VR Gaming - Hands On

The use and effects of bare hand gestures as an interaction method in multiplayer Virtual Reality Games.

Author Abraham Georgiadis
Supervisor Shahrouz Yousefi
Examiner Ilir Jusufi
Exam date 31 May 2017
Subject Social Media and Web Technologies
Level Master
Course code 5ME11E
Abstract

The field of virtual reality (VR) is getting increasing attention from the scientific community and it is being portrayed by advertisements as the user interface (UI) of the future. This is a fair statement since the prior uses of VR that used to exist only in fiction movies and books, are now widely available in many forms and settings to the public. One of the most interesting outcomes from this technological evolution is that now VR can be experienced through the use of a mobile phone and the addition of some inexpensive means typically in a form of a headset. The combination of the phone’s screen as attached to the headset, creates a form of Head Mounted Display (HMD) which can be utilized in order for the user to be immersed within a virtual environment (VE). The argument here is that even if the means to get access to VR are cheap, this should not be the case with the experience as well. On the contrary, the low entry requirements in combination with a high quality experience are the basis for the medium’s success and further adoption by the users. More specifically, the capability of utilizing a three dimensional space (3D) should not limit the medium’s use on just that but instead, this space should be used in order to offer immersive environments which make the user feel as if he is there.

There are many factors that contribute to that result and significant progress has been made to some such as the quality of screen or other hardware parts that allow the user get immersed into the virtual scenery, however, little progress has been made towards the conceptual means that allow the user of better experiencing this VE. Most of the VR applications so far are specifically designed for a single user session. This creates an isolation of the user from any other type of communities which further increases the stigma of VR being a solitary experience. Another issue is the interaction method that is available to users in order to interact with the VE. The use of buttons in most of the available headsets is a counter intuitive method for a person to interact with an environment that wants to be called real. The technological advancements in the field of image processing have resulted in many new methods of interaction and multimodal manipulation within VE and it would be worthy of exploring their effects on the user experience (UX) when used as an interaction method.

For these reasons, this thesis used the case of VR games as a setting to study how UX can be enhanced from its current state by introducing a bare hand gesture interaction method and expanding the VR setting in order to host two users in shared VE. Two individual studies were conducted where user feedback was collected in order to describe the effects of this approach in both a qualitative and quantitative manner. As results indicate, by utilizing gesture analysis on a headset equipped with a smartphone, it is possible to offer a natural and engaging solution of VR interaction capable of rich UXs while maintaining a low entry level for the end users. Finally, the addition of another player, significantly affected the experience by influencing the emotional state of the participants in the game and further enforcing their feeling of presence within the VE.

Keywords. VR, Hand Gestures, Bare-hand interaction, VR gaming, Multiplayer VR, HCI, Usability analysis
Acknowledgements

Many were the times that I asked myself why I am writing this thesis and the answer was always the same. For the others. For the others that helped me get here and for the others that will follow after me. Because it is as the words of Isaac Newton saying “If I have seen further, it is by standing on the shoulders of giants” the way I was able to know more. For this reason, I would like to thank those who helped me in their own way throughout the different stages of this journey.

First of all, I would like to thank my teacher Shahrouz as his guidance and constant support were the means that drove me in every step of the road but most importantly in the moments of need. I had little respect towards the academia, but your advices were an inspiration to me and the reason to understand how wrong I was. Also, I would like to thank, Nuno Otero, for his lectures on pedagogy and ethics of technology. You made the LNU experience so much interesting. Moreover, I would like to thank Ola Petersson and Hans Lundberg from the Computer Science department for their Agile Development course. I will always remember your quote “Create it because it brings value to someone, not because you simply can.”. I would like also to give my special gratitude to the Manomotion family. You trusted me with your technology and sparked my creativity to use it in meaningful ways. But most importantly, for making every day of this process to be fun while working alongside you guys. I also feel the need to express my gratitude towards Helena Sundlöf for helping me organize the study settings as well as all the people that participated in it. I owe my gratitude to both of my friends, Mohamad and Julio. Thank you guys for your honest and thoughtful feedback but most importantly for your friendship, I hope I can pay it back. Finally, to my beloved parents George, Despoina and Daria. Your love minded concern to understand what I was working on, made me feel so special and also get a different perspective of how technology appeals to other people.

Thank you all, this thesis would not have been the same without your help.
# Contents

1 Introduction ................................................................. 7  
  1.2 Motivation .............................................................. 8  
  1.3 Research questions ................................................... 9  
  1.4 Approach ................................................................. 10  
  1.5 Contribution ............................................................ 11  
  1.6 Thesis structure ....................................................... 11  

2 Theoretical framework .................................................... 12  
  2.1 Core concepts .......................................................... 12  
    2.1.1 VR ................................................................. 12  
    2.1.2 Immersion and presence ........................................ 14  
    2.1.3 Gesture based interaction ..................................... 14  
    2.1.4 Games ............................................................ 15  
  2.2 Related work ............................................................ 17  
    2.2.1 Gesture-based interaction technologies ...................... 17  
    2.2.2 Shared virtual spaces and multiplayer VR games .......... 19  

3 Methodology ..................................................................... 20  
  3.1 Research design ......................................................... 20  
  3.2 Studies sample ........................................................... 21  
  3.3 Data collection process and methods ............................... 21  
    3.3.1 Study 1 - Game controllers ................................... 21  
    3.3.2 Study 2 - Presence of an additional user .................. 24  
  3.4 Statistical treatment - Data analysis methods ................... 24  

4 Prototype ...................................................................... 26  
  4.1 Game scenario ............................................................ 26  
  4.2 Game mechanics .......................................................... 27  
  4.3 Interaction methods ..................................................... 28  
  4.4 Implementation ........................................................... 30  
    4.4.1 Expert Validation ............................................... 32  

5 Results and analysis ...................................................... 33  
  5.1 User study 1 - Interaction Methods NASA - TLX ................ 33  
  5.2 User study 1 - Interaction Methods Questionnaire ............ 34  
  5.3 User study 1 - Interaction Methods interview ................... 48  
  5.4 User study 1 - Interaction methods, choice of controller. .... 50  
  5.5 User study 1 - Interaction Methods, additional observations. 50  
  5.6 User study 2 - Multiplayer session, interview ................... 51  
  5.7 User study 2 - Multiplayer session, additional observations. 53
6 Discussion
   6.1 Interaction methods session ........................................ 54
   6.2 Multiplayer session .................................................. 55
   6.3 Research questions .................................................. 56

7 Conclusion
   7.1 Summary ................................................................. 58
   7.2 Limitations ............................................................. 59
   7.3 Future work ............................................................. 59

References ................................................................. 61

Appendices ............................................................... 64

A Methodological tools .................................................. 65
   A.1 Interaction methods questionnaire ................................. 66
   A.2 NASA - TLX ........................................................... 70

B Additional results ...................................................... 71
   B.1 Numerical input used for the word histograms. ............... 71
# List of Figures

2.1 VR means and settings. ................................................................. 13  
2.2 Formats of VR experience stated by L.E.K (Super VR, Medium VR, Casual Mobile VR) .................................................. 13  
2.3 Gestural taxonomy that describes and differentiates the hand gestures. ................................................................. 15  
2.4 Visual representation of Crawford’s game definition. ................................................................. 16  
2.5 Game market forecast (2015-2019) and distribution to different mediums. ................................................................. 17  
2.6 Biomechanical means for gesture recognition in VR experience. ................................................................. 18  
2.7 Computer vision methods for hand and gesture recognition. ................................................................. 19  
3.1 Users on synthesized background, game methods (Gesture,Gesture,Gesture,Touchpad). ................................................................. 22  
3.2 Users participating in multiplayer VR game. ................................................................. 24  
4.1 Top view of the game map. ................................................................. 27  
4.2 View from the networked camera of the player avatar supplying a depleted tower. ................................................................. 28  
4.3 Samsung’s GearVR touchpad and tap method. ................................................................. 29  
4.4 Hand representation transferred in the virtual world. ................................................................. 30  
4.5 Sequence of hand poses to perform Click gesture as defined by Manomotion. ................................................................. 30  
4.6 Networked game system logic. ................................................................. 32  
5.1 TLX histogram.Experience load as perceived by users ................................................................. 33  
5.2 Histogram of general system responsiveness as perceived by users in both interaction methods (Q1). ................................................................. 34  
5.3 Histogram of user awareness of the external environment, while in the VR experience (Q2). ................................................................. 35  
5.4 Histogram of interaction methods experience as perceived by users (Q3). ................................................................. 36  
5.5 Histogram of user awareness of the display and control mechanisms (Q4). ................................................................. 37  
5.6 Histogram of information consistency as perceived by the user senses (Q5). ................................................................. 38  
5.7 System responsiveness derived interaction method responsiveness as perceived by user actions (Q6). ................................................................. 39  
5.8 Histogram of interaction based, experience quality within the virtual environment (Q7) ................................................................. 40  
5.9 Histogram of participation evaluation as perceived by users within the experience (Q8). ................................................................. 41  
5.10 Histogram of evaluation regarding the distracting nature of the control mechanism (Q9). ................................................................. 42  
5.11 Histogram of user evaluation regarding the delay as perceived in the VR experience (Q10). ................................................................. 43  
5.12 Histogram of transition ease in the VR experience affected by the interaction method (Q11). ................................................................. 44  
5.13 Histogram of interaction methods interference with the game tasks (Q12). ................................................................. 45  
5.14 Histogram of interaction methods interfering in the decision making process (Q13). ................................................................. 46  
5.15 Histogram of post game, interaction method proficiency (Q14). ................................................................. 47  
5.16 Wordcloud from the most frequent words users used in the touchpad method. ................................................................. 48  
5.17 Wordcloud from the most frequent words users used in the gesture method. ................................................................. 49
5.18 Major and minor themes emerged from user’s multiplayer session interview. . . . . . . 52

A.1 Interaction methods questionnaire demographics. . . . . . . . . . . . . . . . . . . . . . 66
A.2 Interaction methods questionnaire - UX questions 1. . . . . . . . . . . . . . . . . . . . . . 67
A.3 Interaction methods questionnaire - UX questions 2. . . . . . . . . . . . . . . . . . . . . . 68
A.4 Interaction methods questionnaire - UX questions 3. . . . . . . . . . . . . . . . . . . . . . 69
A.5 NASA Task Load Index questionnaire. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70
List of Tables

5.1 Statistical values of general system responsiveness as perceived by users in both interaction methods (Q1). ................................................................. 34
5.2 Statistical values of the user awareness of the external environment, while in the VR experience (Q2). ................................................................. 35
5.3 Statistical values of interaction methods experience as perceived by users (Q3). .... 36
5.4 Statistical values of the User awareness of of the display and control mechanisms (Q4). 37
5.5 Statistical values of information consistency as perceived by the user senses (Q5). .... 38
5.6 Statistical values of interaction based, experience quality within the virtual environment (Q7) ................................................................. 40
5.7 Statistical values of participation evaluation as perceived by users within the experience (Q8) ................................................................. 41
5.8 Statistical values of evaluation regarding the distracting nature of the control mechanism (Q9). ................................................................. 42
5.9 Statistical values of user evaluation regarding the delay as perceived in the VR experience (Q10). ................................................................. 43
5.10 Statistical values of transition ease in the VR experience affected by the interaction method (Q11). ................................................................. 44
5.11 Statistical values of interaction methods interference with the game tasks (Q12). .. 45
5.12 Statistical values of interaction methods interfering in the decision making process (Q13). 46
5.13 Statistical values of post game, interaction method proficiency (Q14). .................. 47

B.1 Top 10 words users used for Button Tap interaction method. ......................... 71
B.2 Top 10 words users used for Gesture interaction method. ........................... 72
Glossary

3D Three dimensional.

ASL American Sign Language.

C.A.V.E Cave Automatic Virtual Environments.

HDM Head Mounted Display.

LNU Linnaeus University.

PC Personal Computer.

RQ Research Question.

SM Social Media.

TD Tower Defense.

TV Television.

UI User Interface.

UX User Experience.

VE Virtual Environment.

VR Virtual Reality.
Chapter 1

Introduction

Even though prominent in our era, the concept of VR experiences originates back in the 60s where Sutherland (1968) demonstrated a mechanical and computer aided system that was able to show changing perspective images through a HMD. This idea of VEs, became the subject of study by many researchers as the degree of freedom in a VE can bypass the material limitations of the real world. This allows the simulation of many scenarios that would otherwise be impossible to implement or be highly demanding in resources. Another great advantage of this approach is the ability to utilize the three dimensional (3D) space in contrast with other mediums such as a mobile phone or a personal computer (PC), where the environment and interactions are taking place in two dimensions. This expands the threshold in the variable of realism within the experience, however, it remains a challenge, both technically as well as conceptually, to further evolve this medium. The reason for this claim is that users have their own definition of reality as perceived by their senses and processed by their logic, which when exposed into a VE acts as a reference point of comparison. An example of this case can be seen in Plato’s Cave allegory Steinicke, F. (2016), where the story’s participants refused to acknowledge the existence of a different reality even when confronted with evidence of its existence. Of course, the objective of VR is not to replace reality itself but rather to host experiences that include the make belief factor. Referring on this objective, Burdea & Coiffet (2003) suggest that in order to provide the make belief factor, the focus should be balanced between technology and UX.

As an outcome of the above, the quality of a VR environment depends on the technology that supports it and the application’s scenario. This statement is further enforced by the studies of Seibert (2014) that brought evidence of how the quality of HMD affected the experience of users. Additionally, Sherman & Alan (2003) give emphasis on the tasks within a scenario and propose a more conservative guideline of VR scenarios such as walkthroughs. This last suggestion, even though suitable for many of the cases, limits the boundaries of how VR can be utilized. Of course, as the authors stated, it was the hardware limitations that prevented them from stating otherwise. However, a decade later, the increase in computational power and higher resolution HMDs, are now capable of supporting complex scenarios in areas such as physics, architecture, engineering, design or even games. This allows us to further explore aspects within the VEs that can enhance the UX. In order to do so, it is necessary to evaluate what the current state is, what is it missing from it and how can it be expanded.

The first issue explored in this thesis is how the interaction methods of the medium can be enhanced in order to offer a better experience. This, however, is not be made at the expense of any additional hardware since the immersion of the VR experience already depends on the quality of the visual means, typically in a form of a headset. As the VR medium is not well established within the consumer market, it was considered highly important to use a technological non invasive solution instead, in order to introduce hand interactions without any external dependencies. This has the additional benefit of being used with already existing HMD without requiring from the users to purchase any additional equipment. As an outcome, a software based, bare hand gesture interaction is introduced in order to
expand the modality and increase the type of -appropriate- communication tools that are going to be used in order to support and carry the information. As Nigay & Coutaz (1993) explain, modality “covers the way an idea is expressed or perceived, or the manner an action is performed”. This is a vital part of the VR experience where the simulation of a multimodal means creates a more compelling experience for the user. As Rautaray & Agrawal (2015) note “The use of hand gestures provides an attractive and natural alternative to these cumbersome interface devices for human computer interaction. Using hands as a device can help people communicate with computers in a more intuitive way”. The second issue is the make belief factor through a social perspective. More specifically, VR applications are most commonly offered in a setting where a single user is experiencing a VE. However, even in the cases of complex scenarios, this type of experience is more close to the one of a simulation as the possible outcomes unfold from a predefined sequence of events. In the cases where interaction with the environment is allowed, the presence of a user is limited to its intended purpose with the scenario itself. Brey (2008) suggests that the major difference between simulations and VR is the predefined set of parameters simulation has that yield no surprise or variety, whereas VR has the multi-user capability, referred as networked VR, which allows telepresence, real people in virtually generated scenarios. This has an outcome of increased diversity which can differentiate a VR experience from simulation. For this reason, social norms from the real world should be integrated into the VR one. Behrendt (2012) explains that multipersonal participation and collaboration are the aggregation of a set of conscious and willing actions into a complex behavior that has culturally involved into a social norm. These norms classify as existing behaviors that are worthy of transferring to the virtual world and further enrich the user presence and UX.

The setting of this study takes place within a VR Tower Defence (TD) game as experienced through GearVR. The reasons for this choice are the low entry requirements of GearVR and the benefits of conducting research in this type of scenario as stated by Avery, Togelius, Alistar & Van Leeuwen (2011). Additionally, the fun factor makes the study’s environment suitable for introducing new technologies to users. The aforementioned issues were explored with a user oriented mindset as both quantitative and qualitative methods were used in order to better understand the UX.

1.2 Motivation

Even though technology is ubiquitously around us, the emerge of a new medium will always spark a wave of enthusiasm among researchers and users as well. In the case of VR, as previously explained, the magnitude of enthusiasm this medium has, does not rely on its novelty anymore but rather on its increasing availability. As a practical example of this, VEs can now be experienced through a mobile phone which makes the medium’s hardware-based requirements affordable to greater audiences. This creates the perfect settings for the medium’s evolution as the demand for content is inevitably going to increase. At the same time though, the quality of this content should be expanded in order to allow the experiences within this medium to be appropriate and distinct. As an outcome, the focus should be placed in instilling the feeling of presence when within a VE in order to make the user feel as if he was indeed transferred into another world.

As an expansion to the content importance, Trendforce¹, highlights the venture growth of this venture and points towards the importance of software solutions - applications. Based on this, it is safe to assume that games will be a big part of VR as happened in other mediums as well. As an outcome, game experiences will have to be transferred in the new medium a process that has a great degree of freedom and challenges in order to do so. As Zyda (2005) suggests, the future of games should be developed through a design agent and this study can be a reference point of how different methods of interaction can be used within a VR game. The user centered mentality of this

research makes the findings significant towards enhancing the UX of VR games as different aspects are investigated in depth by users that are relevant to the concept of games.

Additionally, as seen in other cases of interactive media such as computer games, the addition of multiplayer functionality has over time affected the medium giving birth to the genre of MMORPG that is being described by Ducheneaut et al (2006) as a “phenomenon of growing cultural, social, and economic importance, routinely attracting millions of players”. With this expectation of a rapid increase in users that have access in VR technology and hardware, this type of behavior could likely take place within a virtual world. Even though it is relatively early to guarantee the previous prediction, this study brings forth insights of how a multiplayer session is perceived by the gamers and how does the addition of another player affects their gaming experience.

In order to do so, some fundamental issues of this experience should be improved such as the way we interact with the VE. For the most cases, if the interaction is indeed allowed, it is based on methods that rely on haptic feedback and mechanical means in order to do so. This raises the question of how natural and intuitive this interaction is when in the real world we mostly use our own hands to interact with our surroundings. It is also worthy to wonder how natural this interaction feels and what impact it has on the user experience. Additionally, what is the worth of a virtual world if we can not share it with others. Instead of limiting the virtual experiences in individual sessions, the virtual space could be shared instead and occupied by multiple users in order to form an experience together. This could potentially increase the depth of how the VE is perceived by users and make it a more compelling place to visit again.

In essence, the capabilities of 3D space in an immersive environment are a blank canvas where stories can be told and experiences can be formed. As an outcome of this, it is the need for better experiences that drives this study’s progress both technologically and conceptually.

1.3 Research questions

Asides the obvious benefit of conveniently simulating a scenario that would otherwise require a lot of resources, VR can offer much more than just being an alternative option for illustration purposes. Since the word reality is present in the medium’s description, VR should not be characterized only by its subparts such as the UI and interaction methods. Instead, the emphasis should be placed in the UX as it is the final outcome of the VR exposure that summarizes the user’s attitude. Having this in mind, it is now clear that the design is the one that shapes the UX and it is a constant challenge to use the appropriate means to form it. Based on this, two were the main problems that were identified and initiated this research. The way we interact within the VE and how is the environment populated.

In the context of strategy games, a user is typically involved in the scenario as an external element. This means that even in the cases where the camera placement is in first person, the user does not immediately interact with the game but through a controller. The studies of Cairns, Wang & Nordin (2014) discuss the immersion differences by using a different kind of controllers, however, the core of the problem is in the design of interaction itself. As an analogy, a controller that does not actively participate in the VR experience is similar to the strings of a puppeteer as he controls the puppets. This has an immediate effect to the UX as the presence of the user is divided among him and the object he controls. This is contradicting to the VR experience as the user should be experiencing things as if he was there. Following the previous analogy, the user should instead be treated as an actor where any visual illustrations of his presence should be the equivalent of a mask or a costume. The same reasoning applies to the controllers. It is unorthodox the fact that we use buttons and joysticks in a VR experience when none of them is a necessary part of us. Of course, they can be scenarios where they could make sense, but in most of the cases, a button is something a user interacts with and not the interaction method itself. Instead, it makes more sense to include the hands of the user in the VR experience and use a design that is centered towards this.
The other issue is the fact that VEs are mostly offered in sessions intended for individuals. In an age that technology allows us to become increasingly connected with each other, virtual spaces should be a logical extension of this. The freedom of VR, among others, can also provide the playground of the future. It is truly amazing and unpredictable how users would interact with each other and what could they achieve if only we allowed them to do so.

As a summary of the above, the interaction methods within VR and the presence of more than one user in the experience, are the main points of interest of this thesis within the scope of a game. In order to address those issues in a concrete manner, two research questions (RQs) were formed to better explore the aforementioned issues and bring forth data in order to improve the UX.

- **RQ1**: How does the use of hand representation and gesture-based interactions affect the gaming experience and feeling of presence in VR in comparison with the default method offered by GearVR?
- **RQ2**: How does the presence of an additional player affect the gaming experience within a cooperative VR environment.

1.4 Approach

There are existing methods capable of detecting the hands and gestures from a user, however, as a general rule they depend on either external artifacts or they require additional features from a mobile phone such as a depth sensor. Later in this thesis, these technologies are explained in further detail, however, these requirements lower the scalability of a project as they possess increased dependencies that might not be universally available. This issue becomes more profound if we consider the fact that in most immersive VR settings, additional means are required with the most typical of them being a HMD. Since the multiplayer functionality is also being put to the test here, it would be rather troublesome to increase the entry requirements for the users as this could prevent them from getting access the game. Instead, a seamless technology should be used in order for the entry stage to be as transparent as it can be.

In his work, Yousefi (2014) presents a technology that uses only the phone’s RGB camera in order to provide an interaction method by the use of hands. He also poses the same question, of what the future of virtually shared spaces can be.

The use of this technology will serve as the bare hand interaction method through the use of Manomotion SDK², while for the hardware part, Gear VR³ was selected as the VR headset of choice for its simplicity of use, and wide availability to most of the Android users. In Chapter 2, Gear VR is further explained and taxonomized among other types of VR systems, however, its increased availability makes the default haptic interaction method it offers, through a sidebutton, a worthy case of comparison in order to potentially optimize the VR experience.

Finally, as an appropriate game concept for this thesis studies, TD was chosen as the VR scenario the players would experience. As further explained in Chapter 4, TD games require from the user to manage a series of tower-like objects in order to fend off enemies as part of the game requirements. Game versions with different interaction methods in the management system were implemented in order to study the effects of them.

---

²https://developers.manomotion.com/
³http://www.samsung.com/global/galaxy/gear-vr/
1.5 Contribution

According to Trendforce\(^4\), there is an estimation that VR consumption is going to be increased from 9 million units in 2016 to 50 million units in 2020. This will additionally generate $70 billions in sales in 2020 driven mostly by software and not hardware. This forecast, although market and revenue based, signifies the rapid space VR is taking as a medium.

This thesis utilizes the aforementioned bare hand interaction method within a game setting in order to investigate what differences will occur from its use compared to the already established interaction methods. This venture focuses on UX and how it is shaped by the use of these interaction methods. By using an inductive reasoning, quantitative data were used as variables of reference in order to study the relationships between them and understand how the experience is shaped. Additionally, the users perspective was considered highly valuable in order to investigate the aspects of the experience in a greater depth. As an outcome of these, this study’s objective is to understand the effects of using a user’s hands and gestures as an interaction method in order to bring forth data towards its potential as well as its limitation. Moreover, the implications and significance of a shared virtual experiences were also explored through a cooperative scenario in order to better understand how additional users in a VE affect the UX.

1.6 Thesis structure

The thesis is structured in 7 chapters and their subsections. Chapter 2 provides the theoretical framework that is used as the basis of information as presented by the core concepts explored in this thesis as well as by referencing to literature and related technologies. Chapter 3 offers the methodology and an overall description regarding the approach used in this research as described by the data collection, study sample and statistical treatment. Chapter 4 is dedicated to the game prototype that was used in this thesis and the relevant information regarding the game mechanics, scenario and interaction methods used in it. Chapter 5 presents the results as captured and analysed while chapter 6 expands on the discussion upon them. Finally, chapter 7 concludes the thesis by summarizing the key aspects as well as listing the limitations of the study and providing ways to expand it.

Chapter 2

Theoretical framework

This chapter offers information regarding the various terms and technologies utilized in this thesis. Additionally, it includes the review of relevant literature as well as the investigation of already established applications that operate within the same scope.

2.1 Core concepts

2.1.1 VR

Virtual reality is the term used in order to describe the means, settings but most importantly the mindset of a participant within an environment. Extensive arguments can be made towards the virtuality as well as the reality part of this term, however, it would be in vain to try and generalize it since it relies heavily on the technological means that support it and the participant’s perception. A very good definition is offered by Earnshaw (2014) where he defines VR to be “The illusion of participation in a synthetic environment rather than external observation of such an environment. … VR is an immersive, multisensory experience.”. This definition is indeed fairly accurate for the time being, as it brings forth the modularity within VR as described by the participation element and the multisensory engagement. The temporal limitation was intentionally placed due to the quality of synthesis the environment is created. At the moment, the means of VR experience consist of peripherals and attachments in order to induce the sensual input from the user with information that is relevant to the VE. As an outcome of this, the definition of VR is time dependent in a sense that since the external means of VR are heavily depending on the technological state and computational strength it is certain to claim that the sensation within VR can easily be labeled as an illusion or as a simulation at best for the time being. However, this is something bound to change since the advancements in the aforementioned areas could make the distinction of what is an illusion and what is not, a very challenging task which in turn would create different scales among what we categorize to be VR.

This brings forth the true impact of VR which in turn makes it able to be categorized as a new medium. The novelty in the outcome of means and settings is the intensity in which a user experiences the environment. In other words, the difference of VR with other mediums that stimulate the senses is the magnitude of this stimulation and the perspective of events. In comparison with TV where a participant passively receives visual information, VR places the focus on the user instead as the information is presented through a VE that the user can interact with. Additionally, by isolating the user’s vision to only what he can see inside the VE, aside from the first person perspective, the VE can be experienced in higher intensity as it is less likely to be interrupted by external stimuli that are not relevant to the environment. In essence, VR presents information in an event manner where the user’s mindset is pushed towards the environment he is experiencing.

As prior mentioned, VR is implemented by systems that utilize appropriate means in an attempt
to override the senses of a participant in order to replace the information of the real world with the one of the VE. An example of those means can be seen in Figure 1, where the visual and haptic input and feedback are supported by different artifacts. A clear distinction among the visual means can be seen in the case of the C.A.V.E. system compared to the stereoscopic glasses and headset. The difference here is that the user is placed in a dedicated room where all the walls, as well as the floor, are projection screens whereas in the cases of wearables they can be utilized independently from the surrounding space. However, the use of one mean not necessarily exclude the use of another and as seen in Figure 2.1, it is possible to combine different artifacts in order to achieve a better transition to the VE.

Figure 2.1: VR means and settings.

Additionally, VR experience is categorized by L.E.K (2015) into 3 formats, SuperVR, MediumVR and CasualVR, Figure 2.2. This categorization is based on the collection of hardware and technology that supports the experience. The significant difference in the price range of the 3 categories is also an indication of a non saturated market where there is room for experimentation thus great variety in offered experiences. Hence, there is no established standard for VR use that applies de facto since as the same source suggests, “A format war is breaking out among these three approaches, as well as among the individual companies using each approach.”. This statement allegorically summarizes the vast growth in the field and the increasing interest from both the scientific community as well as the corporate world.

Figure 2.2: Formats of VR experience stated by L.E.K (Super VR, Medium VR, Casual Mobile VR)
2.1.2 Immersion and presence

As VR settings extend the way users experience an environment, the term of immersion is frequently used in order to describe the phenomenon of a participant transitioning from the physical “here” to the virtual “there”. Biocca & Levy (2013) refer to immersion as the process to which a virtual environment submerges the perceptual system of the user by using computer generated stimuli. This reference is in line with the statement of Mestre, Fuchs, Berthoz & Vercher (2006) as they state that immersion is achieved by removing as many real world sensations as possible, and substituting these with the sensations corresponding to the VE. However, Calleja (2014) brings forth the duality of this phenomenon by differentiating it to immersion as the user adsorption to the VE and the immersion as a transition state to it. As an outcome of these, the immersion in the context of VR can be summarized as the transition and constant occupation towards an artificial environment that is capable of sustaining the make belief factor. As Eichenberg (2011) suggests immersion is one of the key requirements in a VE in order to claim and maintain the variable of realism in the experience.

On the other hand, presence incorporates the feeling of a transition in a VE but in a sense that is highly related to the degree which immersion is achieved and ensured. In the context of VR, Slater & Wilbur (1997) describe presence as the “sense of being there”. This, of course, is a highly subjective metric as it relies on the perception of the participant and it is highly influenced by their prior experiences and emotional state within a VE. However, the same researchers argue that presence is indeed a measurable variable since a comparison of the stimuli effects between the VE and the real world can indicate the degree of presence a user is feeling. More specifically they argue that users being highly present should experience the VE as more the engaging reality than the surrounding physical world. Additionally, they claim that an effect of increased presence sensation can be seen as participants will tend to respond to events in the VE rather than in the real world.

These statements not only bring forth the significance of presence in a virtual experience but also indicate that a qualitative measurement can be an adequate assessment for it. However, it would be more beneficial to expand the notion of presence into subparts, capable of describing a broader sensation of presence within a virtual world. As a virtual world the definition of Maratou, Chatzidaki & Xenos (2016) is used, in order to state it as “a digital, persistent, 3D, graphical environment that can be occupied by multiple concurrent users via the network.” The addition of the human factor over a network raises the issues of the presence being perceived in different levels now. More specifically, Sheridan (1992) states that presence is the sense of being in a computer-generated world, and telepresence, the sense of being at a real remote location. Additionally, Campos-Castillo (2012) expands the definition further by defining copresence, the sense of “being together” with others. Finally, Nowak (2001) has presented extensive research towards social presence, a metric that in the context of VR refers to the degree of salience between the participants in the same VE. This allows the transfer of social behaviors and interactions as presented by Duvall (1979) into the virtual world.

2.1.3 Gesture based interaction

Gestures are a form of communication that is used both consciously and unconsciously in human to human interaction. They are the most expressive form of body language and they convey very direct information regarding the intent. Gestures have been a subject of interest among the scientific community because of the added value towards human computer interaction (HCI) it yields.

In their work Davis & Shah (1994) pointed it out the importance of gestures that are used in the everyday life, specifically as seen by the ASL, which signifies their great potential in HCI. Additionally, Pavlovic, Sharma & Huang (1997) identify the use of gestures in everyday life as mainly a tool of communication while their definition for gestures in HCI gestures are described by the pose of a hand and/or arm as well as the spatial position within an environment. As an outcome, the gesture is represented by the trajectory of hand poses within the suitable interval.
Additionally, the same researchers proposed a classification system of gestures according to their purpose. As seen in Figure 2.3, a gesture is always intentional and meaningful. Its purpose of interaction is either to manipulate the positional or rotational attribute of an object or to convey information. This information is usually a complementary mean towards verbal communication thus it is either used in a form of an action performed by the hand poses or by symbolically representing the information. One of the most significant parts of Symbols is the one of referentials as it can be used to communicate an idea or an action that is occurring in the real world. An example of this can be seen in the Microsoft’s latest Hololens\textsuperscript{5} system that incorporates these kinds of gestures as well as in other systems described later in this thesis.

![Figure 2.3: Gestural taxonomy that describes and differentiates the hand gestures.](image)

2.1.4 Games

Games can be characterized as the medium of entertainment. However, this statement is relative to the context of the discussion as the attempt to define what a game is, yields diverse outcomes which can also conflict with each other. Throughout time, researchers coming from different backgrounds and disciplines theorized definitions in an attempt to universally describe what games are.

Fullerton’s (2014) definition describes a game to be “a closed formal system, that engages players in structured conflict, and resolves in an unequal outcome.”. Very similar to this, Salen & Zimmerman (2004) state that game “is a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome.”. Even though it is unclear what a “system” can be described as and what happens in the case when the rules are changed or overridden, the important notion in this definition is the mention of engagement in an activity and the outcome which of course it does not required to be specifically quantifiable. The reason in this claim is in line with the definition of Schell (2014) that states that “A game is a problem-solving activity, approached with a playful attitude.”. This definition is abstract enough to cover a broad range of what can be regarded as a game since it brings forth the emotional state as presented in the playful attitude. This reveals the close connection of games with entertainment in a sense that they lead to the sensation of amusement and joy. This, of course, is a highly subjective matter that is influenced by the social, emotional and cognitive background of a gamer. This means that something that is considered to be a game by someone, it is not guaranteed to be also considered as a game by someone else.

It is also significant to mention the definition of Crawford (1984) where he initially described games to be represented as ”a closed formal system that subjectively represents a subset of reality”\textsuperscript{6} and later Crawford (2003) he described games through a classification model as seen in Figure 2.4. This

\textsuperscript{5}https://support.microsoft.com/en-us/help/12644/hololens-use-gestures/
enhancement of definition coming from the same researcher indicates that it is indeed challenging to propose a definition that is both universal and time enduring.

However, the most interesting addition in order to describe what a game is, comes from maybe the most unorthodox source, academically speaking. Being a game designer himself, Koster (2013) uses game derived examples in order to describe what games are and notes that “games are teaching us the skills we might need in real life in a safe, low-stakes environment.”. This indeed is so true if we think that in most cases or in the definitions prior used, games are not played by themselves but instead have the human factor thus the cognition, experiences and skills from the users that are involved. As an extension to Koster’s note, and as a derived from the rest of definitions, games could be in fact be categorized as a philosophical approach. The reason for this claim is that since philosophy is the study of general and fundamental problems concerning matters such as existence, knowledge, values and reason, games can be the setting where this study could happen in a low stake environment. Additionally, the notion of philosophy is explained by Deleuze & Guattari (1991) as “knowledge through pure concepts” the game setting can combine the intuition of a person within the appropriate experience. The difference with the traditional approaches as presented by Honderich (2005) is the playful attitude.

![Figure 2.4: Visual representation of Crawford’s game definition.](image)

Evidence shows that we increasingly consume and play games more and more every year and market analysts predict a further increase as seen in Figure 2.5. A fact that supports this claim, is the expansion on the game semantics and the evolution of games in both the conceptual and tangible level. As an example of this, as Kent (2010) notes, since 1961 games have been available in the early versions of personal computers thus creating the term “computer games” a term that became a standard of our time. The same principle was applied in the case of “video games”, where an intelligent system - or console - was dedicated solely to the gaming purpose via the execution of software that was created with an entertainment oriented mindset.

The same principle applies in the case of games that either implemented solely for the VR experience, or that the concept of pre existing games was instead transferred in a VE. In the case of this thesis, the term VR games are often used in order to describe the gaming process within a VE.
2.2 Related work

2.2.1 Gesture-based interaction technologies

In order to list the different methods that manage to offer solutions which enable the use of hands in the interaction with the UI, it is necessary to differentiate them according to the technology and means they utilize in order to achieve that.

The initial step in most of the cases is to detect the hand in order to be used as a reference point towards further detail such as hand poses, finger movement and spatial position. As mentioned earlier, the aggregation of hand and finger movements in the spatial space in combination with the different poses can be treated as a gesture according to the definition of Pavlovic, Sharma & Huang (1997). As Mitra & Acharya (2007) also explain, gesture recognition is the process of recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. In contradiction to the human senses where gestures are mostly perceived visually, gesture recognition can be applied in HCI by both computer vision or other sensory aided means where the above variables can be mapped and perceived as an output of a meaningful gesture. As an outcome the following methods can be placed into categories according to the way they manage to achieve the hand detection and gesture recognition.

Approaches such as the CyberGlove by Kevin, Ranganath & Ghosh (2004) utilize sensory data in order to capture the motion and subtract the gesture from a user using hardware embedded on a glove. Similarly, Fröhlich & Wachsmuth (1997) presented the systematic capture of hand gestures by biomechanical means as the positional, rotational as well as joint placement is captured by external means that are attached on top of the hand. The same approach was followed by Parvini & Shahabi (2005) with especial importance towards the joint movement. However, as the same researchers suggest the main issue of this approach is that it highly depends on the individual manner the gestures are performed by a user as captured by the external device. Moreover, as seen in Figure 2.6 the different accessories that are placed on the user’s hand, could have a different fitting which in turn would introduce noise in the tracking process and also require a form of calibration in order to aggregate gestures performed with a slight difference from unique users.
On a different side, computer vision is one of the most frequently used approaches in order to achieve the same goal. As a typical approach, a form of image capturing is utilized with the major difference among the methods being the type of optical sensors they use and the treatment of the visual input towards the hand and gesture detection. Additionally, the placement of the cameras used in this process also differentiates the methods as seen from a UX perspective. More specifically, in their approach Ge, Liang, Yuan & Thalmann (2016) use the heat maps that indicate the joint position in conjunction with depth maps as captured from a depth sensor as an input to a convolutional network. This approach brings forward the flexibility of computer vision techniques compared to the prior biomechanical approach. The reason for this claim is that as the neural network is trained, it would be capable of detecting different hands and gestures with increasing accuracy as the captured data can also become part of the dataset if needed. Similarly, Sharp et al (2015) in their supplementary material they explain how they utilize a data set of hand poses in order to compare them with the depth and RGB values from a camera equipped with a depth sensor. In their approach the color of the different parts of a hand, as described by the RGB pixel value, is utilized in order to identify a region of the image captured as part of a hand while the transition among their prototype poses could also be treated as a gesture.

Yousefi (2014) and Manresa, Varona, Mas & Perales (2005) proposed systems that utilize only the RGB input from the camera without the use of a depth sensor. The similarity in these two methods is the hand segmentation from the background based on color information and hand element references. However, even though this process can be used in mobile settings while produces similar results with the approaches that utilize more sensors, it is highly dependent on a good segmentation in order for the hand and gestures to be recognized. For this reason, they have an included functionality of calibration which tried to optimize the segmentation process and limit the noise in the detected hand. Li & Kitani (2013) also bring forth the challenges of hand detection in an egocentric camera as the process is occurring from a first person perspective. Challenges that affect this process, as well as extensive research in the field of computer vision, has been made by Moeslund & Granum (2001) as they review 130 publications which cover and compares the methods proposed for the past 2 decades setting the standard of effectiveness in the values of robustness, accuracy and speed.

Finally, a list of commercial artifacts such as the Leap Motion⁶, Microsoft Kinect⁷ and Hololens⁸ and others as seen in Figure 2.7, are utilizing depth sensors embedded in their camera input to offer a hand detection and gesture recognition solution.

---

⁶https://developer.leapmotion.com/orion/
2.2.2 Shared virtual spaces and multiplayer VR games

There is already a number of existing applications that have extended their presence to the VR settings as well as some innovative tools that have been designed specifically for the purpose of virtual world collaboration. MuVR, a multi-user virtual reality platform proposed by Thomas, Bashyal, Goldstein & Suma (2014) is an example of portable VR system that can support multiple users. All the rendering calculations are made based on spatial sensors that are hosted on a vest-like uniform and then illustrated on the Oculus rift HMD. Although, MuVR was a great inspiration towards this study, the high entry level regarding equipment, related to SuperVR, shifted the focus of this research to more affordable means that would be more appealing the end user. Similar to Kaufmann & Schmalstieg (2003), this study is aimed at collaboration in real time without limiting the users at physical colocation. The difference here is the representation layer where the task is taking place in a virtual scenery instead of an augmented one. Finally, a number of entertaining based applications such as Minecraft VR9, VR LAN party10, Zero Latency11 and The world inside the Web12 are showcases of collaboration being introduced in a virtual environment.

Even though the potentials of those shared virtual spaces were early mentioned by Lea, Honda & Matsuda (1997) the studies of Grinberg, Careaga, Mehl & O’Connor (2014) brought forth the impact on the participant’s behavior as studied through a Second Life scenario. Finally, in a purely VR setting, Greenwald, Corning & Maes () proposed a VR framework capable of supporting multiple concurrent users within the CocoVerse platform.

---

10https://www.engadget.com/2016/08/19/bigscreen-oculus-avatars-sound-vr/
11https://zerolatencyvr.com/press
12http://www.janusvr.com/
Chapter 3

Methodology

This chapter presents the discussion on the research methodology of the study, the subjects, sampling technique, research tools, the procedure of data gathering, and statistical treatment that will be used for data analysis and interpretation.

3.1 Research design

According to Rubin & Babbie (2012), the key objectives of a research is to explore, explain and evaluate a situation or event. In the context of this thesis, UX is the main point of interest and how the different variables such as control mechanisms and the presence of an additional player affect it. As an outcome of this, the research approach was designed in a way capable of understanding these effects in both quantitative and qualitative form. The reason for this choice is the human factor within an experience and the necessary depth needed in order to adequately comprehend it.

For the purpose of this thesis, two individual studies were conducted in order to study the effects of those variables. The first study was aimed to address the use of the player’s hands and gestures as an interaction tool in comparison with the GearVR touchpad. The study took place in two single player sessions, each one with a different controller. The second study involved pairs of users playing the game together in one multiplayer session in order to study the effects of an additional player presence within the virtual environment.

As a required step towards achieving the above and answering the research questions (see section 1.3), a game was created (see chapter 4) in order to serve as the environment where users would experience a scenario within a virtual world. The scenario consisted of a set of actions the user should take in order to complete the given task. Additionally, the design of the environment was purposefully increased in the challenge in order to introduce more variables into the problem solving procedure and push the user into actively participating to the event. As a necessary step, in order to ensure emotional consistency regarding the game outcome, the game was intentionally coded to be unbeatable. The reason for this choice is the observation of Wilson & Kerr (1999) that states that winning produces a range of pleasant emotional outcomes and reduces arousal while losing according to Quick & Cannon (1990) has a physiological to an emotional chain reaction that leads to an arousal increase and a mixed emotional state.

Based on the aforementioned theories and issues, we can now ensure that the game outcome will always be negative thus set the emotional state of the users. An additional argument that supports this choice, is that by following this approach it is guaranteed that 100% of the users would lose the game instead of depending on their skill capabilities. This leads to a state normalization where users have the same outcome thus most likely having the similar emotional state and increased arousal. Since the game scenario (see section 4.1) is both mentally and physically demanding, a failing outcome would also encourage the users into reflecting what went wrong and thoroughly express their thoughts and
emotions in order to be stored as data.

As a summary of the above the research approach followed in both studies, was to put users into an “uncomfortable” situation where they adopt the hero mentality in order to perform the given tasks and fulfill a goal. The task performance may vary among different users however, the outcome will always be negative thus ensuring the emotional and arousal state in their responses.

3.2 Studies sample

Due to the limited timeframe available for user studies, convenience sampling was utilized for the purpose of this thesis. Bachelor and Master’s students of Linnaeus University, from various departments, were reached through Social Media and asked to participate in both studies. The studies took place in the Linnaeus University - Media Technology Department and lasted for 10 days.

Since the VR experience was offered within a game context of increased difficulty and demands, it was considered appropriate to filter the population and select participants that had prior experience with games. This attribute was considered capable of increasing the performance of the given tasks and utilized their experience as a criterion that would increase the validity of the study. From this process, 23 users were recruited. All 23 of them participated in the first single player study, however, due to external obligations and traveling arrangements only 12 of them participated in the multiplayer one.

3.3 Data collection process and methods

3.3.1 Study 1 - Game controllers

Study number 1 was performed in order to investigate the issues regarding RQ1. All 23 users participated in two gaming sessions. Each of the sessions had the same VR game with the only difference being the control mechanism. In order to ensure the reliability of the results affected by the first impression of the game and the controllers, the user pool was divided into two, roughly equal, parts (11 and 12) and they were asked to try the VR game in an opposite order. 11 users had the hand gestures as their first experience and touchpad as the second, while 12 users had the touchpad as the first experience and hand gestures as the second. In order to ensure minimal interference regarding the game performance, the study followed the proposed light conditions proposed by Manomotion for optimal performance. Thus, for the sessions where the players were using their hands and gestures, the user was placed in a room with no direct lights pointing towards the phone’s camera and the surrounding environment was of the same color that fulfills the proposed standards\(^{13}\). Finally, all the sessions were monitored in real time by an ingame camera that was also networked in the scenery and the session was stored in a video log.

\(^{13}\)https://developers.manomotion.com/getting-started/
Figure 3.1: Users on synthesized background, game methods (Gesture, Gesture, Gesture, Touchpad).

Figure 3.1 visually represents the interaction methods and in the following fields, it is presented in further detail the steps, methods and procedures that were used in order to retrieve data from the users.

3.3.1.1 Learnability phase

In this phase, users were introduced into the study’s settings. Initially, they were informed about the game’s scenario that would, later on, be used as a point of reference. After the game settings, tasks and goals were clearly understood by the users, they received visual information about the game by experiencing it through a version meant to be used through a mobile phone without the use of HMD. This version implemented the same scenario they would experience in VR but since the phone’s screen was visible by both the participant and the instructor, it was easier to discuss and visually pinpoint any issues they might occur. Additionally, this mobile non-VR version used the phone’s gyroscope in order to move the camera around and exhibit the experience the users would have while turning their head in VR settings. This allowed the user to get familiar with the game environment, UI and interactive elements.

The next step was to inform the players about the two types of controllers, followed by a demonstration from the instructor. The GearVR headset was thoroughly presented to them and significant attention was placed to the interactive surface located on the right of the device as the use of this touchpad would be the tangible interaction method they would have to use. On the same manner, the hand detection process from the phone’s camera was presented to users alongside with examples performed by the instructor of good hand placement and distance from the device in order to minimize the occurrence of errors. Followed to that, the gesture detection process was explained and presented to the users once again by the use of examples performed by the instructor as he demonstrated how the hand gesture should be performed in order to be recognized by the device. Finally, the users were given the mobile phone and they were asked to test both types of controllers to ensure that their use was understood. This process was repeated until the participant felt comfortable using both types of controllers in the given scenario. Notes from this process was kept as qualitative data regarding the initial user reaction to the controllers.

3.3.1.2 Interaction methods interview

After their gaming session being completed, users were asked to describe their thoughts in an unstructured interview. In order to capture a broader field of information, this interview was split into two parts. The first part took place immediately after the user had exited the playroom. In order capture

14The use of touchpad button tap was simulated by tapping on the mobile phone screen explaining the relation to the physical button while the use of hand gestures was demonstrated as is.
clearly the feedback and allow the users to freely express themselves without any hesitation, the interview was presented to them as a reaction video response where they had to express their thoughts about the interaction method they used, in 30 seconds or a bit more. For this part of the process, the user responses were captured in the video as the interviewee had a monologue without the instructor intervening in any part. The second part of the interview took place as the users were asked to rest while waiting for the next gaming session to commence. In the second phase of the interview, the instructor engaged in a freeform discussion with the user regarding the interaction method they had just used. Depending on the user responses, the instructor followed up with more questions on aspects he considered to be of greater importance and value. During the second phase of the interview, notes were kept of the most relevant, significant and surprising responses.

This approach was followed for both gaming sessions a user had in order to get results for both interaction methods. An average duration for each interview was 10 minutes, however, on the second iteration, more time was invested as users were prompted to compare the UX in the different settings.

3.3.1.3 NASA Task Load Index

NASA - TLX is a proposed tool from NASA that is used in order to provide a workload assessment. According to Hart & Staveland (1988) workload is a measurable entity and its definition should be independent from personal interpretations as it is susceptible to personal biases. Instead, a quantified measurement of 6 factors allows the assessment of a workload based on the mental, physical and temporal demands as well as a person’s performance self assessment, the effort needed and frustration levels.

In the context of this thesis, NASA - TLX is an appropriate tool in order to allow users to assess the workload of playing the game in a quantified manner. For this reason, they were asked to fill a paper version of the NASA - TLX (see Appendix A2) in both gaming sessions right after they had completed the interview.

3.3.1.4 Interaction methods questionnaire

After considering Van Baren & IJsselsteijn (2004) guide on measuring the presence, a shorter version, of appropriate questions from the Witmer, & Singer (1998) questionnaire was chosen in order to get user feedback regarding the UX while playing the game with a certain controller.

The main purpose of this questionnaire is to collect quantified data that can potentially lead into generalizations that can be transferred from this study’s case to a broader one. As seen in the chosen questions (see Appendix A1) the main theme is the sense of presence the user had during the gaming session and how did the control mechanisms affect it.

3.3.1.5 Choice of controller

Since the user pool for both studies consisted of the same participants, as a final metric towards answering RQ1, users were asked to choose the game version they would like to use for the multiplayer session. By this choice, they were referring to which control method they would like to use in the upcoming gaming session.

This metric can not be used as the reference point towards answering the research question on its own, as it depends on many factors and this choice might be altered depending on the exposure time the user has with the different control methods. However, it yields significant information regarding the impression those control methods gave to the user.
3.3.2 Study 2 - Presence of an additional user

Study number 2 was performed in order to investigate the issues regarding RQ2. The game scenario was the one from study 1, however, this time two concurrent users (as seen in Figure 3.2) would be within the virtual world cooperating in order to achieve the goal. The participants were 12 users that had previously participated in study 1 and had already declare which version of the game they would like to use in the multiplayer session. The pairs of users were chosen according to relative similarities within the gaming behavior that were captured in the video log from study 1 and the choice of controls in order to prepare the study’s setting accordingly.

As an initial step, users were prompted of the game scenario and control methods in order to ensure that they were aware of the necessary game information. Following to that, users that had chosen the gestures method were placed in the low light room, previously used in study 1. Users that had chosen the touchpad method were placed outside of the room within a 2 meter distance to ensure that they would be able to communicate without physically colliding with each other. In the cases where both users had the touchpad as an interaction method, they were placed within 2 meters distance with each other for the same reasons as previously mentioned. Intentionally, they were no pairs of users using gestures as an interaction method as the study settings were not capable of hosting them. The activities of pairs\textsuperscript{15} using the touchpad method within study 2 were captured in video log and also both version activities monitored by the networked ingame camera.

![Figure 3.2: Users participating in multiplayer VR game.](image)

3.3.2.1 Interview

After the completion of the multiplayer session, users were asked to participate in an unstructured interview with the instructor. In order to avoid users influencing each other on their responses, one to one interview sessions were held. The main theme of the questions is how the user experienced the game within the multiplayer session and how the presence of an additional player affected it. The whole processes lasted roughly for 30 minutes and the interview was captured in a video log.

3.4 Statistical treatment - Data analysis methods

As seen in the previous sections, both qualitative and quantitative data were collected in order to answer the RQ following an inductive reasoning. As a necessary step towards interpreting the data the following procedures were followed:

- **NASA - TLX**: Numerical mean values and standard deviation were calculated for all 6 factors within MS Excel. This process was followed for both interaction methods in order to observe the estimated workload in both cases.

- **Interaction methods questionnaire**: Histogram sorted values as well as numerical mean values and standard deviation were calculated for all question within MS Excel. Additionally,\textsuperscript{15}The low light conditions limited the ability to keep video logs as the capture was lacking light.
median values were also calculated in order to investigate the existence of outliers within the response pool. This process was followed for both interaction methods in order to examine how different factors of the interaction methods affected the UX.

- **Interaction methods interview:** The notes of the initial user reaction to the interaction methods as well as the second part of the interview was serialized and integrated within a text document. The same reasoning was followed for the video logs as the user responses were also turned into text manually and included in the same document. With the use of Python scripting, the body of text was normalized where all characters was turned into lower case. After that, through the use of NL Ranks, English stopword dictionary\textsuperscript{16} the text was filtered from unnecessary words. Finally, for the purpose of content analysis, the body of responses was sorted for the most frequent words appearing in it. Similar words in semantics were grouped together, such as difficult and hard, as well as words deriving from the same root such as repair and repairing. This process was done manually as the lemmatization process and the fuzzy matching were not yielding adequate results. The same approach was used for both interaction methods in order to get a quantifiable metric alongside the interpretation of the user responses.

- **Choice of controller:** As a mixed approach both the qualitative and quantitative information in the user choice was stored as a reference point towards the interaction method of preference.

- **Multiplayer interview:** Relation analysis was used in order to treat the user responses from the video logs. In the first step, the content of the videos was parsed in order to form a general impression. Followed to that, rough labels from relevant sections, phrases and words were kept as an interpretation guideline. These initial labels involved actions, activities, opinions and concepts. As part of this open coding process, multiple cycles were performed in order to discover patterns of sub themes that could inductively form major ones. The main criteria in the coding process was a repeating frequency of an information, a clear statement from the users that this was considered important by them and finally a response that was considered surprising to the researcher within the context of this study.

\textsuperscript{16}http://www.ranks.nl/stopwords
Chapter 4

Prototype

This chapter offers the information regarding the game prototype that was created in order to study the factors that affect the UX and answer the RQ. In the following fields the game scenario, mechanics and, interaction mechanisms as well as the implementation process, are presented in further detail.

4.1 Game scenario

The game used in the data collection process implements a first person Tower Defence scenario. TD games are a subgenre of strategy games where a player is given the objective to defend a certain location of importance from an opponent, which in most of the cases is operated by computer AI. As the name implies the main tool that is used in order to achieve this goal are towers which they attack the enemies as they come close to them. Usually, the player is the one that decides where the towers are going to be placed and as seen in many of these cases, through an ingame currency the player is able to upgrade them as well. The most common setting for a TD game is a predefined path that connects the enemy starting point and the location the player has to protect. The enemies are walking on this path and unless stopped by the towers they will reach the endpoint thus making the player fail on his objective. Depending on the difficulty level, a player is given a certain amount of lives which represent the margin of failure he is allowed to have before the game finishes. The final evaluation of the player’s performance occurs either when his lives have reached to zero, or when there is no active enemy and the remaining lives indicate how well he performed.

From a visual perspective, TD games traditionally were played in a top down or birds eye view, where the player is able to see the entire terrain and occurring events. More recently, third person TD games were introduced where the player is controlling an avatar object thus having the camera following it.

In comparison with the above, the game scenario of the prototype follows the same principles of a classic TD game where the player needs to stop the enemies from reaching the end point. As seen in Figure 4.1, 3 towers were placed alongside the path that connects the start point to the end point. When an enemy would be in range the tower would shoot arrows at it. The difference witnessed here is that the towers have a predefined position without the player being able to change that. Instead, the scenario asks for the player to supply the towers with arrows as they will run out of ammunition. Finally, the camera was placed on top the player’s avatar in a first person setting giving the feeling that instead of controlling a secondary character, the player is the hero himself.
4.2 Game mechanics

As explained in the previous section, the objective of the game is to stop the enemies from reaching the base before his lives run out, which in this case were 20. As a necessary task, the player needs to supply the towers with arrows so they can keep shooting down the enemies. Each tower constantly detects if there are enemies within a 30 meter radius from it and if they are, it shoots an arrow to the closest one every 1.5 seconds. By doing so, the shot takes away 1 ammunition point out of the starting 5 it has. The ammunition point is translated to 20% of the bar that is being placed on top of the tower to keep track of the ammunition levels, the same bar updates the values with different colors gradually as they change. The tower will alert the player with smoke signals that it’s running out of arrows when the bar reaches to 20% and if the ammunition is depleted the tower will become inactive and stop shooting.

In order for the player to supply the towers he has to visit a tower spot and interact with arrowstack that is placed next to it. This interaction requires the player to look at the arrowstack in order to activate it. An invisible raycast extends from the player’s camera center point and upon colliding with an arrowstack it will turn its outline blue signifying that it is active. In order to make the aiming process easier, a crosshair, symbolized as a circle, is placed in the center of the player’s camera. While the arrowstack is active the player can interact with it depending on the game version he is in. For the touchpad version, the helping orange text on top of the arrowstack will ask the player to tap in order to supply while the gesture version will ask the player to use click instead. Upon a successful interaction, the player receives an audio confirmation and he increases the ammunition counter of the tower by 0.3 points.

Every arrow shot by the tower is guaranteed that it will hit its target and it will inflict 40 points of damage. In both game versions all 50 enemies, represented by goblin avatars, had a health pool of 100 points and if their life was set lower than 0 they would die.

As there are 3 towers in the game, the player needs to supply them all and possibly at the same time. In order to do so, he needs to travel among the tower spots using the teleport map. This
mechanic allows the user to move into different locations on the map, specifically to the places where the towers are. The approach is similar to the to the supply mechanism, however, this time the user has to aim in the interactive map next to him and specifically to the tower icon he wants to be teleported. All three towers have unique colors in order for this attribute to be used as a reference point. As the player aims to the yellow tower icon on the map, the outline of it will become blue signifying that the user can interact with it and change his location.

As an outcome of the above, the general meaning in the actions a user can take is supply and teleport. An example of this process can be seen in Figure 4.2, as the player can be seen interacting with the arrowstack of the red tower in order to activate it again.

Figure 4.2: View from the networked camera of the player avatar supplying a depleted tower.

Even though the game was offered in a first person experience the player presence was enhanced by an avatar that differentiated him from the enemies and displayed his position with a large blue diamond hovering on top of him. This attribute was especially useful in the multiplayer session as players could look around and receive visual information about their teammate’s position.

Based on Hsu,Wen & Wu (2007) theory which states that “Challenge is a major contributing factor to fun” a series of design choices were made in order to push the user’s physical and cognitive limits and provoke them into adopting a hero’s mentality. First off, the fast pace of the towers arrow depletion forces the player into constantly moving in order to keep them active. While doing so, the player has to teleport to different towers using the teleport mechanism, however in order to prevent him from knowing the exact location that he will land on, two secret destinations for each tower were linked to the teleport map. These locations, even though closely placed, differ in the way they are rotate thus preventing the player from forming a strategy that has the tower location for granted. Finally, by studying Alger’s (2015) recommendations for comfortably placed elements in a scene, the interactive elements of the game were placed in angles that were impossible for the user to interact with just by turning his head. Instead, in order for the player to switch from the arrowstack to the teleport map, he had to twist his entire body and access the curiosity zone. This constant struggle alongside the background music of a marching tune were intentionally created in order to inspire the player to actively take action.

4.3 Interaction methods

As mentioned already, the VR game was offered in two versions. The difference in them was the interaction method that allowed the player to use the game mechanics that were described above.
The first version utilized the touchpad which is the default method offered by the Gear VR. As seen in Figure 4.3, the touchpad is a rectangular shaped area on the right side of the HMD and has a cylindrical extension that functions as a button. The contrast of the surface texture, where the touchpad area is smooth while the button is extruded from it, offers a haptic feedback to the user in order to locate where the button is. The functionality is similar to the one of a mouse button as the user is required to press on the surface in order to trigger an action.

As a typical setting, users have their hand hovering over that area and perform the tap action by moving their index finger to the button position. In the cases where the finger coordinates do not match the location of the button, through the touch sense the user can identify the distance and try again. Alternatively, they can place their palm forward to the touchpad and use this fixed position as a reference point that minimizes the error of margin. This approach is also more comfortable as the weight of the hand is not suspended in free form but rather placed in the side strap of the Gear VR. Finally, the integration of the button functionalities was provided by the Oculus SDK\textsuperscript{17} where the button tap is recognized as an event that can be used in order to initiate different types of interaction.

![Figure 4.3: Samsung’s GearVR touchpad and tap method.](image)

The second version uses the phone’s camera in order to allow the user to interact with the environment through hand gestures. More specifically, instead of using a tangible device, the user has to place his hand in front of him and follow a sequence of hand motions in order to perform an action. By using Manomotion’s SDK\textsuperscript{18} computer vision technologies are able to translate the raw input from the phone’s camera into meaningful information such as hand detection and gesture recognition.

The user is receiving constant feedback of his control state and actions as his hand representation is projected within the scene. To achieve that, the stream coming from the phone’s camera is constantly checked in order to see if a hand is detected. A binary image of dimensions 320 x 240, highlights with a white color the area that is believed that a hand is present and the rest of the pixels are marked as black. The next step is to use this binary image within an array in order to have a reference point towards describing the hand position, size, rotation and segmentation. Followed to that, an RGB image of the same dimensions is stored from the camera’s stream in order to have information regarding the hand’s texture. Finally, as seen in Figure 4.4, the array of the binary image is used in order to determine the opacity of the RGB image in order to offer a hand representation. For a given frame, the array position of the binary image that has a value of 0, meaning black, are used as an index for the RGB image array in order to set the opacity of those pixels to 0. This process creates a mask for the RGB image that uses the binary array as a reference towards the opacity of its pixel array, leading into an illustration that shows only the hand.

As long as a hand is being recognized by the camera’s input, the user will receive visual feedback of his hand being present in the VR scene in real time. Additionally, if the hand detection is not accurate\textsuperscript{17}https://developer.oculus.com/downloads/\textsuperscript{18}https://developers.manomotion.com/
the noise will be mapped in the binary image leading into a distortion of the hand representation. In that case, the user is prompted to calibrate the hand detection process by tapping on the touchpad. The calibration stage as prior mentioned is a necessary functionality in order to fix a situation where the hand segmentation can not or is poorly achieved. Calibration is a manual process that adjusts the parameters of the detection process in order to remove any noise and successfully extracting the hand from the rest of the pixels. This functionality was assigned to a touchpad tap when the user was using gestures within the VE and on a screen tap when the user was testing the scenario in the mobile non VR version.

![Figure 4.4: Hand representation transferred in the virtual world.](image)

The gesture recognition follows the same logic of the computer vision techniques that was used for the hand representation. The camera’s input is scanned for a hand presence and the hand pose as formed by the finger position is examined for patterns that can form a gesture. As seen in Figure 4.5 a transition of hand pose from an open pinch to a closed one and back to an open is defined as the dynamic gesture of Click. This dynamic gesture is the equivalent of the user tapping on the GearVR button, as seen in the previous interaction method and it was used by players to interact with both the arrowstack and teleport map after they were active.

![Figure 4.5: Sequence of hand poses to perform Click gesture as defined by Manomotion.](image)

4.4 Implementation

The game was implemented using the waterfall model (Royce, 1987) as it was considered the most viable for the purpose of this thesis. In the initial steps the game specifications and design were laid out, followed by the individual models that were implemented and tested in order to be integrated into the game. The game was developed within the Unity game engine using C#. Unity is highly evaluated by Farouk, Simon & Ksentini, (2015) since it is the environment choice for many game
developers because it allows the creation of high fidelity games in a very fast pace. Moreover, the use of Manomotion’s and Oculus SDKs aided in the task of implementing the control methods that were used in the game.

The game logic and functionalities (Figure 4.6), even though implemented in an external environment, were eventually packaged into a cross platform compatible application, capable of operating in Android VR settings. The application was installed on two phones and by following the Service Oriented Architecture, users were connecting to the game server, in this case, Photon\(^{19}\), in order to receive the content. The logic of the network operation involves only the objects of the scene that have modifiable variables while the static objects such as the game terrain, sky and miscellaneous ones were loaded from the application itself. As an example, throughout the duration of the game, the position of a player or an enemy might change whereas the position of the sky will remain the same. For this reason, the dynamic elements of the scene such as the towers, players, enemies and lives have to update their status not only locally but through the network as well. For this reason, these objects were treated as networked entities and given a unique network identifier in order to distinguish themselves and be able to be reached from other entities as well. This way a synchronized state through the different network entities was formed which translates into a game state that is commonly shared by all participants.

However, since this synchronization requires information to be sent through the network, a series of choices were made in order to prevent any latency occurring from a communication overflow. First of all, the communication of the network entities and the server used the UDP protocol which is minimal message-oriented and requires no prior system dialogues, such as the triple handshake. This ensures that messages will arrive fast to their destination and in the cases where there are package duplicates the Photon Server uses the package’s timestamp in order to sort them out and filter them. In the cases where a message was for any reason not delivered, an entity would keep the previous value it had until a new package would update it. In the game context this means a delay of events happening such the position of an enemy being wrong, however since enemies update their position as they move, their positional information will be sent again through the network very frequently. Finally, in order to minimize the amount information networked variables such as a tower’s ammunition would require a certain event in order to advertise their values. Examples of this, are the shooting function of the tower and the supply function of the player that alter the ammunition variable of the tower thus needed to be sent through the network.

Both versions of the game were implemented for GearVR. The reason for this choice is that it falls into the MediumVR format chosen for its ease of use and deployment as well as the low entry of knowledge required for users. The content is being rendered on the phone’s screen that is being hosted inside the GearVR while the phone’s camera is used for the hand detection and gesture recognition. The ambient sound of the scenery, as well as the audio feedback, are being played by the phone’s speakers.

\(^{19}\)https://www.photonengine.com/en-US/onpremise
4.4.1 Expert Validation

As a final step of quality assurance, the VR version of the game was tested in order to detect any operational flaws and ensure that the study sessions with the game would be functioning as intended. The process involved 2 engineers from Manomotion that specialize in VR interaction methods. They individually used the application and were asked to perform the tasks of teleporting and supplying the towers with arrows as the game would progress. In the meantime, the instructor also participated from a computer hosted version of the game in order to simulate network traffic that would occur in the game. As a secondary step, the experts were told to explore the game in free form in order to provide general feedback. The outcome yielded no errors although both experts agreed that the VR experience was highly challenging and demanding. Additionally, interference with the hand detection was noted that led into a distorted hand presence and low accuracy in gesture detection. The reason in this issue was appointed to the physical environment conditions such as light and background colors. As an outcome of this, two one on one sessions with both engineers were held in order to take notes of how the study’s environment should be set and how the process of hand detection should be presented to the users.
Chapter 5

Results and analysis

This section presents the results that were gathered during the two studies of the thesis. For cohesion purposes, the most meaningful results of statistical analysis are presented through graphs and tables as they are a very effective way of accessing and interpreting information. The structure of the section follows the method order that was previously described in the statistical treatment (see section 3.4). 23 users participated in the study 16 of them being male and 7 female. 14 of the users claimed that were fairly aware and accustomed to the touchpad interaction method prior to the study, while only 3 did the same for the gesture method. Finally, for the study of interaction methods, even though the background colors and light conditions were ensured, errors in the hand detection were observed which introduces a degree of error in the results.

5.1 User study 1 - Interaction Methods NASA - TLX

Originating from the design phase, the game experience was calibrated in order to be challenging. As results are indicating, users found the VR experience to require a lot of effort. Additionally, they consider it to be of high and equal demand both mentally as well as physically. Moreover, the frustration levels and the high variance in the mean value indicates that the design has achieved its intended purpose. As seen in Figure 5.1, even though the values among the two interaction methods do not diverge significantly, there is a clear pattern of increased values in the case of hand gestures. Because of this, we can safely observe that users regarding this interaction method to be of a higher challenge.

![Figure 5.1: TLX histogram. Experience load as perceived by users](image-url)
5.2 User study 1 - Interaction Methods Questionnaire

Regarding the responsiveness of the interaction methods, users reported that by tapping on the touchpad they could interact with the VE and get immediate feedback as this method performed better. For most of the cases, the gestures did not pose much trouble as seen in Table 5.1, as the mean value and standard deviation are in acceptable levels. However, the outliers in the Figure 5.2 showcase the severe problems 6 users had with this method’s responsiveness.

In comparison, there is a clear pattern in the use of the touchpad with most of the user’s responses having a positive hue, while in contrast the responses for the gesture method are more symmetric and have a higher concentration towards the average value. As an outcome, it is safe to claim that by using the touchpad, the interaction with the environment was fully responsive whereas the gesture method shows evidence of less responsiveness as described by the user answers.

![Histogram of general system responsiveness as perceived by users in both interaction methods (Q1).](image)

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>3.52</td>
<td>5.86</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.7</td>
<td>1.32</td>
</tr>
<tr>
<td>Median Value</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5.1: Statistical values of general system responsiveness as perceived by users in both interaction methods (Q1).
While wearing the HMD and immersed in the VE, users declared that they were not fully aware of what was happening in the outside world. As seen in the values of Figure 5.3, both methods have a left skew which indicates that the general feedback agrees with this statement. More specifically, the identical median value and the similar mean values as seen in Table 5.2, indicate that the experience of the VE was little interfered by real world stimuli.

However, the increased distribution, as seen by the touchpad’s standard deviation, indicates a greater interference compared to the gesture method. This is more obvious in the case of the 2 extremely high values in the touchpad method pointing in an case where the users were not immersed and fully aware of what was happening in the real world. Additionally, it is also important to further examine the histogram, where the appearance of negative immersion regarding the gesture method, appears in a lower value while for the touchpad method it is significantly higher. This signifies that, even though immersion was generally ensured, there is indeed a difference among the two methods.

Figure 5.3: Histogram of user awareness of the external environment, while in the VR experience (Q2).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>3.17</td>
<td>3.3</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.3</td>
<td>1.98</td>
</tr>
<tr>
<td>Median Value</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5.2: Statistical values of the user awareness of the external environment, while in the VR experience (Q2).
Regarding the nature of the interaction methods, users reported a significant difference among them. As seen in Figure 5.4, users noted that overall the use of gestures offered a natural way to interact with the VE. On the contrary, by using the touchpad, users reported that the interaction seemed foreign to them. More specifically, there is a clear accumulation of responses towards the right side of the histogram when it came to the gesture method while for the touchpad the distribution is more symmetric.

An interesting outcome is the low distribution for the gesture method as most of the values are placed around the mean, while the standard deviation is relatively low. However, this is not the case for the touchpad that even though the median and mean value as seen from Table 5.3 are in accepted levels, outliers are still present in the histogram. This is an interesting result indeed as the methods do not differ significantly in their median value, however, the mean value difference among them presents a different information regarding the nature of the methods.

Figure 5.4: Histogram of interaction methods experience as perceived by users (Q3).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>5.21</td>
<td>3.78</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.95</td>
<td>1.44</td>
</tr>
<tr>
<td>Median Value</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 5.3: Statistical values of interaction methods experience as perceived by users (Q3).
In both cases of interaction methods, users reported that were mostly aware of the hardware components and peripherals that supported the experience. As seen in Figure 5.5, most of the values for both interaction methods are located on the positive side of the axis. There is little difference, that can be more clearly seen in Table 5.4, where the difference in median value indicates that the touchpad method was fairly more noticeable than the gesture method.

This metric by itself hard to evaluate as the awareness of the components has a different effect based on the method. In the case of the gestures, being aware of the display and the illustration of a user’s hand, is an indication that could potentially be linked with increased presence. In the case of touchpad though, since the feedback from the method is tactile it could potentially be counter-immersive. As an outcome this questions result indicate that users were indeed aware of the display and control devices, however, this notion should be combined with other resources in order to better understand the effect of it.

![Histogram of user awareness of the display and control mechanisms (Q4).](image)

Figure 5.5: Histogram of user awareness of the display and control mechanisms (Q4).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>4.73</td>
<td>5.17</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.42</td>
<td>1.58</td>
</tr>
<tr>
<td>Median Value</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5.4: Statistical values of the User awareness of the display and control mechanisms (Q4).
An interesting metric was the flow of information as perceived by users while playing the game in VR settings. As seen in Figure 5.6, users regarded the touchpad to have more consistency in the way it stimulated the senses. As a matter of fact, the illustration shows the values of this method to be placed mostly on the left side of the axis where the sole peak appears. For the case of the gestures on the other hand, the values of the responses are more evenly spaced in the axis which indicates a lack of rigidly describing the consistency.

However, as seen in Table 5.5 both methods had acceptable values regarding the consistency of information with regards towards the quality of the experience. The increased mean value of the gesture method and the difference between the median values of those gestures is an indication of difference. There are many factors that might have contributed to this response however, based on the prior knowledge of the methods performance, noise in the segmentation process might be the justification for these answers.

![Histogram of information consistency as perceived by the user senses (Q5).](image)

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>3.52</td>
<td>2.91</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.62</td>
<td>1.64</td>
</tr>
<tr>
<td>Median Value</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5.5: Statistical values of information consistency as perceived by the user senses (Q5).
As a continuous investigation towards the responsiveness of the both methods, users were prompted to respond regarding their ability to anticipate the outcome of their actions. As seen in the Figure 5.7, the vast majority of users were indeed capable of coping with the interaction method they used and in turn anticipate the outcome of their actions.

However, since this is a closed type question, the answers are yielding the general hue of the method’s capabilities and of course they are not in any position to indicate fine aspects of the experience it self because of the limited range in available responses. This is more clear in the cases where the users were in fact unable to anticipate what would happen next. Even though for both methods the values are really low with respect to the studies population, the increased number of users unable to anticipate the outcome of their actions when using gestures, is a metric that indicates that the individual factors that lead to this outcome should be identified. indicates

![Chart showing system responsiveness derived interaction method responsiveness as perceived by user actions (Q6).](image)

Figure 5.7: System responsiveness derived interaction method responsiveness as perceived by user actions (Q6).
Another interesting metric was the ability level which users interacted with the VE. As seen in Figure 5.8, there is a clear indication of the major difference witnessed here. Both methods have a distinct categorization which is clearly represented by the median values as seen also in Table 5.6.

For the case of the touchpad, users responded that the interaction was excellent with most of the responses having high to extremely positive values. Even for the cases of users that they encountered some difficulties, their responses was still on the average side. This was not however, the case for the gestures as both the mean and median value are positioned in the average spectrum. Additionally, again 6 were the cases that reported they could not use this method efficiently to interact with the VE. The number of respondents matches patterns of similar questions which indicates that there might be a correlation worthy of exploring. Overall though, both methods were accepted by users having though a significant difference on the quality of interactions they experienced.

![Histogram of interaction based, experience quality within the virtual environment (Q7)](image)

<table>
<thead>
<tr>
<th></th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>3.21</td>
<td>5.69</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.67</td>
<td>1.1</td>
</tr>
<tr>
<td>Median Value</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5.6: Statistical values of interaction based, experience quality within the virtual environment (Q7)
Regarding the engagement, users mostly reported that they were highly involved while in the VE. As seen in Figure 5.9, both interaction methods have a right skew and the similar mean values towards the positive side of the histogram support this claim. However, even though the median values of both methods are the same, as seen in Table 5.7, a closer inspection to the histogram reveals a greater tendency while using the gesture method.

Another interesting fact is that even though both methods provided high engagement in the VE, for the case of the touchpad the appearance of an outlier in the extremely negative values indicates a notion worth of further exploring. For the most part, the responses follow the same principle however the notable issues here are the peak noticed in the gesture method and the fact that the increased standard deviation of the touchpad method is highly affected by the outlier cases.

![How involved were you in the virtual environment experience?](image)

Figure 5.9: Histogram of participation evaluation as perceived by users within the experience (Q8).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>5.9</td>
<td>5.47</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.71</td>
<td>1.34</td>
</tr>
<tr>
<td>Median Value</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5.7: Statistical values of participation evaluation as perceived by users within the experience (Q8)
One of the biggest differences between the two methods was pointed to the distraction level introduced in the experience. It is challenging to decode how this was interpreted by the users as the histogram of Figure 5.10 gives a different perspective when compared to the mean values of the methods as seen in Table 5.8. In the average values we can observe a non significant difference among the methods, however the histogram pictures a different message as the touchpad’s distribution is clearly left side while the gesture’s is more evenly spread.

For the case of the touchpad, half of the responses are placed near the median value while the rest of them are declining reach towards the other end. The fact that there is no other peak signifies that the major theme here is indeed placed in the median value without a strong opposition. Of course, the other half, disagrees with this opinion and commented that they indeed felt distracted by this method but in a lower magnitude compared to the other great cluster.

Gesture in the other hand shows no clear peaks in the user responses and this signifies that there is indeed no clear description of how users felt during its use in the session. By using the median value as a reference point it is significant to observe the negative part of the scale thus claim that more users felt that they were distracted than the ones they didn’t.

![Histogram of evaluation regarding the distracting nature of the control mechanism (Q9).](image)

Figure 5.10: Histogram of evaluation regarding the distracting nature of the control mechanism (Q9).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>3.86</td>
<td>3.17</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.76</td>
<td>1.64</td>
</tr>
<tr>
<td>Median Value</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5.8: Statistical values of evaluation regarding the distracting nature of the control mechanism (Q9).
While playing the game with the two interaction methods, users reported a significant difference in the amount of delay they encountered when they performed any actions. For the case of the touchpad, as seen in Figure 5.11, the values are skewed to the left side, giving a positive indication regarding the experience. Moreover, the two peaks are located in the extremely positive values which could characterize the overall experience. This is not certain though as there are 3 cases which can be labeled as outliers that fall way out of the general opinion. This can be also seen in the Table 5.9, as the method’s standard deviation is increased by the appearance of those values. Even though the general outcome is highly acceptable, illustrated by the especially low median value, the appearance of those outliers are worthy to be noted.

There is a completely different picture for the case of the gesture though. There was no consistency in the experience, a fact that is illustrated by the symmetry in Figure 5.8. From the extremely positive to the extremely negative side of the scales, there is a huge gap in the experience’s delay when it came to this method. This can be further observed in the Table 5.7 with the high standard deviation. Additionally, both the mean and median values indicate a problematic experience as users commented that generally experienced a significant amount of delay.

![Histogram of user evaluation regarding the delay as perceived in the VR experience (Q10).](image)

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>3.91</td>
<td>2.47</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2</td>
<td>1.8</td>
</tr>
<tr>
<td>Median Value</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 5.9: Statistical values of user evaluation regarding the delay as perceived in the VR experience (Q10).
As previously refereed to the virtual "there" and the process of transferring the user to a VE, users have generally responded that the interaction method had little difference in this process. As seen in Figure 5.12, both methods are right skewed which indicates that the transition time was low for both of them. However, it is interesting to observe the outlier value for the gesture method which in turn indicates the inability of a user to be immersed in the experience. This in turn leads to the assumption that this specific user had a harder game experiencing the game with this method as the transition time was significantly higher than the average as seen in Table 5.10.

On a general scope though, as seen in Table 5.10, the statistical values for both methods indicate little problems towards the immersion process. This can be seen by both the high mean and median values that are on values towards the extremely positive side.

Figure 5.12: Histogram of transition ease in the VR experience affected by the interaction method (Q11).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>4.95</td>
<td>5.89</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.63</td>
<td>1.17</td>
</tr>
<tr>
<td>Median Value</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5.10: Statistical values of transition ease in the VR experience affected by the interaction method (Q11).
Similar to the previous metric, users rated the interface each method introduced, in a very distinct manner. As seen in Figure 5.13, there are peak values for both methods, placed at the exact opposite side of the scale. This is a clear indication of how the experience was differently shaped while using a different method.

For the case of the touchpad, users experienced little to none interference while playing the game. Most of their responses are placed near the median value, as seen in Table 5.11. Almost half of them described the perfect scenario where the method did not interfere with their performance. However, it is also noteworthy that the other half of the user pool experienced the exact opposite of it with their responses being at the negative side of effect. This issue can be also seen in the very high standard deviation and the high difference between the mean and median value. This is a clear indication that there were, in fact, two major opinions regarding the touchpad method as half of the sample pool reported either very low interference or very high.

Regarding the gesture method, the histogram indicates a high concentration towards the negative side of the scale as both peaks are located in high or very high interference. This is further enforced by the median value which showcases the difficulties users encountered with this method. However, as with the prior method, gestures had a minority that reported the exact opposite effect which can also be seen by the relatively high standard deviation and the difference between the median and mean value.

As an outcome, even though the methods returned significantly different results, the division between opinions in them individually, raises additional questions that need further exploring.

![Histogram of interaction methods interference with the game tasks (Q12).](image)

**Table 5.11: Statistical values of interaction methods interference with the game tasks (Q12).**

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>4.91</td>
<td>3.65</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.85</td>
<td>2.24</td>
</tr>
<tr>
<td>Median Value</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>
While playing the game with the two different methods, users reported having a different experience concerning the ability to be concentrated while on a task. This difference is indeed significant as the histogram of the users responses seen in Figure 5.14, shows a high concentration for the touchpad method and a great sparse for the gesture. More specifically, the appearance of a peak in the touch method signifies a clear group of users, in this case, half of the population, that declared they had little problems concentrating while using the touchpad. As the rest of the histogram is right skewed, it is safe to claim that this also the general opinion of users, with some variance of course. In fact, the close values of mean and median as seen in Table 5.12, as well as the low standard deviation is clear suggestion that users were not distracted while using this method.

On the contrary, the responses regarding the use of gesture, have a greater standard deviation, a great sparsity as seen in the histogram as well. Based on the median value, we can observe two groups of responses of almost equal supporters, claiming that exact opposite. This translates into two subsets of users, one of them had little trouble concentrating while using this method while the other one did. Finally, the appearance of 2 extremely negative cases is also noteworthy as it indicates that 2 users were not able at all to concentrate on their tasks while using the gesture method.

![Histogram of interaction methods interfering in the decision making process (Q13).](image)

**Figure 5.14:** Histogram of interaction methods interfering in the decision making process (Q13).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>4.04</td>
<td>5.73</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.74</td>
<td>0.86</td>
</tr>
<tr>
<td>Median Value</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

**Table 5.12:** Statistical values of interaction methods interfering in the decision making process (Q13).
From a cognitive perspective, both methods appeared to be wieldy to the users. As results indicate the majority of them feel capable of using both methods after their exposure to them. The difference in values presented in Figure 5.15, can be arguable due to the pre existing knowledge however the general schema of the histogram proves the above claim and yields some additional information about how it is formed.

The responses for the touchpad method are very concrete and rigidly placed in the right side of the scale, mostly around the median. As seen in Table 5.13, both mean and median values are very close to each other and the standard deviation is at low levels. This enforces the statement that users feel immensely proficient with this method and even the lowest values of their responses are in very accepted levels.

Similarly, users overall commented the same about the gesture method. However, there is a difference in magnitude when compared to the touchpad. Even though right skewed, gesture responses were noticeably more sparse with the peak located near the median value. This comparison can be further seen in the comparison of the mean value of both methods as seen in Table 5.13. Additionally, while using the gestures, 3 users reported that have no confidence in using this method for any future sessions. This is a clear outlier as it does not follow the general tendency formed by the responses of the majority of the study participants.

Figure 5.15: Histogram of post game, interaction method proficiency (Q14).

<table>
<thead>
<tr>
<th>Value</th>
<th>(Method) Gesture</th>
<th>(Method) Button</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>4.17</td>
<td>5.78</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.69</td>
<td>0.85</td>
</tr>
<tr>
<td>Median Value</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 5.13: Statistical values of post game, interaction method proficiency (Q14).
5.3 User study 1 - Interaction Methods interview.

Upon its introduction, the touchpad’s functionality was parallelized from users with the functionality of PC mouse. This made it very clear of how it operated and how they could interact with it. No further explanations were asked from the user’s side regarding the method nor was it considered hard to understand.

After using it, users reported that they mostly placed their fingers on the side strap of GearVR and used the index finger to tap on the active surface. As explained by them, they adopted this approach because they considered it very problematic to successfully tap with their hand hovering over it. More specifically, users claimed that it was almost impossible to know if they were tapping on the right spot unless they were holding their hand on the side, utilizing it as a reference point. By doing so they achieved a better performance, however as they noted this hand position acted as an anchor point that grounded them to the real world as they were constantly getting feedback from their touch sense.

![Figure 5.16: Wordcloud from the most frequent words users used in the touchpad method.](image)

In addition to that, as seen in Figure 5.16, teleporting was considered a difficult task to achieve since rotating their bodies with their hand in that position was considered challenging. However, due to the fast tapping pace, they said it was significantly easier to repair the towers.

Regarding the game experience with this method, users commented that the game was hard yet much enjoyable. They had little problems using the touchpad which made the experience very comfortable, however, they reported that it was hard for them to get the feeling they were actually there since the task they were performing was tedious especially when the enemies were walking next to them. For the gesture method, it was interesting to observe the enthusiasm players had when the
technology was presented to them. Most of the users spent a lot of time exploring it as they were very curious about how the device was able to illustrate their hand in a virtual scenery. In addition, they asked many questions regarding the limitations and functionality of this method. Following to this, the references to the hands and fingers were very frequent as it contributed to the whole experience from many aspects.

One of them was the high physical demands it required during the game session. As users reported, they became physically tired by holding their hand in front of the screen without resting. Additionally, by performing the click gesture repeatedly, they felt a great workload focusing on their fingers. Moreover, the fast pace of the game did not allowed them to find a comfortable position for their hands nor their fingers. More specifically, as they were worried about the accuracy of the method, they were not willing to experiment with an already working hand position since they were concerned that it would take time to adjust thus failing in their tasks.

Figure 5.17: Wordcloud from the most frequent words users used in the gesture method.

However, users pointed out that these concerns made them eager to try harder as the scenery music and game scenario put them in a mentality that they had to work hard in order to beat the game. As seen in Figure 5.17, fun was very frequently mentioned and it was usually appointed into a roleplay setting where they felt they were trying to defeat the enemies. Additionally, the majority of them declared that they felt that they were actually participating in an event rather than playing a game, however, they would like to use more gestures or even interact immediately by just using their hand.

Finally, even though they could move more comfortably in the VE without worrying about their balance in the real world, the noise coming from the hand detection, forced them to calibrate and in
some cases more than once, a fact that led into wasting time as they noted.

5.4 User study 1 - Interaction methods, choice of controller.

Following the end of study number 1, 2 out of the 12 participants chose gestures as their interaction method of preference for study number 2. The rest 10 participants chose the touchpad, however, there was an interesting factor that most of the participants included in their responses. This factor was the element of competition that even though was not intended in the study, it was formed by most of the participants. As they declared, they wanted to surpass the other “teams”, this is why they chose the touchpad in order to be faster. This was an unexpected turn of events as there was no “team” composition in the study’s design, nor a definite scoreboard in order to have a competition. The only variable they had constant access to, was the remaining lives, but as they had no experience of the game finishing with lives having a non zero value, it remains unclear how this setting was formed. In addition to that, there were 2 more cases of users participating in study number 2, that after consulting with their game partner, they changed their minds and switched to the touchpad method before they played the multiplayer session. It is unclear what was it specifically that made them take this choice, however, the competition spirit was obvious to observe. The player’s arousal was increased prior to the game session, compared with the single player one, and after completing it, they engaged into strategy discussions involving ways that they could have performed better. Additionally, all of the participants asked how their performance was compared to the other “teams”. This made very clear that indeed the participants were, in fact, competing with each other.

On the following discussion for the multiplayer session interview, when players were told that the game was impossible to win. 3 more users, which makes them a total of relative 7, admitted that they wanted to use gestures as their interaction method but they were hesitant towards their “team’s” performance.

5.5 User study 1 - Interaction Methods, additional observations.

The study’s setting took place within a VR game which, as observed by the researcher, made the participants very enthusiastic about having a session with this technology. Especially interesting was the reaction of players when they were introduced into the hand detection technology through a mobile version of the game. Almost all participants reacted to it as it was the game itself and they proceeded into further exploring it by testing different background settings, distance from the camera and gestures.

Even though the study’s setting was explained to the users, where they would have to use the game without any prior practicing, all of the users upon completing the game, asked to have multiple extra sessions after they had completed the two meant for the study. On average, if we exclude the two sessions of game exposure dedicated to the study, the users voluntarily used the game for an additional 30 minutes, or 7 sessions. Most of that time was spent exploring the hand gesture technology in the mobile settings followed by playing in VR using their hands. Additionally, the users asked for more content in the sense of other games, and they followed up with suggestions that could expand the one they had just played. The least amount of time spent, or 2 sessions on average, was on the touchpad version of the game and it was always the last choice before they stopped playing in general.

Finally, as seen in the video captures with the users avatars, users that were using the gesture method tended to move more compared to the users using the touchpad. This observation is referring to the body’s rotation as any other movement in the space was not able to be represented in the game scene.
5.6 User study 2 - Multiplayer session, interview.

Through the corpus of the user responses, common elements and attitudes were noted, capable of describing the general experience from the multiplayer session. As users suggested, this experience was something different from the single player one as the fact that they had a teammate, transformed the game into something else. Even though there was not a clear definition or a constantly appearing explanation of what the session was transformed into, 3 major themes emerged from their responses.

As seen in Figure 5.18, the addition of another player brought forth one more variable that shaped the way users sought information. It was considered highly valuable to communicate their thoughts, concerns and plans with the other person and in return, they expected the same from him. As they described, it was important to keep the information constantly flowing in order to be aware of what is happening but mostly what is going to happen also. In order to do so, they chose to verbally communicate information that was urgent to be delivered while they used their vision to collect data that they considered to be important. Positioning and logistics were the key variables here as the players were constantly seeking information about them. Knowing where their teammate was, what the ammunition levels were and where the enemies are currently at, were the most crucial information that would have allowed them to win the game.

Based on the above information, the next major theme that was detected was the problem solving procedure. The players described that they did no longer interpret the game as a series of tasks but instead as a situation, they were involved in. The majority of them treated the game as a scenario and that they would have to use a strategy in order to fulfill their purpose and overcome their failure from the singleplayer experience. This strategy highly involved the other player and the avatar that was used to illustrate him was a constant reminder that they not only had to work together but also to help each other. More specifically, they expressed their willingness to help their teammate by teleporting to the same spot they were because they felt bad looking the enemies pass through their towers and as teammate was unable or struggling to stop them.
While the majority treated the game as a cooperative scenario, there was a small part of the sample that entered the game with a puzzle solving mentality. 4 of the participants declared that they believed that the game would be won only by using certain patterns of collaborative actions. More specifically, they orchestrated a set of teleporting sequences that were focused only in keeping the towers up all the time and ignored any other type of information in the scene. They also declared that they were not interested in experiencing the game while solving this type of puzzle and their teammate was considered as a tool in order to achieve the goal.

Regarding the experience itself, users regarded the multiplayer session more rewarding compared to the single player one. The summary of their description can be labeled as the fun factor that was formed from the individual pieces of the experience. Users commented that by sharing both the virtual and real space made them feel more relaxed and willing to be entertained. In an addition to that, the unpredictable behavior from the different personalities involved in the experience introduced a new variable that was not controlled by the users. More specifically, most of them declared that it was a
funny challenge to cooperate with their teammate as they did not have a clear communication protocol with each other from the beginning. This made possible for events to happen that were not predefined or agreed upon. The addition of fun to the prior challenge mention was linked by users to a real world situation of them gradually getting to know a new person. As a side note, it was interesting to hear from users that knew each other prior to the study that they discovered a new behavior from their teammate within the VR game. As users pointed out, this was highly linked with the fact that they were taking up a role that fixed their mindset towards achieving a goal. This process was regarded highly entertaining and the visual elements in the scene greatly added to it. Almost all participants referred to their teamplayer by using his avatar description pointing out the funny appearance and animations. Once again, the element of competition was clearly spotted from the users responses as winning the game while other teams wouldn’t be regarded as a very rewarding motive while they were taking part in the experience.

Finally, there was a unanimous statement that the multiplayer setting would be their choice for any future sessions as they had a lot of fun playing in it. Their argumentation was based in the different kind of entertainment they had received and the different experience they were exposed to. The single player setting was referred to as the practice environment they would often use in order to become better at the game, but as users said, it is more meaningful and fun to play this kind of game with another person, either being in the same location or not.

5.7 User study 2 - Multiplayer session, additional observations.

Similar to the single player session, users requested to have additional time allowed in order to play the game in the cooperative setting. The enthusiasm levels were significantly higher compared to the single player session. As an addition to that almost all of them shared in social media content taken from the study’s setting. However, when they were informed that the game was designed to be unbeatable they lost almost all their enthusiasm and willingness to play again. Instead, they requested to use again the hand detection versions of the game in singleplayer settings both in VR and mobile.
Chapter 6

Discussion

This chapter discusses the results from the two user studies reported in chapter 5. Additionally, the issues that emerged from the results as well as their contribution to the research questions, are also presented here.

6.1 Interaction methods session

The difference of the two methods as far as UX is concerned was observable from the early stages of the study. Upon their introduction, users formed a completely different first impression. More specifically, the touchpad was regarded as a technology that was easy to relate to as the functionality was connected to existing physical devices that have button interfaces. This formed a sensation of stability as users immediately felt comfortable using this method. Their interest was shifted instead to the VR setting and they were more interested in exploring the game mechanics and visual elements. On the contrary, when introduced to the gesture method, users were very enthusiastic about the technology involved in this method. Compared to the touchpad, they spent significantly more time exploring it and they engaged in conversation with the instructor in order to discover more about it. The new to them hand representation and gesture functionality captured their attention and initiated an exploring behavior. The challenges of a new interaction method eased the users into building up an emotional state that was fitting to the game scenario. This was further obvious by their main focus which was to identify their role in the experience.

The above facts bring forth evidence of how differently users entered the VR experience. In the case of the touchpad, users had a more passive mentality and their concerns were located into how to receive information from the game rather on how to interact with it. On the other hand, gestures introduced a more egocentric view as users regarded themselves being the center of the game instead. They were frequently referring to their presence in the game, how they should interact with it and the impact their actions would have to the final outcome. It is a reasonable argument of how the lack of prior experience with this method initiated this kind of behavior, however even in cases of users that had prior knowledge of it, there was no difference in their initial reactions. In the case of the gestures, users vividly imagined themselves being the protagonist as if in a movie while for the touchpad method the low levels of arousal indicated that users treated the setting as just another game.

The same logic was followed after users had their first experience with the game. Their responses were more vivid and they were clearly more excited when using the gesture method. Even though, as intended, none of the participants was able to win the game in any of the game versions, it was interesting to observe the different aspect the interaction methods introduced. Upon losing the game while using the touchpad, users were more concerned on how the game mechanics developed the experience while in the case of gestures the focal point was themselves. Additionally, even though it was reported that the game was highly demanding in both versions, users referred to it less when
using the gesture method. Instead, since their mindset was closer fixed to a challenging scenario, they pointed out specific areas where they felt uncomfortable. As seen in the two word clouds, users unanimously agreed that the game was hard, however, the fatigue experienced while using the gesture method was more specific as users pointed out that they also experienced the high workload physically in their hands and fingers. The controversy here though is that this weariness did not prohibit them from having fun but instead they claimed to have more of it since their description to the experience was more cheerful and they referred into the fun statement more frequently. Regarding the difficulties introduced by each method, a series of issues emerged from feedback coming in many different cases. When using the touchpad version of the game, users reported that by holding on to the side strap of the device they found themselves to be very accurate and relaxed by also very distant from the experience. This is further enforced by their responses to questions Q1 and Q7 in the interaction questionnaire where they reported having little trouble interacting with this method and the interaction to be very robust. However, they also noted that by holding their hand firmly on the side of the device they essentially anchored themselves to the real world by constantly receiving haptic feedback from something that was not part of the VE. Moreover, their body stance made it harder for them to twist their bodies and as seen from the ingame camera, indeed players rotated less than when using the gesture method. This is a reasonable outcome, since the limitation in balance this grip introduced, significantly brought forth the thought of falling over which inevitably made the users less mobile. Effects of this case can also be seen in Q8 where users felt less involved in the VE. On the same matter, the picture of the gesture experience is completely different when it comes to difficulties it introduced. The main issue can be pinpointed to the interference users experienced while using this method. If we follow the logic unveiled by the responses to Q1, Q7, Q9, Q10, and Q13 it is possible to understand that users faced difficulties when using this method such as low responsiveness and delay in their action’s outcomes. This made them feel distracted and not able to interact with the environment as they wished so. This interference was appointed by users to the cases where they experienced noise in their hand and gesture detection which led them into calibrating the device. This action was indeed thought to be highly disliked as it interrupted the game flow in many of the cases and took away play time. This effect is significantly important since as users reported in Q3 and Q8 they had a great sensation of natural interaction with the VE in an experience they felt they were highly involved. Even though this interference did take away their entertainment, the inconsistency of game flow is highly disruptive in otherwise highly enjoyable experience.

As an outcome of the above, users formed a different opinion about the two methods towards the effect they have on an experience. Since they were asked to choose which one of them they would use in the multiplayer session, users with a competitive mindset mostly chose the touchpad as it was very robust and accurate, which would have allowed them to compete with other teams not caring so much about the experience but rather on the result. On the other hand, when it came purely to entertainment, users frequently asked to use the gesture method outside the scope of the study where they claimed there was no competition involved. They instead spend a lot of time exploring the game and interacting with the VE simply because they just wanted to have fun.

6.2 Multiplayer session

As stated in chapter 5, there were a plethora of comments regarding how the multiplayer experience differed from the single player one. Among the major themes that were formed, the social perspective is very imminent as the references to the team member is present in all of them. This was something that was expected since an additional player is technically one more variable in the game, however, it was surprising to observe it in such a magnitude since it greatly affected how users experienced the game.

The first interesting part was how users approached the gaming session even before it began. The
fact that they gave serious thought and even changed their controller choice in order to be a better fit in what they theorized to be a team, indicates that in fact, they were willing to share this experience. Of course, there were different motives involved in this process, however, this indicates the basis of a social behavior. This was further enforced by the high expectations that led into theorycrafting and strategy planning before the game began. Even though there were some similar strategies, the difference in personalities and how the participants communicated their ideas was a clear resemblance of a social interaction towards a matter of the real world. This indicates that the gaming session was treated equally in importance to the one of the real world.

Having previously experienced the game, users were now familiar with the scenario and fully aware of all the mechanics involved. The only variable they could not control was the behavior of their teammate. This was a key factor that shaped the experience, as users tried to prepare for it before the game and constantly tried to predict while the game was in progress. On the individual level, this was very reasonable for them to do so, as they wanted to get the maximum value from the additional player. However, as users reported, this was easier said than done since the different personalities, embodied by the game avatar, had to be socially interacted with. More specifically, users pointed out that it was highly important for them to know their team player’s actions and intentions while maintaining knowledge of what is currently happening in the game. This process was very challenging since the game was hard by itself and the addition of cooperation increased both the mental workload but the rewards as well.

As an extension to this, users reported that the ambiguous moments within the game were many. This happened because either they had performed individually very well or because they were cooperating efficiently with the other player. The same principle was applied in the case where they thought they made a mistake or because they believed they were not synchronized with their teammate. This by itself is a quantifiable metric that says that the additional player introduced more challenges, however, the most interesting part was how it affected the emotional state of the users as well. The key reference here was towards their team player’s avatar since as they noted it was the only friendly visual element in the scene. This reveals an additional notion that is closer related to the co-presence in the scene and the shared goal among the players. It was no longer only the towers that helped the player achieve the goal but the other player as well. The game was also less linear regarding the emotional state as the additional player could introduce moments of frustration with an action the user would disagree with or he would introduce joy as he would cover up a mistake that otherwise would be impossible to dismiss when playing alone. It was finally surprising to hear that users felt bad seeing their team player’s avatar struggling to supply the tower with arrows, which made them teleport to their location and help them out.

As an outcome of this, not only the avatar evoked emotions to the participants but a sense of empathy is being hinted out as users would try to understand the emotional state of the other player. The same effect was observed right after the experience was finished as the participants rushed to discuss with their teammate their thoughts and often apologizing for not playing well. This real-life behavior made the users claim that even though the game scenario was the same, the multiplayer setting was far more compelling to them.

6.3 Research questions

Based on the results of the two user studies presented in chapter 5 and the discussion offered in this chapter, it is now possible to answer the initially stated research questions.

- **RQ1:** How does the use of hand representation and gesture-based interactions affect the gaming experience and feeling of presence in VR in comparison with the default method offered by GearVR?
As an initial outcome, users considered this method to be highly amusing since the innovative concept of gesture interactions as well as the visual representation of their hands, immediately captured their attention. While using it within the gaming session, users also reported that it felt more natural since it allowed them to reach into the VE. By doing so, players had the feeling that they were actually there as the visual information they were receiving included other parts of their physical body which gave them the sensation that they were in fact somewhere else. Compared to the touchpad method, the position of their hands did not interfere with their balance, which in turn allowed them to turn their bodies with more confidence while exploring the virtual world. In fact, the hand illustration as captured by their hand in the real world was used as a reference point which allowed users to turn their head with a better accuracy. However, the non resting position of the arm, and the repeated actions performed in the gesture process increased the physical workload of the experience. This gave the sensation of fatigue to users as they were becoming more tired as time progressed. At the same time though, due to the game scenario, users felt that they were more involved in the experience because of this exhaustion and their affective state was influenced significantly.

Finally, in the cases where the method failed to detect a gesture or poorly illustrated the hand, users were greatly annoyed as the game flow was immediately broken. The calibration method had the same effects on the experience as users regarded it as an interference that did not allow them to play. However, as they reported this sensation for most of the cases was momentary as they would immediately focus again on playing the game by using their hands. This effect brings forth once again, the impact on presence this method has, as it was able to recuperate from a great affective transition.

- **RQ2:** How does the presence of an additional player affect the gaming experience within a cooperative VR environment?

As an initial outcome, participants felt that they were no longer alone in the virtual world which in turn had an impact on their behavior and the UX as well. More specifically, the addition of another player increased the challenges within the game as the actions a player made were highly affected by the position and actions of his teammate. Players now shared both the virtual space and the experience as well since their in game avatars were used as a reference point of action from their team player.

The most impactful, however, was the occurrence of social behaviors within the VE. The intention to help the other player and the expectation to receive help back was noted not only to the common goal but also on individual parts of the game where users consider the state of the other player. As an extension to this, the planning process towards the in game cooperation can be regarded as a practice that indicates the existence of social interactions.

As an outcome of this, the biggest difference the additional player introduces is the expansion of presence in both participants experience. The telepresence artifact as illustrated by the game avatar was used a variable towards the problem solving procedure. Moreover, the visual element, alongside the performed actions of a player enforced the sensation of co-presence to the other player. Finally, the game plan and efforts to achieve it established the social presence for both participants which had as an effect for users to regard the multiplayer setting to be more vivid, compelling and rewarding.
Chapter 7

Conclusion

As a final step, this chapter concludes the thesis by summarizing the general notions and outcomes that were brought forth in the previous chapters. Additionally, the limitations as well as the future work section, provide useful information of how this work could be continued and expanded.

7.1 Summary

As the technological means of VR are rapidly progressing, it is important to constantly evaluate the experiences that this medium is offering in order to improve them and at the same time expand the possibilities in the VEs. VR is in a mid development state, a fact that can be witnessed by the great amount of prototyping involved in these technologies and the high interest among the scientific community as well as the corporate world. As the entry requirements towards such an experience are lowered, an increasing amount of consumers has now access to VR which inevitably pushes the need for quality in virtual experiences.

For this reason, this thesis assessed the conceptual and technological means that support this medium in order to specify areas that can be improved and enhance the UX within VR. This venture was initiated in order to bring forth scientific evidence towards UX in VR, however, the reference point of this thesis was centered towards user feedback as it was considered highly important to get them involved in the medium’s progress.

The setting of this thesis took place within a VR game that used a TD scenario. As mentioned in the theoretical background, the playful attitude the games provide was considered a fitting case where users could comfortably experience this medium.

The topics that were investigated, were regarding the effects on the UX when using software assisted, bare hand gestures in comparison with the default haptic method as provided by GearVR’s touchpad. Additionally, the networked architecture of the game was able to host multiple users which allowed the observation of the experience as it was affected by the sharing of the virtual space and the cooperation scenario involved in the experience.

As results indicate, the use of gestures is still a challenging approach since the segmentation process of the hand is a crucial factor towards the usability of this method. However, the experience, as perceived by users, was more compelling since they stated that this form of interaction feels more natural. Additionally, the direct reach into the VE while users were using their hands, lowered the postural requirements in the GearVR setting, which in turn allowed users to have more stability and confidence in the navigation within the VE. These benefits though were overshadowed when the method failed to respond accurately to their gestures. The hand illustration of the user was a highly significant factor that made the user feel that was actually transferred in the VE. Additionally, the gestures empowered the presence of the users as they felt more engaged in the experience. However, this sensation of presence was dramatically decreased when noise in the segmentation process was
introduced. This indicates that the potentials in this method are truly great since the negative outcomes heavily rely on the technical implementation while the conceptual part of this method is ideal for use within VR games. As an outcome, further research on this method should be investigated as it has the potential to be a leading approach for the future of VR gaming. Regarding the addition of another player within the gaming session, users were highly amused and had a positive reaction towards an experience of this kind. Through their responses, users noted that the virtual scenery seemed more real when they were cooperating with a teammate and definitely more intense from an emotional aspect. The results from this study agree with the ones presented by the literature and further indicate that the human factor is an important part of the virtual experience. As an outcome of this study, the future of VR games should be in line with the norms that are followed in games of the other media which already incorporate multiplayer functionality. However, this should be done while bearing in mind that the affective state of users in VE is significantly high, which in turn indicates a very high sense of presence in the game. Thus, the design of both the game mechanics and especially the way social interactions are designed within the environment should be in balance with the norms seen in the real world.

7.2 Limitations

Even though the studies performed in this thesis were meant to be used in a gaming context, some of the design choices limit the scope of where these results can have an impact. First of all, the fact that the game was unbeatable, restrains the spectrum of the UX. It is highly expected that users would have a different experience if the outcome was positive for them. This variable, however, was not regarded of great significance towards the interaction methods nor the VE’s population and instead it was considered more reasonable to be further explored in future studies.

Additionally, the multiplayer session was populated by pairs of users only, which prevents the study of exploring further behaviors and interactions if more users would share the VE at the same time.

Finally, the interference to the hand and gesture detection due to the study’s settings, lower the significance of the results as the noise in the interaction method highly influenced the perceived experience from the users. Any future studies should regard this variable highly, as it is very significant to the quality of the experience.

7.3 Future work

Improving UX within games is a perpetual task. Even in the case of VR, where the medium is not well established within the consumer market, it is very important to maintain a user centered mentality in order to offer to them new ways they can experience a game and expand the entertainment boundaries. Having this in mind, the enthusiasm, or the lack of it, expressed by the users participating in these studies, was a very good indication of this work’s impact and where it should proceed from now on.

The results from the two different methods revealed a clear difference among them and especially on the effect they had on the experience. Out of the many, gestures were considered very natural while the use of touchpad was considered to be very precise. It is clear now that both methods bring something unique and significant to the experience and instead of having to choose one of them it would be interesting to see how the experience is formed by combining them. By utilizing the better body balance offered by the gesture method and the resting point offered by the touchpad, we could offer a reasonably more mobile experience where equilibrium is ensured.

In order to do so, the nature of these methods should be clearly mapped on the effect they have within the VE. As users suggested, seeing their hands interacting with the VE was something very
vigorous to them thus the use of gestures should be of equal impact. There is no need to assign actions to gestures that are tedious or frequently repeating as it will have diminishing returns on how they are perceived within the experience. Instead, gestures should be meaningful and used appropriately in the given scenario whereas the use of touchpad should be utilized with actions of lower significance that require high precision and robustness. Additionally, visual elements on top of the hand illustration such as cosmetic items or visual enhancements can potentially trigger the user’s imagination.

Regarding the occupation of the virtual space by many participants, this was something that was highly appreciated by the users. The human factor greatly influenced this experience both within the game setting as well as in the real world afterwards. For this reason, the multiplayer use should be further expanded by increasing the type of dependencies among the users. By introducing different roles within the game, users would be able to identify themselves differently according to the abilities and functionalities of this role. This will increase the complexity of the scenario thus making less probable of becoming boring and also expand the possible cooperation ways within the game. Additionally, by assigning different avatars according to this role, users would have a greater sense of presence within a game and receive additional information about their teammate.

Finally, it would be very interesting to expand the game dependencies even outside the VE. Since users were very fond of sharing their experience by posting pictures to SM after the game was finished, maybe a shareable event taking place in real time would be a fitting addition. Since the game is already networked, it would be an interesting approach to offer non VR applications which connect to the same server and allow spectators in the game session. The presence of spectators would potentially increase the actor mentality in the VR participant and by offering interaction methods to the spectators such as empowering their actions of the VR participant, the game dependencies are now taking place not only in the virtual world but to the real one as well.
References


computer vision and pattern recognition (pp. 3593–3601).
Kent, S. L. (2010). The ultimate history of video games: from pong to pokemon and beyond... the story behind the craze that touched our lives and changed the world. Three Rivers Press.


Appendices
Appendix A

Methodological tools

A.1 Interaction methods questionnaire

Figure A.1: Interaction methods questionnaire demographics.
**UX Questions**

How responsive was the environment to actions that you initiated (or performed)?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fully responsive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How aware were you of events occurring in the real world around you?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I was fully aware</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How natural did your interactions with the environment seem?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completely natural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How aware were you of your display and control devices - or mechanisms?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completely aware</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How inconsistent or disconnected was the information coming from your various senses?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completely</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure A.2: Interaction methods questionnaire - UX questions 1.
Were you able to anticipate what would happen next in response to the actions that you performed?

- Yes
- No

How well could you interact or manipulate objects in the virtual environment?

Not at all: 1 2 3 4 5 6 7

Extremely: 8 9

How involved were you in the virtual environment experience?

Not at all: 1 2 3 4 5 6 7

Completely: 8 9

How distracting was the control mechanism?

Not at all: 1 2 3 4 5 6 7

Completely: 8 9

How much delay did you experience between your actions and expected outcomes?

Not at all: 1 2 3 4 5 6 7

Very much: 8 9

Figure A.3: Interaction methods questionnaire - UX questions 2.
How quickly did you adjust to the virtual environment experience?

It took me a lot of time 0 0 0 0 0 0 0 Very fast

How much did the control devices - or mechanism - interfere with the performance of assigned tasks or with other activities?

Not at all 0 0 0 0 0 0 0 Very much

How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

Not at all 0 0 0 0 0 0 0 Completely

How proficient in interacting with the virtual environment did you feel at the end of the experience?

Not at all 0 0 0 0 0 0 0 Completely

Figure A.4: Interaction methods questionnaire - UX questions 3.
A.2 NASA - TLX

Figure A.5: NASA Task Load Index questionnaire.

**NASA Task Load Index**

Hart and Staveland’s NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

<table>
<thead>
<tr>
<th>Name</th>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
</table>

- **Mental Demand**: How mentally demanding was the task?
  - [ ] Very Low
  - [ ] Very High

- **Physical Demand**: How physically demanding was the task?
  - [ ] Very Low
  - [ ] Very High

- **Temporal Demand**: How hurried or rushed was the pace of the task?
  - [ ] Very Low
  - [ ] Very High

- **Performance**: How successful were you in accomplishing what you were asked to do?
  - [ ] Perfect
  - [ ] Failure

- **Effort**: How hard did you have to work to accomplish your level of performance?
  - [ ] Very Low
  - [ ] Very High

- **Frustration**: How insecure, discouraged, irritated, stressed, and annoyed were you?
  - [ ] Very Low
  - [ ] Very High

Figure A.5: NASA Task Load Index questionnaire.
Appendix B

Additional results

B.1 Numerical input used for the word histograms.

<table>
<thead>
<tr>
<th>Value</th>
<th>Frequency</th>
<th>Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>82</td>
<td>game</td>
</tr>
<tr>
<td>2</td>
<td>73</td>
<td>tap</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>fun</td>
</tr>
<tr>
<td>4</td>
<td>69</td>
<td>hard</td>
</tr>
<tr>
<td>5</td>
<td>54</td>
<td>repair</td>
</tr>
<tr>
<td>6</td>
<td>48</td>
<td>teleport</td>
</tr>
<tr>
<td>7</td>
<td>41</td>
<td>map</td>
</tr>
<tr>
<td>8</td>
<td>38</td>
<td>enemy</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>lives</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>position</td>
</tr>
</tbody>
</table>

Table B.1: Top 10 words users used for Button Tap interaction method.
<table>
<thead>
<tr>
<th>Value</th>
<th>Frequency</th>
<th>Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>94</td>
<td>game</td>
</tr>
<tr>
<td>2</td>
<td>87</td>
<td>tap</td>
</tr>
<tr>
<td>3</td>
<td>57</td>
<td>fun</td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>hard</td>
</tr>
<tr>
<td>5</td>
<td>42</td>
<td>repair</td>
</tr>
<tr>
<td>6</td>
<td>38</td>
<td>teleport</td>
</tr>
<tr>
<td>7</td>
<td>34</td>
<td>map</td>
</tr>
<tr>
<td>8</td>
<td>41</td>
<td>enemy</td>
</tr>
<tr>
<td>9</td>
<td>28</td>
<td>lives</td>
</tr>
<tr>
<td>10</td>
<td>24</td>
<td>calibration</td>
</tr>
</tbody>
</table>

Table B.2: Top 10 words users used for Gesture interaction method.