in Mixed Reality Environments". In: *Augmented Cognition: Intelligent Technologies*. Ed. by Dylan D. Schmorrow and Cali M. Fidopiastis. Cham: Springer International Publishing, 2018, pp. 355–368. ISBN: 978-3-319-91470-1. https://doi.org/10.1007/978-3-319-91470-1_29.

[34] J. G. Grandi, H. G. Debarba, and A. Maciel. "Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 127–135. DOI: 10.1109/VR.2019.8798080. https://doi.org/10.1109/VR.2019.8798080.

[35] Jan Gugenheimer et al. "FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality". In: Apr. 2018, pp. 1–13. DOI: 10.1145/3173574.3173628. https://doi.org/10.1145/3173574.3173628.

[36] Jan Gugenheimer et al. "ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 4021–4033. ISBN: 9781450346559. DOI: 10.1145/3025453.3025683. https://doi.org/10.1145/3025453.3025683.

[37] Chad Harms and Frank Biocca. "Internal consistency and reliability of the networked minds measure of social presence". In: (2004). https://core.ac.uk/download/pdf/87479.pdf.

[38] Kiran Ijaz et al. "Player Experience of Needs Satisfaction (PENS) in an Immersive Virtual Reality Exercise Platform Describes Motivation and Enjoyment". In: *International Journal of Human–Computer Interaction* 36.13 (2020), pp. 1195–1204. DOI: 10.1080/10447318.2020.1726107. https://doi.org/10.1080/10447318.2020.1726107.

[39] Akira Ishii et al. "ReverseCAVE: CAVE-Based Visualization Methods of Public VR Towards Shareable VR Experience". In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA '18. Montreal QC, Canada: Association for Computing Machinery, 2018, p. 1. ISBN: 9781450356213. DOI: 10.1145/3170427.3186593. https://doi.org/10.1145/3170427.3186593.

[40] Sebastian Kapp et al. "Augmenting Kirchhoff's laws: Using augmented reality and smartglasses to enhance conceptual electrical experiments for high school students". In: *The Physics Teacher* 57.1 (Jan. 2019), pp. 52–53. ISSN: 0031-921X. DOI: 10.1119/1.5084931. https://doi.org/10.1119/1.5084931.

[41] Muhammad Sikandar Lal Khan, Haibo Li, and Shafiq Ur Réhman. "Tele-Immersion: Virtual Reality Based Collaboration". In: *HCI International 2016 – Posters' Extended Abstracts*. Ed. by Constantine Stephanidis. Cham: Springer International Publishing, 2016, pp. 352–357. ISBN: 978-3-319-40548-3. DOI: 10.1007/978-3-319-40548-3_59. https://doi.org/10.1007/978-3-319-40548-3_59.

[42] Konstantina Kilteni, Raphaela Groten, and Mel Slater. "The Sense of Embodiment in Virtual Reality". In: *Presence: Teleoperators and Virtual Environments* 21.4 (Nov. 2012), pp. 373–387. DOI: 10.1162/PRES_a_00124. https://doi.org/10.1162/PRES%5C_a%5C_00124.

[43] Dawon Kim and Yosoon Choi. "Applications of Smart Glasses in Applied Sciences: A Systematic Review". In: *Applied Sciences* 11.11 (2021). ISSN: 2076-3417. DOI: 10.3390/app11114956. https://www.mdpi.com/2076-3417/11/11/4956.

[44] Thomas Kirks et al. "Evaluation of the Application of Smart Glasses for Decentralized Control Systems in Logistics". In: *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. 2019, pp. 4470–4476. DOI: 10.1109/ITSC.2019.8917159. https://doi.org/10.1109/ITSC.2019.8917159.

[45] Barbara Kitchenham. "Procedures for performing systematic reviews". In: *Keele, UK, Keele University* 33.2004 (2004), pp. 1–26. https://www.researchgate.net/publication/228756057_Procedures_for_Performing_Systematic_Reviews.

[46] Robert E. Kraut, Darren Gergle, and Susan R. Fussell. "The Use of Visual Information in Shared Visual Spaces: Informing the Development of Virtual Co-Presence". In: *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*. CSCW '02. New Orleans, Louisiana, USA: Association for Computing Machinery, 2002, pp. 31–40. ISBN: 1581135602. DOI: 10.1145/587078.587084. https://doi.org/10.1145/587078.587084.

[47] Donald R. Lampton et al. "Side Effects and Aftereffects of Immersion in Virtual Environments". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 38.18 (1994), pp. 1154–1157. DOI: 10.1177/154193129403801802. https://doi.org/10.1177/154193129403801802.

[48] JR Landis and GG Koch. "The measurement of observer agreement for categorical data". In: *Biometrics* 33.1 (Mar. 1977), pp. 159–174. ISSN: 0006-341X. DOI: 10.2307/2529310. https:%20//doi.org/10.2307/2529310.

[49] Aleksandra Landowska et al. "Within-and between-session prefrontal cortex response to virtual reality exposure therapy for acrophobia". In: *Frontiers in human neuroscience* 12 (2018), p. 362. DOI: 10.3389/fnhum.2018.00362. https://doi.org/10.3389/fnhum.2018.00362.

[50] Jiwon Lee, Mingyu Kim, and Kim Jinmo. "RoleVR: Multi-experience in immersive virtual reality between co-located HMD and non-HMD users". In: *Multimedia Tools and Applications* 79 (Oct. 2019). DOI: 10.1007/s11042-019-08220-w. https://doi.org/10.1007/s11042-019-08220-w.

[51] Jiabao Li, Honghao Deng, and Panagiotis Michalatos. "CatEscape: An Asymmetrical Multiplatform Game Connecting Virtual, Augmented and Physical World". In: *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '17 Extended Abstracts. Amsterdam, The Netherlands: Association for Computing Machinery, 2017, pp. 585–590. ISBN: 9781450351119. DOI: 10.1145/3130859.3130860. https://doi.org/10.1145/3130859.3130860.

[52] Lan Li et al. "Application of virtual reality technology in clinical medicine". In: *American journal of translational research* 9.9 (2017), p. 3867. https://pubmed.ncbi.nlm.nih.gov/28979666/.

[53] Henry E. Lowood. *Virtual reality*. Sept. 16, 2023. https://www.britannica.com/technology/virtual-reality (visited on 09/21/2021).

[54] Yoan-Daniel Malinov, David E. Millard, and Tom Blount. "StuckInSpace: Exploring the Difference Between Two Different Mediums of Play in a Multi-Modal Virtual Reality Game". In: *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 2021, pp. 501–510. DOI: 10.1109/VR50410.2021.00074.

[55] Yoan-Daniel Grigorov Malinov. "Characterising the Benefits of Multi-Modal Play in Virtual Reality". In: *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '20. Virtual Event, Canada: Association for Computing Machinery, 2020, pp. 10–11. ISBN: 9781450375870. DOI: 10.1145/3383668.3419955.

[56] Susan Michalak. *White Paper — Guidelines for Immersive Virtual Reality Experiences*. Tech. rep. Intel Corporation, 2017. https://software.intel.com/sites/default/files/managed/78/95/GuidelinesforImmersiveVirtualRealityExperiences.pdf.

[57] Paul Milgram et al. "Augmented reality: A class of displays on the reality-virtuality continuum". In: *Telemanipulator and Telepresence Technologies* 2351 (Jan. 1994). DOI: 10.1117/12.197321. https://doi.org/10.1117/12.197321.

[58] Marvin Minsky. "Telepresence". In: (1980).

[59] Ahmed Mostafa et al. "VRSpineSim: Applying Educational Aids Within A Virtual Reality Spine Surgery Simulator". In: June 2019, pp. 380–387. ISBN: 978-3-030-22513-1. DOI: 10.1007/978-3-030-22514-8_34. https://doi.org/10.1007/978-3-030-22514-8_34.

[60] Michael Narayan et al. "Quantifying the Benefits of Immersion for Collaboration in Virtual Environments". In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST '05. Monterey, CA, USA: Association for Computing Machinery, 2005, pp. 78–81. ISBN: 1595930981. DOI: 10.1145/1101616.1101632. https://doi.org/10.1145/1101616.1101632.

[61] Kristine L. Nowak and Frank Biocca. "The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments". In: *Presence: Teleoperators and Virtual Environments* 12.5 (2003), pp. 481–494. DOI: 10.1162/105474603322761289. https://doi.org/10.1162/105474603322761289.

[62] Odd Raven Studios. Carly and the Reaperman - Escape from the Underworld. Game [Microsoft Windows]. Odd Raven Studios, Stockholm, Sweden. Last played December 2019. Stockholm, Sweden, June 2018.

[63] Perkins Coie LLP. *2019 Augmented and Virtual Reality Survey Report*. Tech. rep. "1201 Third Avenue, Seattle, Washington, USA", Mar. 2019. https://www.perkinscoie.com/images/content/2/1/v4/218679/2019-VR-AR-Survey-Digital-v1.pdf.

[64] Valeria Ivanova Petkova, Mehrnoush Khoshnevis, and H Henrik Ehrsson. "The perspective matters! Multisensory integration in ego-centric reference frames determines full-body ownership". In: *Frontiers in psychology* 2 (2011), p. 35. DOI: 10.3389/fpsyg.2011.00035. https://doi.org/10.3389/fpsyg.2011.00035.

[65] Thammathip Piumsomboon et al. "Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: Association for Computing Machinery, 2018. ISBN: 9781450356206. DOI: 10.1145/3173574.3173620. https://doi.org/10.1145/3173574.3173620.

[66] Iana Podkosova and Hannes Kaufmann. "Mutual Collision Avoidance during Walking in Real and Collaborative Virtual Environments". In: *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. I3D '18. Montreal, Quebec, Canada: Association for Computing Machinery, 2018. ISBN: 9781450357050. DOI: 10.1145/3190834.3190845. https://doi.org/10.1145/3190834.3190845.

[67] D. Pohl and C. F. de Tejada Quemada. "See what I see: Concepts to improve the social acceptance of HMDs". In: *2016 IEEE Virtual Reality (VR)*. 2016, pp. 267–268. DOI: 10.1109/VR.2016.7504756. https://doi.org/10.1109/VR.2016.7504756.

[68] Priscilla Ramsamy et al. "Using Haptics to Improve Immersion in Virtual Environments". In: *Computational Science – ICCS 2006*. Ed. by Vassil N. Alexandrov et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 603–609. ISBN: 978-3-540-34382-0. https://doi.org/10.1007/11758525_81.

[69] Holger Regenbrecht et al. "Mixed Voxel Reality: Presence and Embodiment in Low Fidelity, Visually Coherent, Mixed Reality Environments". In: *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2017, pp. 90–99. DOI: 10.1109/ISMAR.2017.26. https://doi.org/10.1109/ISMAR.2017.26.

[70] Jacek Ruminski et al. "The data exchange between smart glasses and healthcare information systems using the HL7 FHIR standard". In: *2016 9th International Conference on Human System Interactions (HSI)*. 2016, pp. 525–531. DOI: 10.1109/HSI.2016.7529684. https://doi.org/10.1109/HSI.2016.7529684.

[71] Joseph P Salisbury et al. "Concussion Assessment With Smartglasses: Validation Study of Balance Measurement Toward a Lightweight, Multimodal, Field-Ready Platform". In: *JMIR Mhealth Uhealth* 6.1 (Jan. 2018), e15. ISSN: 2291-5222. DOI: 10.2196/mhealth.8478. http://www.ncbi.nlm.nih.gov/pubmed/29362210.

[72] Eva-Lotta Sallnäs. "Collaboration in Multimodal Virtual Worlds: Comparing Touch, Text, Voice and Video". In: (Jan. 2002). DOI: 10.1007/978-1-4471-0277-9_10. https://doi.org/10.1007/978-1-4471-0277-9_10.

[73] Atsushi Sato and Asako Yasuda. "Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership". In: *Cognition* 94.3 (2005), pp. 241–255. DOI: 10.1016/j.cognition.2004.04.003. https://doi.org/10.1016/j.cognition.2004.04.003.

[74] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. "The Experience of Presence: Factor Analytic Insights". In: *Presence: Teleoperators and Virtual Environments* 10.3 (2001), pp. 266–281. DOI: 10.1162/105474601300343603. https://doi.org/10.1162/105474601300343603.

[75] Martijn J Schuemie et al. "Research on presence in virtual reality: A survey". In: *CyberPsychology & Behavior* 4.2 (2001), pp. 183–201. DOI: 10.1089/109493101300117884. https://doi.org/10.1089/109493101300117884.

[76] Stephanie Schulze et al. "The Effects of Embodiment in Virtual Reality on Implicit Gender Bias". In: *Virtual, Augmented and Mixed Reality. Multimodal Interaction*. Ed. by Jessie Y.C. Chen and Gino Fragomeni. Cham: Springer International Publishing, 2019, pp. 361–374. ISBN: 978-3-030-21607-8. https://doi.org/10.1007/978-3-030-21607-8_28.

[77] Shervin Shirmohammadi and Nicolas D Georganas. "An end-to-end communication architecture for collaborative virtual environments". In: *Computer Networks* 35.2-3 (2001), pp. 351–367. DOI: https://doi.org/10.1016/S1389-1286(00)00186-9. https://doi.org/10.1016/S1389-1286(00)00186-9.

[78] Mel Slater. "Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire". In: *Presence: Teleoperators and Virtual Environments* 8.5 (1999), pp. 560–565. DOI: 10.1162/105474699566477. https://doi.org/10.1162/105474699566477.

[79] Mel Slater. "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1535 (2009), pp. 3549–3557.

[80] Mel Slater and Sylvia Wilbur. "A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments". In: *Presence: Teleoperators & Virtual Environments* 6.6 (1997), pp. 603–616.

[81] Mel Slater et al. "First person experience of body transfer in virtual reality". In: *PloS one* 5.5 (2010), e10564. DOI: 10.1371/journal.pone.0010564. https://doi.org/10.1371/journal.pone.0010564.

[82] Mel Slater et al. "Small-group behavior in a virtual and real environment: A comparative study". In: *Presence: Teleoperators & Virtual Environments* 9.1 (2000), pp. 37–51. DOI: 10.1162/105474600566600. https://doi.org/10.1162/105474600566600.

[83] Steel Crate Games. Keep Talking and Nobody Explodes. Game [Microsoft Windows]. Steel Crate Games, Ottawa, Ontario, Canada. Last played January 2020. Ottawa, Ontario, Canada, July 2015.

[84] Jonathan Steuer. "Defining virtual reality: Dimensions determining telepresence". In: *Journal of communication* 42.4 (1992), pp. 73–93. DOI: https://doi.org/10.1111/j.1460-2466.1992.tb00812.x. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x.

[85] M P Strzys et al. "Physics holo.lab learning experience: using smartglasses for augmented reality labwork to foster the concepts of heat conduction". In: *European Journal of Physics* 39.3 (Mar. 2018), p. 035703. DOI: 10.1088/1361-6404/aaa8fb. https://dx.doi.org/10.1088/1361-6404/aaa8fb.

[86] SIE Japan Studio Team ASOBI! The Playroom VR. Game [Playstation 4]. Team ASOBI!, SIE Japan Studio, Tokyo, Japan. Tokyo, Japan, Oct. 2016.

[87] Theophilus Teo et al. "Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–14. ISBN: 9781450359702. DOI: 10.1145/3290605.3300431. https://doi.org/10.1145/3290605.3300431.

[88] Vankrupt Games. Pavlov VR. Game [Microsoft Windows]. Vankrupt Games, Vancouver, British Columbia, Canada. Vancouver, British Columbia, Canada, Feb. 2017.

[89] Maria Virvou and George Katsionis. "On the Usability and Likeability of Virtual Reality Games for Education: The Case of VR-ENGAGE". In: *Comput. Educ.* 50.1 (Jan. 2008), pp. 154–178. ISSN: 0360-1315. DOI: 10.1016/j.compedu.2006.04.004. https://doi.org/10.1016/j.compedu.2006.04.004.

[90] VRChat. VRChat. Game [Microsoft Windows]. VRChat, San Francisco, CA, United States. Last played November 2019. San Francisco, CA, United States, Feb. 2017.

[91] S.G. Weinbaum. *Pygmalion's Spectacles*. 1935.

[92] Whirlybird Games. VR The Diner Duo. Game [Microsoft Windows]. Whirlybird Games, Uddevalla, Sweden. Uddevalla, Sweden, Nov. 2016.

[93] Margaret Wilson. "Six views of embodied cognition". In: *Psychonomic bulletin & review* 9.4 (2002), pp. 625–636. https://doi.org/10.3758/BF03196322.

[94] Bob G. Witmer and Michael J. Singer. "Measuring Presence in Virtual Environments: A Presence Questionnaire". In: *Presence: Teleoperators and Virtual Environments* 7.3 (1998), pp. 225–240. DOI: 10.1162/105474698565686. https://doi.org/10.1162/105474698565686.

[95] Josef Wolfartsberger, Jan Zenisek, and Norbert Wild. "Data-Driven Maintenance: Combining Predictive Maintenance and Mixed Reality-supported Remote Assistance". In: *Procedia Manufacturing* 45 (2020). Learning Factories across the value chain – from innovation to service – The 10th Conference on Learning Factories 2020, pp. 307–312. ISSN: 2351-9789. DOI: https://doi.org/10.1016/j.promfg.2020.04.022. https://www.sciencedirect.com/science/article/pii/S235197892031060X.

[96] Michael Yong et al. "Application of augmented reality to surgical practice: A pilot study using the ODG R7 Smartglasses". In: *Clinical Otolaryngology* 45.1 (2020), pp. 130–134. DOI: https://doi.org/10.1111/coa.13460. https://onlinelibrary.wiley.com/doi/abs/10.1111/coa.13460.

[97] Shanyang Zhao. "Toward a taxonomy of copresence". In: *Presence: Teleoperators & Virtual Environments* 12.5 (2003), pp. 445–455. DOI: 10.1162/105474603322761261. https://doi.org/10.1162/105474603322761261.

[98] Zhuoming Zhou et al. "Astaire: A Collaborative Mixed Reality Dance Game for Collocated Players". In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '19. Barcelona, Spain: Association for Computing Machinery, 2019, pp. 5–18. ISBN: 9781450366885. DOI: 10.1145/3311350.3347152. https://doi.org/10.1145/3311350.3347152.

Systematic Review Papers

[S1] L. Almeida et al. "Towards natural interaction in immersive reality with a cyber-glove". In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. 2019, pp. 2653–2658. DOI: 10.1109/SMC.2019.8914239. https://doi.org/10.1109/SMC.2019.8914239.

[S2] Ayman Alzayat, Mark Hancock, and Miguel A. Nacenta. "Quantitative Measurement of Tool Embodiment for Virtual Reality Input Alternatives". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–11. ISBN: 9781450359702. https://doi.org/10.1145/3290605.3300673.

[S3] Sahar Aseeri et al. "Embodied Realistic Avatar System with Body Motions and Facial Expressions for Communication in Virtual Reality Applications". In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2020, pp. 580–581. DOI: 10.1109/VRW50115.2020.00141. https://doi.org/10.1109/VRW50115.2020.00141.

[S4] Steven Baker et al. "Interrogating Social Virtual Reality as a Communication Medium for Older Adults". In: *Proc. ACM Hum.-Comput. Interact.* 3.CSCW (Nov. 2019). DOI: 10.1145/3359251. https://doi.org/10.1145/3359251.

[S5] Sidney Bovet et al. "The Critical Role of Self-Contact for Embodiment in Virtual Reality". In: *IEEE Transactions on Visualization and Computer Graphics* 24.4 (2018), pp. 1428–1436. DOI: 10.1109/TVCG.2018.2794658. https://doi.org/10.1109/TVCG.2018.2794658.

[S6] Lila Bozgeyikli and Evren Bozgeyikli. "Tangiball: Dynamic Embodied Tangible Interaction with a Ball in Virtual Reality". In: *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*. DIS '19 Companion. San Diego, CA, USA: Association for Computing Machinery, 2019, pp. 135–140. ISBN: 9781450362702. DOI: 10.1145/3301019.3323904. https://doi.org/10.1145/3301019.3323904.

[S7] Filippo Brizzi et al. "Effects of Augmented Reality on the Performance of Teleoperated Industrial Assembly Tasks in a Robotic Embodiment". In: *IEEE Transactions on Human-Machine Systems* 48.2 (2018), pp. 197–206. DOI: 10.1109/THMS.2017.2782490. https://doi.org/10.1109/THMS.2017.2782490.

[S8] Polona Caserman, Augusto Garcia-Agundez, and Stefan Göbel. "A Survey of Full-Body Motion Reconstruction in Immersive Virtual Reality Applications". In: *IEEE Transactions on Visualization and Computer Graphics* 26.10 (2020), pp. 3089–3108. DOI: 10.1109/TVCG.2019.2912607. https://doi.org/10.1109/TVCG.2019.2912607.

[S9] Zubin Choudhary et al. "Virtual Big Heads: Analysis of Human Perception and Comfort of Head Scales in Social Virtual Reality". In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2020, pp. 425–433. DOI: 10.1109/VR46266.2020.00063. https://doi.org/10.1109/VR46266.2020.00063.

[S10] Diane Dewez et al. "Influence of Personality Traits and Body Awareness on the Sense of Embodiment in Virtual Reality". In: *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2019, pp. 123–134. DOI: 10.1109/ISMAR.2019.00-12. https://doi.org/10.1109/ISMAR.2019.00-12.

[S11] Carlos Flavián, Sergio Ibáñez-Sánchez, and Carlos Orús. "The impact of virtual, augmented and mixed reality technologies on the customer experience". In: *Journal of Business Research* 100 (2019), pp. 547–560. ISSN: 0148-2963. DOI: https://doi.org/10.1016/j.jbusres.2018.10.050. https://www.sciencedirect.com/science/article/pii/S0148296318305319.

[S12] Jakob Fröhner et al. "An embodiment paradigm in evaluation of human-in-the-loop control". In: *IFAC-PapersOnLine* 51.34 (2019). 2nd IFAC Conference on Cyber-Physical and Human Systems CPHS 2018, pp. 104–109. ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2019.01.036. https://www.sciencedirect.com/science/article/pii/S2405896319300382.

[S13] Henrique Galvan Debarba et al. "Self-attribution of distorted reaching movements in immersive virtual reality". In: *Computers & Graphics* 76 (2018), pp. 142–152. ISSN: 0097-8493. DOI: https://doi.org/10.1016/j.cag.2018.09.001. https://www.sciencedirect.com/science/article/pii/S0097849318301353.

[S14] Lynda Gerry, Sofia Dahl, and Stefania Serafin. "ADEPT: Exploring the Design, Pedagogy, and Analysis of a Mixed Reality Application for Piano Training". English. In: *Proceedings of the 16th Sound & Music Computing Conference SMC 2019*. Vol. 1. Proceedings of the Sound and Music Computing Conference. 16th Sound and music computing conference, SMC2019 ; Conference date: 28-05-2019 Through 31-05-2019. Sound and Music Computing Network, 2019, pp. 241–249. DOI: 10.5281/zenodo.3249333. http://smc2019.uma.es/index.html.

[S15] Lynda Joy Gerry. "Paint with Me: Stimulating Creativity and Empathy While Painting with a Painter in Virtual Reality". In: *IEEE Transactions on Visualization and Computer Graphics* 23.4 (2017), pp. 1418–1426. DOI: 10.1109/TVCG.2017.2657239. https://doi.org/10.1109/TVCG.2017.2657239.

[S16] Mar Gonzalez-Franco, Parastoo Abtahi, and Anthony Steed. "Individual differences in embodied distance estimation in virtual reality". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2019, pp. 941–943. DOI: 10.1109/VR.2019.8798348. https://doi.org/10.1109/VR.2019.8798348.

[S17] Denis Gračanin. "Immersion Versus Embodiment: Embodied Cognition for Immersive Analytics in Mixed Reality Environments". In: *Augmented Cognition: Intelligent Technologies*. Ed. by Dylan D. Schmorrow and Cali M. Fidopiastis. Cham: Springer International Publishing, 2018, pp. 355–368. ISBN: 978-3-319-91470-1. https://doi.org/10.1007/978-3-319-91470-1_29.

[S18] Denis Gračanin et al. "Mixed Library — Bridging Real and Virtual Libraries". In: *Virtual, Augmented and Mixed Reality*. Ed. by Stephanie Lackey and Jessie Chen. Cham: Springer International Publishing, 2017, pp. 227–238. ISBN: 978-3-319-57987-0. https://doi.org/10.1007/978-3-319-57987-0_18.

[S19] Scott W. Greenwald et al. "Investigating Social Presence and Communication with Embodied Avatars in Room-Scale Virtual Reality". In: *Immersive Learning Research Network*. Ed. by Dennis Beck et al. Cham: Springer International Publishing, 2017, pp. 75–90. ISBN: 978-3-319-60633-0. https://doi.org/10.1007/978-3-319-60633-0_7.

[S20] Mohamed Handosa et al. "An Approach to Embodiment and Interactions with Digital Entities in Mixed-Reality Environments". In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 569–570. DOI: 10.1109/VR.2018.8446458. https://doi.org/10.1109/VR.2018.8446458.

[S21] Mohamed Handosa et al. "Extending Embodied Interactions in Mixed Reality Environments". In: *Virtual, Augmented and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation*. Ed. by Jessie Y.C. Chen and Gino Fragomeni. Cham: Springer International Publishing, 2018, pp. 314–327. ISBN: 978-3-319-91581-4. https://doi.org/10.1007/978-3-319-91581-4_23.

[S22] Wen Huang. "Exploring players' user experience in a high-embodied virtual reality game". In: *2019 IEEE Games, Entertainment, Media Conference (GEM)*. 2019, pp. 1–8. DOI: 10.1109/GEM.2019.8902148. https://doi.org/10.1109/GEM.2019.8902148.

[S23] Elena Kalina and Mina C. Johnson-Glenberg. "Presence and Platform: Effects of Embodiment Comparing a 2D Computer and 3D VR Game". In: *2020 6th International Conference of the Immersive Learning Research Network (iLRN)*. 2020, pp. 31–37. DOI: 10.23919/iLRN47897.2020.9155160. https://doi.org/10.23919/iLRN47897.2020.9155160.

[S24] Fengfeng Ke and Xinhao Xu. "Virtual reality simulation-based learning of teaching with alternative perspectives taking". In: *British Journal of Educational Technology* 51.6 (2020), pp. 2544–2557. DOI: https://doi.org/10.1111/bjet.12936. https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12936.

[S25] Konstantina Kilteni, Raphaela Groten, and Mel Slater. “The Sense of Embodiment in Virtual Reality”. In: *Presence: Teleoperators and Virtual Environments* 21.4 (Nov. 2012), pp. 373–387. DOI: 10.1162/PRES_a_00124. https://doi.org/10.1162/PRES%5C_a%5C_00124.

[S26] Martin Kocur et al. “Flexing Muscles in Virtual Reality: Effects of Avatars’ Muscular Appearance on Physical Performance”. In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’20. Virtual Event, Canada: Association for Computing Machinery, 2020, pp. 193–205. ISBN: 9781450380744. DOI: 10.1145/3410404.3414261. https://doi.org/10.1145/3410404.3414261.

[S27] Kathrin Koebel and Doris Agotai. “Embodied Interaction for the Exploration of Image Collections in Mixed Reality (MR) for Museums and Other Exhibition Spaces”. In: *HCI International 2020 - Posters*. Ed. by Constantine Stephanidis and Margherita Antona. Cham: Springer International Publishing, 2020, pp. 291–299. ISBN: 978-3-030-50732-9. https://doi.org/10.1007/978-3-030-50732-9_39.

[S28] Elena Kokkinara, Mel Slater, and Joan López-Moliner. “The Effects of Visuomotor Calibration to the Perceived Space and Body, through Embodiment in Immersive Virtual Reality”. In: *ACM Trans. Appl. Percept.* 13.1 (Oct. 2015). ISSN: 1544-3558. DOI: 10.1145/2818998. https://doi.org/10.1145/2818998.

[S29] Andrey Krekhov, Sebastian Cmentowski, and Jens Krüger. “The Illusion of Animal Body Ownership and Its Potential for Virtual Reality Games”. In: *2019 IEEE Conference on Games (CoG)*. 2019, pp. 1–8. DOI: 10.1109/CIG.2019.8848005. https://doi.org/10.1109/CIG.2019.8848005.

[S30] Claudia Krogmeier and Christos Mousas. “Eye fixations and electrodermal activity during low-budget virtual reality embodiment”. In: *Computer Animation and Virtual Worlds* 31.4-5 (2020), e1941. DOI: https://doi.org/10.1002/cav.1941. https://onlinelibrary.wiley.com/doi/abs/10.1002/cav.1941.

[S31] Kaylee Payne Kruzan et al. “Supporting Self-Injury Recovery: The Potential for Virtual Reality Intervention”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–14. ISBN: 9781450367080. DOI: 10.1145/3313831.3376396. https://doi.org/10.1145/3313831.3376396.

[S32] Marc Erich Latoschik et al. “Not Alone Here?! Scalability and User Experience of Embodied Ambient Crowds in Distributed Social Virtual Reality”. In: *IEEE Transactions on Visualization and Computer Graphics* 25.5 (2019), pp. 2134–2144. DOI: 10.1109/TVCG.2019.2899250. https://doi.org/10.1109/TVCG.2019.2899250.

[S33] Sarah Lopez et al. "Investigating Implicit Gender Bias and Embodiment of White Males in Virtual Reality with Full Body Visuomotor Synchrony". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–12. ISBN: 9781450359702. https://doi.org/10.1145/3290605.3300787.

[S34] Divine Maloney, Guo Freeman, and Donghee Yvette Wohn. """Talking without a Voice": Understanding Non-Verbal Communication in Social Virtual Reality". In: *Proc. ACM Hum.-Comput. Interact.* 4.CSCW2 (Oct. 2020). DOI: 10.1145/3415246. https://doi.org/10.1145/3415246.

[S35] Soroosh Mashal, Gerold Hoelzl, and Matthias Kranz. "Valkyrie Project: Flying Immersion in Virtual Reality". In: *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. 2019, pp. 1–4. DOI: 10.1109/VS-Games.2019.8864545. https://doi.org/10.1109/VS-Games.2019.8864545.

[S36] Catarina Runa Miranda and Verónica Costa Orvalho. "Assessing Facial Expressions in Virtual Reality Environments." In: *VISIGRAPP (3: VISAPP)*. 2016, pp. 488–499. https://api.semanticscholar.org/CorpusID:26954834.

[S37] Jewoong Moon. "Reviews of Social Embodiment for Design of Non-Player Characters in Virtual Reality-Based Social Skill Training for Autistic Children". In: *Multimodal Technologies and Interaction* 2.3 (2018). ISSN: 2414-4088. DOI: 10.3390/mti2030053. https://www.mdpi.com/2414-4088/2/3/53.

[S38] Francesca Morganti. "Embodied Space in Natural and Virtual Environments: Implications for Cognitive Neuroscience Research". In: *Pervasive Computing Paradigms for Mental Health*. Ed. by Silvia Serino et al. Cham: Springer International Publishing, 2016, pp. 110–119. ISBN: 978-3-319-32270-4. https://doi.org/10.1007/978-3-319-32270-4_11.

[S39] Chontira Nimcharoen et al. "Is That Me?—Embodiment and Body Perception with an Augmented Reality Mirror". In: *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2018, pp. 158–163. DOI: 10.1109/ISMAR-Adjunct.2018.00057. https://doi.org/10.1109/ISMAR-Adjunct.2018.00057.

[S40] Juan Nino et al. "Enhancing Mobile VR Immersion: A Multimodal System of Neural Networks Approach to an IMU Gesture Controller". In: *Electronic Imaging* 2019.2 (2019), pp. 184-1-184–6. ISSN: 2470-1173. DOI: doi:10.2352/ISSN.2470-1173.2019.2.ERVR-184. https://www.ingentaconnect.com/content/ist/ei/2019/00002019/00000002/art00011.

[S41] Nami Ogawa, Takuji Narumi, and Michitaka Hirose. "Object Size Perception in Immersive Virtual Reality: Avatar Realism Affects the Way We Perceive". In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 647–648. DOI: 10.1109/VR.2018.8446318. https://doi.org/10.1109/VR.2018.8446318.

[S42] Ye Pan, David Sinclair, and Kenny Mitchell. "Empowerment and embodiment for collaborative mixed reality systems". In: *Computer Animation and Virtual Worlds* 29.3-4 (2018), e1838. https://doi.org/10.1002/cav.1838.

[S43] Mathias Parger et al. "Human Upper-Body Inverse Kinematics for Increased Embodiment in Consumer-Grade Virtual Reality". In: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. VRST '18. Tokyo, Japan: Association for Computing Machinery, 2018. ISBN: 9781450360869. DOI: 10.1145/3281505.3281529. https://doi.org/10.1145/3281505.3281529.

[S44] Federico Pianzola and Luca Deriu. "StoryVR: A Virtual Reality App for Enhancing Reading". In: *Methodologies and Intelligent Systems for Technology Enhanced Learning, 10th International Conference. Workshops*. Ed. by Zuzana Kubincová et al. Cham: Springer International Publishing, 2021, pp. 281–288. ISBN: 978-3-030-52287-2. https://doi.org/10.1007/978-3-030-52287-2_29.

[S45] Mirjana Prpa et al. "The Pulse Breath Water System: Exploring Breathing as an Embodied Interaction for Enhancing the Affective Potential of Virtual Reality". In: *Virtual, Augmented and Mixed Reality*. Ed. by Stephanie Lackey and Jessie Chen. Cham: Springer International Publishing, 2017, pp. 153–172. ISBN: 978-3-319-57987-0. https://doi.org/10.1007/978-3-319-57987-0_13.

[S46] Iulian Radu and Alissa Antle. "Embodied learning mechanics and their relationship to usability of handheld augmented reality". In: *2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual Augmented Reality (KELVAR)*. 2017, pp. 1–5. DOI: 10.1109/KELVAR.2017.7961561. https://doi.org/10.1109/KELVAR.2017.7961561.

[S47] Holger Regenbrecht et al. "Mixed Voxel Reality: Presence and Embodiment in Low Fidelity, Visually Coherent, Mixed Reality Environments". In: *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2017, pp. 90–99. DOI: 10.1109/ISMAR.2017.26. https://doi.org/10.1109/ISMAR.2017.26.

[S48] Nina Rosa et al. "Embodying an Extra Virtual Body in Augmented Reality". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 1138–1139. DOI: 10.1109/VR.2019.8798055. https://doi.org/10.1109/VR.2019.8798055.

[S49] Daniel Roth and Marc Erich Latoschik. “Construction of the Virtual Embodiment Questionnaire (VEQ)”. In: *IEEE Transactions on Visualization and Computer Graphics* 26.12 (2020), pp. 3546–3556. DOI: 10.1109/TVCG.2020.3023603. https://doi.org/10.1109/TVCG.2020.3023603.

[S50] Daniel Roth et al. “Technologies for Social Augmentations in User-Embodied Virtual Reality”. In: *25th ACM Symposium on Virtual Reality Software and Technology*. VRST ’19. Parramatta, NSW, Australia: Association for Computing Machinery, 2019. ISBN: 9781450370011. DOI: 10.1145/3359996.3364269. https://doi.org/10.1145/3359996.3364269.

[S51] Pejman Sajjadi et al. “Mixed or Virtual: Does Device Type Matter in Human-ECA Interactions”. In: *2020 6th International Conference of the Immersive Learning Research Network (iLRN)*. 2020, pp. 230–235. DOI: 10.23919/iLRN47897.2020.9155178. https://doi.org/10.23919/iLRN47897.2020.9155178.

[S52] Stephanie Schulze et al. “The Effects of Embodiment in Virtual Reality on Implicit Gender Bias”. In: *Virtual, Augmented and Mixed Reality. Multimodal Interaction*. Ed. by Jessie Y.C. Chen and Gino Fragomeni. Cham: Springer International Publishing, 2019, pp. 361–374. ISBN: 978-3-030-21607-8. https://doi.org/10.1007/978-3-030-21607-8_28.

[S53] Dong-Hee Shin. “The Role of Affordance in the Experience of Virtual Reality Learning”. In: *Telemat. Inf.* 34.8 (Dec. 2017), pp. 1826–1836. ISSN: 0736-5853. DOI: 10.1016/j.tele.2017.05.013. https://doi.org/10.1016/j.tele.2017.05.013.

[S54] Donghee Shin. “Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience?” In: *Computers in Human Behavior* 78 (2018), pp. 64–73. ISSN: 0747-5632. DOI: https://doi.org/10.1016/j.chb.2017.09.012. https://www.sciencedirect.com/science/article/pii/S0747563217305381.

[S55] Jose L. Soler, Manuel Contero, and Mariano Alcañiz. “VR Serious Game Design Based on Embodied Cognition Theory”. In: *Serious Games*. Ed. by Mariano Alcañiz et al. Cham: Springer International Publishing, 2017, pp. 12–21. ISBN: 978-3-319-70111-0. https://doi.org/10.1007/978-3-319-70111-0_2.

[S56] Erica Southgate. “Conceptualising Embodiment through Virtual Reality for Education”. In: *2020 6th International Conference of the Immersive Learning Research Network (iLRN)*. 2020, pp. 38–45. DOI: 10.23919/iLRN47897.2020.9155121. https://doi.org/10.23919/iLRN47897.2020.9155121.

[S57] Jason Tham et al. "Understanding Virtual Reality: Presence, Embodiment, and Professional Practice". In: *IEEE Transactions on Professional Communication* 61.2 (2018), pp. 178–195. DOI: 10.1109/TPC.2018.2804238. https://doi.org/10.1109/TPC.2018.2804238.

[S58] Yaqiong Xue et al. "MMRPet: Modular Mixed Reality Pet System Based on Passive Props". In: *Image and Graphics Technologies and Applications*. Ed. by Yongtian Wang, Qingmin Huang, and Yuxin Peng. Singapore: Springer Singapore, 2019, pp. 645–658. ISBN: 978-981-13-9917-6. https://doi.org/10.1007/978-981-13-9917-6_61.

[S59] Zhenjie Zhao, Feng Han, and Xiaojuan Ma. "A Live Storytelling Virtual Reality System with Programmable Cartoon-Style Emotion Embodiment". In: *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. 2019, pp. 102–1027. DOI: 10.1109/AIVR46125.2019.00024. https://doi.org/10.1109/AIVR46125.2019.00024.

[S60] Katja Zibrek and Rachel McDonnell. "Social Presence and Place Illusion Are Affected by Photorealism in Embodied VR". In: *Motion, Interaction and Games*. MIG '19. Newcastle upon Tyne, United Kingdom: Association for Computing Machinery, 2019. ISBN: 9781450369947. DOI: 10.1145/3359566.3360064. https://doi.org/10.1145/3359566.3360064.

Appendix A

Participant Information Sheet and Consent Form given to each person for the first experiment

UNIVERSITY OF Southampton

Participant Information Sheet

Study Title: Exploring the difference between two different mediums of play in a mixed reality game
Researcher: Yoan-Daniel Malinov
ERGO number: 54456

You are being invited to take part in the above research study. To help you decide whether you would like to take part or not, it is important that you understand why the research is being done and what it will involve. Please read the information below carefully and ask questions if anything is not clear or you would like more information before you decide to take part in this research. You may like to discuss it with others but it is up to you to decide whether or not to take part. If you are happy to participate you will be asked to sign a consent form.

What is the research about?
I am a PhD Computer Science student researching co-presence in and outside VR. I am conducting this study to see how different mediums of interacting and playing a game alongside a VR player (with a phone or a PC) affect the experience as a whole, how each player is affected by the other and to see if there is any notable differences that might want to be looked more into in the future. You should expect to be asked about your experience with the game that you would be trying together with the other participant, the awareness of the other player, how they affected you, etc.

Why have I been asked to participate?
You have been selected for this experiment as you wished to take part beforehand.

What will happen to me if I take part?

You would be playing a mixed reality game with another participant. You will be chosen to play the VR headset first or phone/PC first randomly.

The game is about an astronaut stuck in space and a scientist here on earth that controls a drone on the space station. Here the VR player would be the astronaut, and the scientist would be the person with the phone/PC. You are tasked from NASA with trying out this “new” system.

1. You will play the game with the other participant, which would take approximately 10 minutes. I would choose randomly if this play session would use the PC or the phone as second player.
    a. If you are using the phone, you will be given a phone with the game preloaded and a controller for you to be able to track inside the virtual world.
    b. If you are chosen to use the PC, you would be situated in front of the desktop and would be given a mouse and a keyboard to play the game.
    c. If you are using the headset, you will be given one controller together with the headset to play the game. Headphones are also going to be provided (you can use your own if you want).
2. I would be observing the playthrough of the game and noting down any interesting interactions in an empty sheet of paper, that would be later scanned and destroyed physically.
3. After finishing the game, you’ll be given a questionnaire about the experience and the usability, which would tale around 5 minutes, after which you would swap roles and play through again. This would take another 10 minutes.
4. When done, you will again complete the questionnaires specifically for the second experience – this would take around 5 minutes.
5. Finally, there is going to be a short interview with both players together where I will start asking you some questions about the experience, about how you felt, how the other player affected your experience, etc. I would be audio recording the interview, as this is the best way of saving what has been said. The audio recording would be then transcribed and

deleted, only the text being left. The recording is required and not optional so consent needs to be given in the Consent form.

6. After the end, you'll receive my gratitude for participating in this study and a £10 Amazon voucher.

The game itself would be logging in the background the time it takes for different tasks to be completed, and I use that just for statistics.

The questionnaires, paired with the observations and the interview would help me find if there is any difference between the two different modes – phone or PC. It would also help me find any interesting ideas that I haven't thought about through the interview.

Are there any benefits in my taking part?
There are no direct benefits to you for taking part in the study. By taking part however you will help expand knowledge different types of interaction in mixed modality games, in this case VR-Phone and VR-PC.

Are there any risks involved?
There is a small risk of motion sickness caused by the Virtual Reality headset. This risk however has been minimised as good practises for VR have been used when creating the game, like not moving the camera within the game and using minimal textures and objects to minimise lag that could cause motion sickness.
There is also a small risk of injury if the player with the headset hits the other player as they can't see what's happening in the physical world while they're playing the game. This risk however is also minimal, as they would see a virtual representation of the other player and will be advised to not make sudden movements.
If at any point you feel nauseous while wearing the headset you can immediately say so and I will come help remove the headset.

What data will be collected?
As part of this study, a signed form will be collected; the consent form will be digitised, and the hard copy version destroyed. These are personal to you and will be kept securely on a password protected university desktop that only I can access.

Data will be collected while playing the game, in the form of any interactions you would be having or any noticeable thing that I deem interesting. This would be written in an empty sheet of paper in the form of "Player 1 helped player 2 by using an in-game item", "Player 2 asked for help", etc. After that, there are going to be questionnaires about the game and the experience. The data would be pseudo-anonymised using your PIN that I would give you at the beginning of the experiment.

The game would also be logging the time it takes for different parts to be completed.

The interview is going to be audio recorded. The recording would be kept until it is transcribed by me, after which it would be deleted and only text kept. Any personally identifiable information would be removed, and I would be using your PIN.

All of this data would help see any differences between the two experiences, the two modes of play (phone and PC) and help me find interesting ideas that could be explored in the future.

Will my participation be confidential?
Your participation and the information we collect about you during the course of the research will be kept strictly confidential.

Only members of the research team and responsible members of the University of Southampton may be given access to data about you for monitoring purposes and/or to carry out an audit of the study to ensure that the research is complying with applicable regulations. Individuals from regulatory authorities (people who check that we are carrying out the study correctly) may require

access to your data. All of these people have a duty to keep your information, as a research participant, strictly confidential.
All audio recordings would be transcribed and pseudo-anonymised while doing so, and the audio would be destroyed.

Do I have to take part?
No, it is entirely up to you to decide whether or not to take part. If you decide you want to take part, you will need to sign a consent form to show you have agreed to take part.

If at any time you decide you do not wish to continue, please just let the me know.

What happens if I change my mind?
You have the right to change your mind and withdraw at any time without giving a reason and without your participant rights being affected.

If you withdraw from the study, I will keep the information about you that has already been obtained for the purposes of achieving the objectives of the study only.

What will happen to the results of the research?
Your personal details will remain strictly confidential. Research findings made available in any reports or publications will not include information that can directly identify you without your specific consent. The results of this study would be used in my 9-month report, as well as any paper that I might write for any conference or journal.
As stated above, your data will be pseudo-anonymised using PIN and the only people who would have access to your data would be me and my supervisor.

Where can I get more information?
You can contact me on my email address should you require more information: ydgm1g16@soton.ac.uk.

What happens if there is a problem?
If you have a concern about any aspect of this study, you should speak to the researchers who will do their best to answer your questions. You can either do that though email (ydgm1g16@soton.ac.uk) or by asking them before starting the study.
If you remain unhappy or have a complaint about any aspect of this study, please contact the University of Southampton Research Integrity and Governance Manager (023 8059 5058, rgoinfo@soton.ac.uk).

Data Protection Privacy Notice
The University of Southampton conducts research to the highest standards of research integrity. As a publicly-funded organisation, the University has to ensure that it is in the public interest when we use personally-identifiable information about people who have agreed to take part in research. This means that when you agree to take part in a research study, we will use information about you in the ways needed, and for the purposes specified, to conduct and complete the research project. Under data protection law, 'Personal data' means any information that relates to and is capable of identifying a living individual. The University's data protection policy governing the use of personal data by the University can be found on its website (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page).

This Participant Information Sheet tells you what data will be collected for this project and whether this includes any personal data. Please ask the research team if you have any questions or are unclear what data is being collected about you.

Our privacy notice for research participants provides more information on how the University of Southampton collects and uses your personal data when you take part in one of our research projects and can be found at
http://www.southampton.ac.uk/assets/sharepoint/intranet/ls/Public/Research%20and%20Integrity%20Privacy%20Notice/Privacy%20Notice%20for%20Research%20Participants.pdf

Any personal data we collect in this study will be used only for the purposes of carrying out our research and will be handled according to the University's policies in line with data protection law. If any personal data is used from which you can be identified directly, it will not be disclosed to anyone else without your consent unless the University of Southampton is required by law to disclose it.

Data protection law requires us to have a valid legal reason ('lawful basis') to process and use your Personal data. The lawful basis for processing personal information in this research study is for the performance of a task carried out in the public interest. Personal data collected for research will not be used for any other purpose.

For the purposes of data protection law, the University of Southampton is the 'Data Controller' for this study, which means that we are responsible for looking after your information and using it properly. The University of Southampton will keep identifiable information about you for **10 years** after the study has finished after which time any link between you and your information will be removed.

To safeguard your rights, we will use the minimum personal data necessary to achieve our research study objectives. Your data protection rights – such as to access, change, or transfer such information - may be limited, however, in order for the research output to be reliable and accurate. The University will not do anything with your personal data that you would not reasonably expect.

If you have any questions about how your personal data is used, or wish to exercise any of your rights, please consult the University's data protection webpage (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page) where you can make a request using our online form. If you need further assistance, please contact the University's Data Protection Officer (data.protection@soton.ac.uk).

Thank you.

CONSENT FORM

Study title: Exploring the effects of co-presence on players in a mixed reality game

Researcher name: Yoan-Daniel Malinov
ERGO number: 54456
Participant Identification Number:

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (*20200120 /V1*) and have had the opportunity to ask questions about the study.	
I agree to take part in this research project and agree for my data to be used for the purpose of this study.	
I understand my participation is voluntary and I may withdraw (at any time) for any reason without my participation rights being affected.	
I understand that taking part in the study involves audio recording *which will be transcribed and then destroyed* for the purposes set out in the participation information sheet.	
I agree to take part in the *interview* for the purposes set out in the participation information sheet and understand that these will be recorded using *audio.*	

Name of participant (print name)..

Signature of participant...

Date.. ....

Name of researcher (print name)..

Signature of researcher ....

Date...

Appendix B

Questionnaires and Interview questions in the first study

System Usability Scale

Participant Identification Number (PIN)

What did you just play on (Circle)?

VR **PC/PHONE**

Instructions: For each of the statements bellow, mark one box that best describes your experience with the game today.

	Strongly Disagree				Strongly Agree
1. I think that I would like to play this game frequently.	☐	☐	☐	☐	☐
2. I found the game unnecessarily complex.	☐	☐	☐	☐	☐
3. I thought the game was easy to play.	☐	☐	☐	☐	☐
4. I think that I would need the support of a technical person to be able to play this game.	☐	☐	☐	☐	☐
5. I found the various functions in this game were well integrated.	☐	☐	☐	☐	☐
6. I thought there was too much inconsistency in this game.	☐	☐	☐	☐	☐
7. I would imagine that most people would learn to play this game very quickly.	☐	☐	☐	☐	☐
8. I found the game very cumbersome to play.	☐	☐	☐	☐	☐
9. I felt very confident playing the game.	☐	☐	☐	☐	☐
10. I needed to learn a lot of things before I could get going with this game.	☐	☐	☐	☐	☐

Co-Presence and Immersion Survey

Participant Identification Number (PIN)

What did you just play on (Circle)?

VR **PC/PHONE**

Optional

Gender? ____________________

Age Group (Circle):

18-24 years old	25-34 years old	35-44 years old	45-54 years old	55 years or older

---

Instructions: For each of the statements bellow, mark one box that <u>best</u> describes your experience with the game today.

	Strongly Disagree						*Strongly Agree*
C1. I noticed the other player.	☐	☐	☐	☐	☐	☐	☐
C2. The other player noticed me.	☐	☐	☐	☐	☐	☐	☐
C3. The other player's presence was obvious to me.	☐	☐	☐	☐	☐	☐	☐
C4. My presence was obvious to the other player.	☐	☐	☐	☐	☐	☐	☐
C5. I was easily distracted from the other player when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C6. The other player was easily distracted from me when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C7. My thoughts were clear to the other player.	☐	☐	☐	☐	☐	☐	☐
C8. The other player's thoughts were clear to me.	☐	☐	☐	☐	☐	☐	☐

	Strongly Disagree						*Strongly Agree*
I1. In the computer world I had a sense of "being there".	☐	☐	☐	☐	☐	☐	☐
I2. I wasn't aware at all of the real world surrounding me while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?	☐	☐	☐	☐	☐	☐	☐
I3. I was not aware of my real environment.	☐	☐	☐	☐	☐	☐	☐
I4. I still paid attention to the real environment.	☐	☐	☐	☐	☐	☐	☐
I5. I was completely captivated by the virtual world.	☐	☐	☐	☐	☐	☐	☐
I6. I had a sense of acting in the virtual space, rather than operating something from outside.	☐	☐	☐	☐	☐	☐	☐
I7. Somehow, I felt that the virtual world surrounded me.	☐	☐	☐	☐	☐	☐	☐

Questions for interview:

Co-presence:

1. Did you feel alone in the play space?
2. How aware were you that the other person was in the same place as you? Did it feel more like they're in the virtual space or more like they're outside?
3. Did you use any type of physical and/or body cues (as in pointing towards an object with a hand, using physical touch to help somebody find something, etc.) in the interaction?
    a. Did you feel like you could use them and still convey some sort of meaning?
    b. Did you communicate through any other channels other than verbal?
4. Did you feel like the other person was there with you in the virtual space and in the real world?
5. Was there a point in the game where you were thinking about the other player's point of view?

Immersion:

1. Did you have a sense of "being there" in the computer-generated world?
    a. In what ways did you feel the sense of "being there"?
2. How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?
    a. How did the other player affect that awareness?
    b. Did thinking about the other player make you less or more aware?
3. Were you aware of the real environment around you?
4. In what ways did the virtual world captivate you?

Appendix C

Participant Information Sheet and Consent Form given to each person for the second experiment

Participant Information Sheet

Study Title: Exploring the difference between two different mediums of play in a mixed reality game

Researcher: Yoan-Daniel Malinov

ERGO number: 54456.A2 V4

You are being invited to take part in the above research study. To help you decide whether you would like to take part or not, it is important that you understand why the research is being done and what it will involve. Please read the information below carefully and ask questions if anything is not clear or you would like more information before you decide to take part in this research. You may like to discuss it with others but it is up to you to decide whether or not to take part. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

I am a PhD Computer Science student researching multi-player interactions in and outside VR. I am conducting this study to see how different mediums of interacting and playing a game alongside a VR player (with a phone or a PC) affect the experience as a whole, how each player is affected by the other and to see if there is any notable differences that might want to be looked more into in the future.

You should expect to be asked about your experience with the game that you would be trying together with the other participant, the awareness of the other player, how they affected you, etc.

Why have I been asked to participate?

You have been selected for this experiment as you wished to take part beforehand.

What will happen to me if I take part?

You would be playing a mixed reality game with another participant. You will be chosen to play the VR headset first or phone/PC first randomly.

Before participation, any surfaces and equipment would be disinfected and sanitised according to University policy. You will also be asked to wear a mask throughout the duration of the study, unless exempt, in accordance with the university's COVID-19 policy.

For more information, you can check the "**COVID-19 (Coronavirus) guidance for research involving human participants – Non-Clinical, v3.1**" at

https://sotonac.sharepoint.com/teams/RIS/SitePages/COVID-19-response.aspx or ask me a copy of the document.

The room would have been ventilated between each study as well.

As the main researcher I would also be tested weekly using the University's Saliva Testing Programme and take all necessary measures to minimise the risk for you when participating in the study. If you have any questions about the steps that have been undertaken you can ask me in person or email me, and I would be happy to answer any worries.

The game is about an astronaut stuck in space and a scientist here on earth that controls a drone on the space station. Here the VR player would be the astronaut, and the scientist would be the person with the phone/PC. You are tasked from NASA with trying out this "new" system.

1. You will play the game with the other participant, which would take approximately 10 minutes. I would choose randomly if this play session would use the PC or the phone as second player.
    a. If you are using the phone, you will be given a phone with the game preloaded and a controller for you to be able to track inside the virtual world.
    b. If you are chosen to use the PC, you would be situated in front of the desktop and would be given a mouse and a keyboard to play the game.
    c. If you are using the headset, you will be given one controller together with the headset to play the game. Headphones are also going to be provided (you can use your own if you want).
2. I would be observing the playthrough of the game and noting down any interesting interactions in an empty sheet of paper, that would be later scanned and destroyed physically.
3. A video recording of the play sessions is going to be made for purposes of easier analysis of any interactions between the two players later. The video itself is only being made for analysis purposes and is **not** going to be shared outside the research team (my supervisors and myself) and current project. It's going to be kept for the duration of my degree.
4. After finishing the game, you'll be given a questionnaire about the experience and the usability, which would take around 5 minutes, after which you would swap roles and play through again. This would take another 10 minutes.
5. When done, you will again complete the questionnaires specifically for the second experience – this would take around 5 minutes.
6. Finally, there is going to be a short interview with both players together where I will start asking you some questions about the experience, about how you felt, how the other player affected your experience, etc. I would be audio recording the interview, as this is the best way of saving what has been said. The audio recording would be then transcribed and deleted, only the text being left. The recording is required and not optional so consent needs to be given in the Consent form.
7. After the end, you'll receive my gratitude for participating in this study and a £10 banknote.

The game itself would be logging in the background the time it takes for different tasks to be completed, and I use that just for statistics.

The questionnaires, paired with the observations and the interview would help me find if there is any difference between the two different modes – phone or PC, and how different strategies for embodiment affect the experience. It would also help me find any interesting ideas that I haven't thought about through the interview.

Are there any benefits in my taking part?

There are no direct benefits to you for taking part in the study. By taking part however you will help

expand knowledge different types of interaction in mixed modality games, in this case VR-Phone and VR-PC.

Are there any risks involved?

There is a small risk of motion sickness caused by the Virtual Reality headset. This risk however has been minimised as good practises for VR have been used when creating the game, like not moving the camera within the game and using minimal textures and objects to minimise lag that could cause motion sickness.

There is also a small risk of injury if the player with the headset hits the other player as they can't see what's happening in the physical world while they're playing the game. This risk however is also minimal, as they would see a virtual representation of the other player and will be advised to not make sudden movements.

If at any point you feel nauseous while wearing the headset you can immediately say so and I will come help remove the headset.

As this is an in-person study and there's a pandemic, there's always a risk of contagion - steps to mitigate such risk have been taken, such as every surface and equipment being used would be disinfected after each playthrough, every participant would be screened beforehand about COVID-19 and asked if they've had any symptoms or contact with any person who was positive. Participants would be encouraged to wear a face mask while in the room according to University policy (linked above), and the room would be ventilated after each experiment.

What data will be collected?

As part of this study, a signed form will be collected; the consent form will be digitised, and the hard copy version destroyed. These are personal to you and will be kept securely on a password protected university desktop that only I can access.

Data will be collected while playing the game, in the form of any interactions you would be having or any noticeable thing that I deem interesting. This would be written in an empty sheet of paper in the form of "Player 1 helped player 2 by using an in-game item", "Player 2 asked for help", etc. A video recording is also going to be taken for both play sessions (only for analysis purposes, not going to be shared via social media), and it's not going to be retained beyond conferment of my degree.

After that, there are going to be questionnaires about the game and the experience. The data would be pseudo-anonymised using your PIN that I would give you at the beginning of the experiment.

The game would also be logging the time it takes for different parts to be completed.

The interview is going to be audio recorded. The recording would be kept until it is transcribed by me, after which it would be deleted and only text kept. Any personally identifiable information would be removed, and I would be using your PIN.

All of this data would help see any differences between the two experiences, the two modes of play (phone and PC) and help me find interesting ideas that could be explored in the future.

Will my participation be confidential?

Your participation and the information we collect about you during the course of the research will be kept strictly confidential.

Only members of the research team and responsible members of the University of Southampton may be given access to data about you for monitoring purposes and/or to carry out an audit of the study to ensure that the research is complying with applicable regulations. Individuals from regulatory authorities (people who check that we are carrying out the study correctly) may require access to your data. All of these people have a duty to keep your information, as a research participant, strictly confidential.

All audio recordings would be transcribed and pseudo-anonymised while doing so, and the audio would be destroyed.

All recorded videos will be kept secure on a password protected University machine until the end of my PhD, and according to the guidelines for data handling in the **Data Protection Privacy Notice** section below.

Do I have to take part?

No, it is entirely up to you to decide whether or not to take part. If you decide you want to take part, you will need to sign a consent form to show you have agreed to take part.

If at any time you decide you do not wish to continue, please just let the me know.

What happens if I change my mind?

You have the right to change your mind and withdraw at any time without giving a reason and without your participant rights being affected.

If you withdraw from the study, I will keep the information about you that has already been obtained for the purposes of achieving the objectives of the study only.

What will happen to the results of the research?

Your personal details will remain strictly confidential. Research findings made available in any reports or publications will not include information that can directly identify you without your specific consent. The results of this study would be used in my final thesis, as well as any paper that I might write for any conference or journal.

As stated above, your data will be pseudo-anonymised using PIN and the only people who would have access to your data would be me and my supervisor.

Where can I get more information?

You can contact me on my email address should you require more information: ydgm1g16@soton.ac.uk.

What happens if there is a problem?

If you have a concern about any aspect of this study, you should speak to the researchers who will do their best to answer your questions. You can either do that though email (ydgm1g16@soton.ac.uk) or by asking them before starting the study.

If you remain unhappy or have a complaint about any aspect of this study, please contact the University of Southampton Research Integrity and Governance Manager (023 8059 5058, rgoinfo@soton.ac.uk).

Data Protection Privacy Notice

The University of Southampton conducts research to the highest standards of research integrity. As a publicly-funded organisation, the University has to ensure that it is in the public interest when we use personally-identifiable information about people who have agreed to take part in research. This means that when you agree to take part in a research study, we will use information about you in the ways needed, and for the purposes specified, to conduct and complete the research project. Under data protection law, ‘Personal data’ means any information that relates to and is capable of identifying a living individual. The University’s data protection policy governing the use of personal data by the University can be found on its website (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page).

This Participant Information Sheet tells you what data will be collected for this project and whether this includes any personal data. Please ask the research team if you have any questions or are unclear what data is being collected about you.

Our privacy notice for research participants provides more information on how the University of Southampton collects and uses your personal data when you take part in one of our research projects and can be found at http://www.southampton.ac.uk/assets/sharepoint/intranet/ls/Public/Research%20and%20Integrity%20Privacy%20Notice/Privacy%20Notice%20for%20Research%20Participants.pdf

Any personal data we collect in this study will be used only for the purposes of carrying out our research and will be handled according to the University's policies in line with data protection law. If any personal data is used from which you can be identified directly, it will not be disclosed to anyone else without your consent unless the University of Southampton is required by law to disclose it.

Data protection law requires us to have a valid legal reason ('lawful basis') to process and use your Personal data. The lawful basis for processing personal information in this research study is for the performance of a task carried out in the public interest. Personal data collected for research will not be used for any other purpose.

For the purposes of data protection law, the University of Southampton is the 'Data Controller' for this study, which means that we are responsible for looking after your information and using it properly. The University of Southampton will keep identifiable information about you for **10 years** after the study has finished after which time any link between you and your information will be removed.

To safeguard your rights, we will use the minimum personal data necessary to achieve our research study objectives. Your data protection rights – such as to access, change, or transfer such information - may be limited, however, in order for the research output to be reliable and accurate. The University will not do anything with your personal data that you would not reasonably expect.

If you have any questions about how your personal data is used, or wish to exercise any of your rights, please consult the University's data protection webpage (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page) where you can make a request using our online form. If you need further assistance, please contact the University's Data Protection Officer (data.protection@soton.ac.uk).

Thank you.

CONSENT FORM

Study title: Exploring the effects of co-presence on players in a mixed reality game

Researcher name: Yoan-Daniel Malinov
ERGO number: 54456.A2
Participant Identification Number*:*

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (*27022022 /V4*) and have had the opportunity to ask questions about the study.	
I agree to take part in this research project and agree for my data to be used for the purpose of this study.	
I understand my participation is voluntary and I may withdraw (at any time) for any reason without my participation rights being affected.	
I understand that taking part in the study involves audio recording *which will be transcribed and then destroyed* for the purposes set out in the participation information sheet.	
I understand that taking part in the study involves video recordings *which will be kept secure on a password protected university machine* for the purposes set out in the participation information sheet.	
I agree to take part in the *interview* for the purposes set out in the participation information sheet and understand that these will be recorded using *audio*.	

Name of participant (print name)..

Signature of participant...

Date.. ....

Name of researcher (print name)..

Signature of researcher ....

Appendix D

Embodiment Questionnaires used in the final study

Embodiment Survey – ASTRONAUT – Emergency lights off

Participant Identification Number (PIN)

What did you just play on (Circle)?

VR **Non-VR (PC/PHONE)**

Optional

Gender? ____________________

Age Group (Circle):

18-24 years old	25-34 years old	35-44 years old	45-54 years old	55 years or older

If you have any experience with VR, how often do you use VR:

Daily	Weekly	Monthly	A few times per year	Rarely to never

Instructions: In the first part of the game where the **emergency lights were off** and the ship had power, consider your answers to the following statements and mark the on that best describes your experience.

	Strongly Disagree						*Strongly Agree*
BO1. I felt as if the virtual avatar (astronaut) was part of my body.	☐	☐	☐	☐	☐	☐	☐
BO2. I felt as if the virtual avatar (astronaut) was part of someone else.	☐	☐	☐	☐	☐	☐	☐
BO3. It seemed as if I might have more than one body.	☐	☐	☐	☐	☐	☐	☐
A1. It felt like I could control the virtual body as if it was my own body/avatar.	☐	☐	☐	☐	☐	☐	☐
A2. The movements of the virtual body were caused by my movements.	☐	☐	☐	☐	☐	☐	☐
A3. I felt as if the movements of the virtual body/avatar were influencing my own movements.	☐	☐	☐	☐	☐	☐	☐
A4. I felt as if the virtual body was moving by itself.	☐	☐	☐	☐	☐	☐	☐
SL1. I felt as if my body was located where I saw the virtual body.	☐	☐	☐	☐	☐	☐	☐

TURN OVER PAGE

	Strongly Disagree						*Strongly Agree*
SL2. In the computer world I had a sense of “being there”.	☐	☐	☐	☐	☐	☐	☐
C1. I noticed the other player.	☐	☐	☐	☐	☐	☐	☐
C2. The other player noticed me.	☐	☐	☐	☐	☐	☐	☐
C3. The other player’s presence was obvious to me.	☐	☐	☐	☐	☐	☐	☐
C4. My presence was obvious to the other player.	☐	☐	☐	☐	☐	☐	☐
C5. I was easily distracted from the other player when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C6. The other player was easily distracted from me when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C7. My thoughts were clear to the other player.	☐	☐	☐	☐	☐	☐	☐
C8. The other player’s thoughts were clear to me.	☐	☐	☐	☐	☐	☐	☐

Embodiment Survey – ASTRONAUT – Emergency lights on

Participant Identification Number (PIN)

What did you just play on (Circle)?

VR **Non-VR (PC/PHONE)**

Optional

Gender? ____________________

Age Group (Circle):

18-24 years old	25-34 years old	35-44 years old	45-54 years old	55 years or older

Instructions: In the second part of the game where the **emergency lights were on** and the ship lost power, consider your answers to the following statements and mark the on that best describes your experience.

	Strongly Disagree						*Strongly Agree*
BO1. I felt as if the virtual avatar was part of my body.	☐	☐	☐	☐	☐	☐	☐
BO2. I felt as if the virtual avatar was part of someone else.	☐	☐	☐	☐	☐	☐	☐
BO3. It seemed as if I might have more than one body/avatar.	☐	☐	☐	☐	☐	☐	☐
A1. It felt like I could control the virtual body as if it was my own body/avatar.	☐	☐	☐	☐	☐	☐	☐
A2. The movements of the virtual body were caused by my movements.	☐	☐	☐	☐	☐	☐	☐
A3. I felt as if the movements of the virtual body/avatar were influencing my own movements.	☐	☐	☐	☐	☐	☐	☐
A4. I felt as if the virtual body was moving by itself.	☐	☐	☐	☐	☐	☐	☐
SL1. I felt as if my body was located where I saw the virtual body.	☐	☐	☐	☐	☐	☐	☐

TURN OVER PAGE

	Strongly Disagree						*Strongly Agree*
SL2. In the computer world I had a sense of "being there".	☐	☐	☐	☐	☐	☐	☐
C1. I noticed the other player.	☐	☐	☐	☐	☐	☐	☐
C2. The other player noticed me.	☐	☐	☐	☐	☐	☐	☐
C3. The other player's presence was obvious to me.	☐	☐	☐	☐	☐	☐	☐
C4. My presence was obvious to the other player.	☐	☐	☐	☐	☐	☐	☐
C5. I was easily distracted from the other player when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C6. The other player was easily distracted from me when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C7. My thoughts were clear to the other player.	☐	☐	☐	☐	☐	☐	☐
C8. The other player's thoughts were clear to me.	☐	☐	☐	☐	☐	☐	☐

Embodiment Survey – DRONE – Emergency lights off

Participant Identification Number (PIN)

What did you just play on (Circle)?

VR **Non-VR (PC/PHONE)**

Optional

Gender? ____________________

Age Group (Circle):

18-24 years old	25-34 years old	35-44 years old	45-54 years old	55 years or older

Instructions: In the first part of the game where the **emergency lights were off** and the ship had power, consider your answers to the following statements and mark the on that best describes your experience.

	Strongly Disagree						*Strongly Agree*
BO1. I felt as if the drone was part of my body.	☐	☐	☐	☐	☐	☐	☐
BO2. I felt as if the drone was part of someone else.	☐	☐	☐	☐	☐	☐	☐
BO3. It seemed as if I might have more than one body/avatar.	☐	☐	☐	☐	☐	☐	☐
A1. It felt like I could control the drone as if it was my own body/avatar.	☐	☐	☐	☐	☐	☐	☐
A2. The movements of the drone were caused by my movements.	☐	☐	☐	☐	☐	☐	☐
A3. I felt as if the movements of the drone were influencing my own movements.	☐	☐	☐	☐	☐	☐	☐
A4. I felt as if the drone was moving by itself.	☐	☐	☐	☐	☐	☐	☐
SL1. I felt as if my body was located where I saw the drone.	☐	☐	☐	☐	☐	☐	☐

TURN OVER PAGE

	Strongly Disagree						*Strongly Agree*
SL2. In the computer world I had a sense of "being there".	☐	☐	☐	☐	☐	☐	☐
C1. I noticed the other player.	☐	☐	☐	☐	☐	☐	☐
C2. The other player noticed me.	☐	☐	☐	☐	☐	☐	☐
C3. The other player's presence was obvious to me.	☐	☐	☐	☐	☐	☐	☐
C4. My presence was obvious to the other player.	☐	☐	☐	☐	☐	☐	☐
C5. I was easily distracted from the other player when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C6. The other player was easily distracted from me when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C7. My thoughts were clear to the other player.	☐	☐	☐	☐	☐	☐	☐
C8. The other player's thoughts were clear to me.	☐	☐	☐	☐	☐	☐	☐

Embodiment Survey – DRONE – Emergency lights on

Participant Identification Number (PIN)

What did you just play on (Circle)?

VR **Non-VR (PC/PHONE)**

Optional

Gender? ___________________

Age Group (Circle):

18-24 years old	25-34 years old	35-44 years old	45-54 years old	55 years or older

Instructions: In the second part of the game where the **emergency lights were on** and the ship lost power, consider your answers to the following statements and mark the on that best describes your experience.

	Strongly Disagree						*Strongly Agree*
BO1. I felt as if the drone was part of my body.	☐	☐	☐	☐	☐	☐	☐
BO2. I felt as if the drone was part of someone else.	☐	☐	☐	☐	☐	☐	☐
BO3. It seemed as if I might have more than one body/avatar.	☐	☐	☐	☐	☐	☐	☐
A1. It felt like I could control the drone as if it was my own body/avatar.	☐	☐	☐	☐	☐	☐	☐
A2. The movements of the drone were caused by my movements.	☐	☐	☐	☐	☐	☐	☐
A3. I felt as if the movements of the drone were influencing my own movements.	☐	☐	☐	☐	☐	☐	☐
A4. I felt as if the drone was moving by itself.	☐	☐	☐	☐	☐	☐	☐
SL1. I felt as if my body was located where I saw the drone.	☐	☐	☐	☐	☐	☐	☐

TURN OVER PAGE

	Strongly Disagree						*Strongly Agree*
SL2. In the computer world I had a sense of "being there".	☐	☐	☐	☐	☐	☐	☐
C1. I noticed the other player.	☐	☐	☐	☐	☐	☐	☐
C2. The other player noticed me.	☐	☐	☐	☐	☐	☐	☐
C3. The other player's presence was obvious to me.	☐	☐	☐	☐	☐	☐	☐
C4. My presence was obvious to the other player.	☐	☐	☐	☐	☐	☐	☐
C5. I was easily distracted from the other player when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C6. The other player was easily distracted from me when other things were going on.	☐	☐	☐	☐	☐	☐	☐
C7. My thoughts were clear to the other player.	☐	☐	☐	☐	☐	☐	☐
C8. The other player's thoughts were clear to me.	☐	☐	☐	☐	☐	☐	☐

Questions for interview:

Embodied Interactions:

1. Did the zoom (mouse wheel or pinch to zoom) feel more natural?
2. For VR player: Did you feel any different when the drone’s zoom “broke”? Did it feel less immersive?
3. Did it feel faster when the drone player was using the embodied interaction?

Audio Streaming:

1. Did you hear the other person through the headphones?
2. Was it harder to hear with or without the audio in the headset?
3. For VR player: Did the spatial audio help you have an idea where the other player was?