Exploring the Importance of Presence in VR Research

Elsa Ryrlén and Ruben Täpptorp

DEPARTMENT OF DESIGN SCIENCES
FACULTY OF ENGINEERING LTH | LUND UNIVERSITY
2026

MASTER THESIS

Exploring the Importance of Presence in VR Research

Is High Fidelity Necessary for Accurate VR Research Outcomes?

Elsa Ryrlén and Ruben Täpptorp

LUND UNIVERSITY

Exploring the Importance of Presence in VR Research

Is High Fidelity Necessary for Accurate VR Research Outcomes?

Published by
Department of Design Sciences
Faculty of Engineering LTH, Lund University
Box 118
SE-221 00 LUND
Sweden

Subject: Virtual Reality and Augmented Reality (MAMM20)
Supervisor(s): Günter Alce, gunter.alce@design.lth.se
Examiner: Joakim Eriksson, joakim.eriksson@design.lth.se

Abstract

This master thesis aims to evaluate how user test participants perform differently in Virtual Reality (VR) environments depending on the level of presence. It also examines how much effort and resources should be invested in increasing the realism of VR environments in a research context. This study was carried out in two primary parts; initially a development phase where a comprehensive information search was made in order to understand what actually results in better realism as well as implementing it, followed by a test phase. The developed environment was an aircraft setting in which an in-flight entertainment system was tested on a screen in front of the user. This serves as a clear example of how VR can significantly simplify testing compared to using a physical aircraft. The user tests were conducted by comparing two virtual environments with different degrees of realism. One of the environments included moving human characters, natural lighting, sound as well as interactive objects, while the other one was more scaled back without extra realistic elements. The results showed that the more realistic environment generated a higher level of presence and was experienced as more engaging and fun by the participants. On the contrary, the participants' performance, such as task completion time or mistakes made, was not affected to any greater extent. Instead it was revealed that other factors, such as the learning curve, had a bigger impact on the performance. The conclusion is that heightened presence in VR can contribute to a more engaging environment, but that the purpose of the study has to be considered when doing a cost-benefit analysis. In environments where the experience as a whole is of great importance, or in instances where participants engage a lot with the environment, a realistic environment can be crucial. However, studies with a high focus on solving specific tasks which do not necessarily require VR to be completed, a simpler environment will likely suffice.

Keywords: Virtual Reality, Presence, Usability Evaluation

Sammanfattning

Denna masteruppsats har som mål att utvärdera på vilket sätt nivån av *presence* inom virtuell verklighet (VR) påverkar användares prestation i en miljö när de utför uppgifter. Den utvärderar även hur mycket ansträngning som bör läggas för att göra VR-miljöer mer realistiska i forskningssammanhang. Denna studie genomfördes i två primära delar; först en utvecklingsfas där en gedigen informationssökning gjordes för att förstå vad som ger ökad realism samt implementera detta, följt av en testfas. Miljön som utvecklades i var en flygplansmiljö där ett underhållningssystem testades på en skärm framför användaren. Vilket är ett tydligt exempel på när testning i VR underlättar markant istället för att använda ett fysiskt flygplan. Användartesterna genomfördes genom att jämföra två virtuella miljöer med olika grad av realism. Resultaten visar att den mer realistiska miljön genererade en högre nivå av *presence* och upplevdes som mer engagerande och trovärdig av deltagarna. Däremot påverkades inte användarnas prestation, räknat i tid eller precision, då de genomförde uppgifter i någon större utsträckning av den ökade graden av realism. Istället framkom att andra faktorer, såsom inlärningseffekter, hade större påverkan på prestation. Slutsatsen är att ökad *presence* i VR kan bidra till en mer engagerande miljö, men att nyttan bör vägas mot kostnaden beroende på studiens syfte. För miljöer där upplevelsen står i fokus eller där deltagaren interagerar mycket med den virtuella miljön kan hög realism vara avgörande, medan studier med hög fokus på att enbart lösa uppgifter, uppgifter som inte behöver VR för att utföras, kan enklare miljöer vara tillräckligt.

Nyckelord: Virtuell verklighet, Presence, Användbarhetsutvärdering

Acknowledgements

Firstly we want to thank the collaborating technology company for their assistance with the work on this master thesis. Access to their office and all that comes with it in combination with great help from our assigned supervisor from the company has been of great help.

We would also like to thank our supervisor, Günter Alce, for his terrific guidance throughout the entire process of making this master thesis.

Lastly we would like to thank everyone who has participated in any of our user tests for taking their time to provide us with valuable information and insights.

Lund, May 2026
Elsa Ryrlén and Ruben Täpptorp

Table of contents

List of acronyms and abbreviations 11

1. Introduction 12
  1.1 Background 12
    1.1.1 Company Collaboration 12
  1.2 Purpose and Goal 12
  1.3 Research Questions 13
  1.4 Sustainable Development Goals 13
    1.4.1 SDG Target 9.5, Enhance Research and Upgrade Industrial Technologies 13
    1.4.2 SDG Target 12.2.1, Sustainable Management and Use of Natural Resources 14
  1.5 Related Work 14
    1.5.1 Prior Work on This Project 14
    1.5.2 Other Papers of Interest 15
  1.6 Ethical Consideration 15

2. Theory 16
  2.1 Virtual Reality Technology 16
    2.1.1 Virtual Reality 16
    2.1.2 Presence and Immersion 16
    2.1.3 Measuring Presence 17
    2.1.4 Uncanny Valley 17
  2.2 Data collection & analysis 18
    2.2.1 Interviews 18
    2.2.2 Affinity Diagram and Thematic Analysis 18
    2.2.3 Brainstorm 18
    2.2.4 MoSCoW Method of Prioritization 19
    2.2.5 Literature review 19
  2.3 Technology 19
    2.3.1 Meta Quest Pro 19
    2.3.2 Unity 19

2.3.3 Meta Horizon Link 20

3. Phase One 21
3.1 Evaluating Prior Projects 21
3.1.1 State of the VR Environment 21
3.1.1.1 Start Menu 21
3.1.1.2 Airplane Cabin 22
3.1.1.3 Eye Tracking 22
3.1.2 Changes to be Made 23
3.1.2.1 The Environment 23
3.1.2.2 Character Design 24
3.2 Development Process 24
3.2.1 Literature Review 24
3.2.2 Brainstorming 24
3.2.3 Prioritization 25
3.2.4 Storyboard 26
3.3 Design Process 27
3.3.1 Human Characters 27
3.3.2 Where to Get the Characters 27
3.3.3 Lights in VR 28
3.3.4 Interactable Objects 28
3.3.5 Sound in VR 29
3.3.5.1 Spatial Sound 29
3.3.5.2 Sound in Airplane 29
3.4 User test 30
3.4.1 Prior to the user test 30
3.4.2 During the user test 30
3.4.3 Post-test 31
3.5 Result 32
3.5.1 Prescreening 33
3.5.2 Sources of Error 33
3.5.3 Observations and Post-test answers 33
3.5.3.1 Presence and Realism 33
3.5.3.2 Characters and Animations 34
3.5.3.3 Sounds and Interactions 35
3.5.4 Takeaways 35
3.5.4.1 Major Takeaway - Lag 35
3.5.4.2 Major Takeaway - Animation 35
3.5.4.3 Major Takeaway - Presence and Realism 37
3.5.4.4 Minor Takeaways 37

4. Phase Two 39
4.1 New Implementations 39

4.1.1 Animations and Sound 39
4.1.2 Performance 40
4.1.2.1 Two Airplanes 40
4.1.2.2 Changing Scene 41
4.1.3 Interactable Objects 41
4.1.3.1 Seatbelt or Tray table 41
4.2 User Test 43
4.2.1 Measuring Presence 43
4.2.2 Test Environments 43
4.2.3 Choice of Touchscreen 43
4.2.4 Choice of Test Participants 44
4.2.5 Set Up 44
4.2.6 Pre-test Screening 45
4.2.7 Conducting the User Test 45
4.2.8 Choice of Tasks and Questions 47
4.2.9 Post-test Questionnaire 47
4.2.10 Post-test Interview 48
4.2.11 Pilot Test 48
4.2.11.1 Removing the Foldable Tray Table 49
4.2.11.2 Fewer Tasks 49
4.3 Result 50
4.3.1 Result of Pre-test Questionnaire 50
4.3.2 Result from IPQ 50
4.3.3 Analyzing the Interviews 51
4.3.3.1 Overall Insights 54
4.3.4 Analyzing User Groups 54
4.3.5 Task Completion Time 54
4.3.6 Errors 57
4.3.7 Takeaways 58
4.3.7.1 Realism Equals Higher Presence 58
4.3.7.2 Presence does not affect performance 58
4.3.7.3 Presence Affects Users Experience 58
4.3.7.4 Gamers are More Susceptible to Non-Realistic Events 59
4.3.7.5 Natural in Higher Presence 59
5. Discussion 60
5.1 Research Questions 60
5.1.1 In what way does the level of presence affect how a test person performs a task in VR? 60
5.1.2 When researching in VR, how much effort should be put into making the environment realistic? 61
5.2 Limitations 63

5.2.1 System Used in User Test 63
5.2.2 Participant Selection 64
5.2.3 Hardware 64
5.3 Technical difficulties 65
5.3.1 Meta Horizon Link 65
5.3.2 Rebuilding the Environment 66
5.4 Reflections 67
5.5 Future Work 67
5.5.1 Research with Different Task and Environment 67
5.5.2 Broader Group of Participants 68
5.5.3 Isolation of Realism Factors 68
6. Conclusions 69
References 70
A. Appendix 73
A.1 Pre Screening phase 1 73
A.1.1 Pre-test Questions 73
A.1.2 Pre-test Answers 73
A.2 Pre Screening phase 2 75
A.2.1 Pre-test Questions 75
A.2.2 Pre-test questionnaire results 76
A.3 IPQ 78
A.3.1 Items in the IPQ 78
A.3.2 Result from the IPQ 79

List of Acronyms and Abbreviations

FPS	frames per second
IFE	in-flight entertainment
IPQ	igroup presence questionnaire
NPC	non-player character
PQ	presence questionnaire
SDG	sustainable development goals
VR	virtual reality

1 Introduction

This chapter introduces the thesis. It contains the background to the subject of virtual reality in research. The chapter introduces the goal and purpose as well as its relation to the sustainable development goals and the research questions. Lastly related work is discussed.

1.1 Background

There are many instances where research can be facilitated by virtual reality (VR) [8]. VR creates opportunities to perform tests that have previously been too expensive and inconsistent but can now be carried out in a more flexible manner. The technology is especially useful in industries where performing prototyping, early tests, and demos is expensive or heavily time consuming, such as the aviation industry.
A lot of research has been done on how to streamline scientific research processes, but there is a lack of exploration in regards to specifically making research in VR more effective [11]. This thesis will explore the possibilities of making research in VR more efficient through performing tests in virtual airplane cabins of varying detail. In general, research will be able to become more efficient in terms of better prioritizing which VR elements to implement.

1.1.1 Company Collaboration

The thesis was done in collaboration with a mid-sized technology consultancy. The company works in the in-flight entertainment (IFE) industry, therefore a virtual airplane environment is relevant to their work and they were capable of providing sufficient aid throughout the project.

1.2 Purpose and Goal

The purpose of this thesis was to explore ways to better facilitate VR research. More precisely, to help businesses and researchers improve their priorities and budget when researching in VR. This was done through developing a realistic airplane cabin environment in VR as well as an environment with less realistic features as it was a stripped down version of the same airplane cabin. Then through comparing human behavior and measuring performance in both environments, it could be determined whether or not it is worth spending extra resources and time on developing the realistic environment.

1.3 Research Questions

The thesis's purpose is here presented more precisely through the research questions.

1. In what way does the level of presence affect how a test person performs a task in VR?
2. When researching in VR, how much effort should be put into making the environment realistic?

1.4 Sustainable Development Goals

This thesis also helps the Sustainable Development Goals (SDG) [22], with focus on target 9.5 and 12.2.

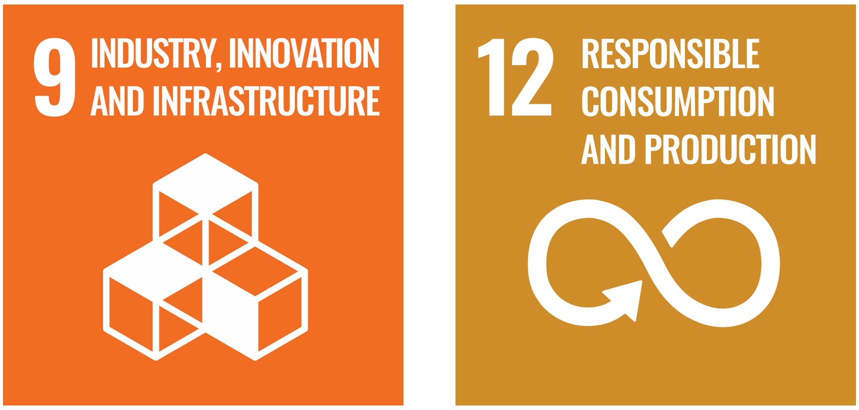


Figure 1.1 Sustainable development goal 9 and 12.

1.4.1 SDG Target 9.5, Enhance Research and Upgrade Industrial Technologies

In United Nations publication *Transforming our World: The 2030 Agenda for Sustainable Development* Target 9.5 aims to enhance scientific research, encourage innovation and *"substantially increasing the number of research and development worker"* [22]. This work aims to better understand the expenses of research in VR and ultimately lower the costs. Making research affordable is a great step to enhance scientific research. Figure 1.1

1.4.2 SDG Target 12.2.1, Sustainable Management and Use of Natural Resources

Target 12.2 states *"By 2030, achieve the sustainable management and efficient use of natural resources"*. Indicator 12.2.1 more specifically states *"Material footprint, material footprint per capita, and material footprint per GDP"* [21] Figure 1.1. This thesis will support the goals of this indicator in two different ways. It will help incentivize the use of VR research by giving investors and research teams an improved understanding of the costs of VR research. Resulting in utilizing less resources compared to having physical prototypes or environments thus minimizing material footprint. In addition, this thesis intends to lower the costs of VR research. Both presenting a clearer picture of the cost as well as possibly lowering the costs promotes research in VR.

1.5 Related Work

The following section presents previous work and relevant research to this thesis. It includes earlier versions of this project as well as scientific literature regarding VR, presence, and evaluation methods used throughout the study.

1.5.1 Prior Work on This Project

What was developed during this thesis is not the first version of this project. This is version three, and two other master theses have been completed working on prior versions of this project. Both of those projects were also done in collaboration with the same technology consultancy as this thesis.
The first version was done in 2023 by Andersson and Voigts during their master thesis [1]. During Andersson and Voigts' master thesis they built the initial version of the airplane cabin environment. The primary objective of their thesis was to develop a modular virtual airplane cabin environment. The purpose was to be able to set up tests in VR in place of expensive tests with a physical airplane cabin.
The second version was done in 2024 by Do and Phung Nguyens during their master thesis [5]. Do and Phung Nguyens used Anderssons and Voigts [1] first version of the airplane testing environment and improved upon it in many ways, but mainly through adding eye tracking to the testing environment. This was done to gather more valuable data from the tests.

1.5.2 Other Papers of Interest

In this master thesis it is highly important to accurately understand and describe presence, given that presence is at the root of the research questions. Therefore, it is of relevance to measure the level of presence in an environment. How to measure presence is not an exact science, but the *igroup*, a research group, including Thomas Schubert and Holger Regenbrecht, developed a questionnaire in the early 2000s to measure the sense of presence [9]. The term presence and how to measure it is further discussed in subsection 2.1.3.

1.6 Ethical Consideration

Being in a VR environment includes only seeing the screens of the glasses and the controls in ones' hands in some shape. This can be a strange setting for many people who have not yet tried it, since one cannot see their own body or the real surroundings anymore. It is meant to feel realistic which also makes animations feel like it is happening to oneself - falling or flying in a VR game feels a lot more real than falling in a TV or computer game. Therefore, it is important to inform participants taking part in a VR test, that they can terminate or pause the test at any point if they feel dizzy or nauseous, but also physically or emotionally uncomfortable [16]. Also that there will be no consequence if they do so, to make sure no one feels obligated to continue even though they do not want to. It is due to the realism of VR that it can trigger real emotions and cognitive load such as stress, fear or agitation. Furthermore, when conducting a test in VR, it is important to monitor the participants to make sure they do not physically injure themselves by bumping into objects or trip on their surroundings.

In addition to the physical and psychological considerations associated with virtual reality exposure, ethical responsibility also includes collecting and handling of participant data. During the VR testing sessions, data is usually collected in the form of questions regarding information about the user, observational notes and recordings of their interaction. All participants needs to be informed in advance about what type of data is going to be collected, the purpose of the data collection as well as how the data will be used within the research. It is also ethical to have voluntary participation where the users are asked for consent as well as being anonymized in the report [16].

2 Theory

This chapter presents the essential theories used in the thesis. It contains theories about Virtual Reality, design concepts and data collection and evaluation methodology.

2.1 Virtual Reality Technology

2.1.1 Virtual Reality

VR is a computer-generated simulation system consisting of three main components: hardware, interactive devices and virtual content. The hardware and software platforms enable real-time rendering of a three-dimensional environment, while interactive devices allow users to navigate and interact with the virtual world usually using hand controllers. VR technology can result in a high level of realism, making it suitable for a wide range of simulation, training, and usability evaluation applications as well as highly interactive games [13].

2.1.2 Presence and Immersion

The key factor for a user to have an enjoyable and realistic experience in a VR environment is to have high presence and immersion [3]. Su et. al define presence, in a virtual scenario, when the user feels physically present in the virtual environment even when one is physically in another place. Unlike immersion which is reached through technical advancements e.g. achieving low latency or high quality video. High immersion is therefore achieved when the technical modules support the person in such a way that they do not notice the technical limitations of the hardware or software. A virtual environment with greater sense of immersion will produce higher levels of presence. The main approach presented on how to design a virtual environment to increase the immersion is to have high interactivity, this promotes cognitive engagement and triggers emotional response, thus making the experience of presence more vivid.

2.1.3 Measuring Presence

Measuring presence is not easy. Presence is a subjective feeling for each person and difficult to distinguish from other parts of the experience in a virtual setting [19]. There have been multiple efforts to try and measure presence with questionnaires, which gives the test participants a chance to express their subjective thought. Physical measures such as eye movements or heart rate have also been used when trying to measure presence.

One widely used questionnaire is the Igroup Presence Questionnaire (IPQ), which was created in 1997 and is still developing till this day [9]. Its goal is to be used as a subjective measurement of presence for virtual environments. It consists of 14 items rated on a 7-point Likert scale, usually between -3 and 3. All the questions in the questionnaire is divided into four subscales:

- **General Presence (GP)** – the overall *sense of "being there"* in the virtual environment
- **Spatial Presence (SP)** – the feeling of being *physically located* within the virtual environment
- **Involvement (INV)** – the level of *attention and engagement* directed toward the virtual environment
- **Experienced Realism (Real)** – the perceived *realism* of the virtual environment

In addition to its structure, the scoring procedure of the IPQ requires careful handling of certain items. On the scale -3 represents the lowest level of agreement and 3 the highest. Before calculating mean scores, it is necessary to be aware of the reverse-coded items, as some questions are negatively worded and their scores needs to be reverted, e.g -2 becomes 2. After reversing these items, mean scores can be calculated for each subscale as well as an overall score.

2.1.4 Uncanny Valley

Uncanny Valley is a psychological concept where human-like objects such as robots, CGI or characters has to much of a resemblance with real humans, causing an unconformable feeling inside us [17]. As artificial agents become increasingly human-like, people's affinity toward them generally increases, however, this positive response only persists up to a certain point. When the likeness becomes highly realistic but still contains subtle imperfections in appearance, movement, or behavior, the emotional response can sharply shift from familiarity to discomfort creating a creepy feeling. This dip in affinity is referred to as the *uncanny valley*, where entities appear almost human but fail to fully meet behavioral expectations, resulting in a disturbing or unsettling experience.

2.2 Data collection & analysis

The following methods describes how data is collected and analyzed throughout this study.

2.2.1 Interviews

Conducting interviews after performing a test is an effective way to gather insightful information regarding users' experiences and opinions. It often provides more detailed and in-depth answers than simple questionnaires. Semi structured interviews consist of a set of pre-decided questions which also can have follow up questions if the interviewers want to gather even more information regarding an interesting topic or dig deeper into a vague answer [4].

2.2.2 Affinity Diagram and Thematic Analysis

A thematic analysis is a qualitative research method used to identify, analyze, and interpret themes within a set of data. The data can be retrieved from interview transcripts, observations, or open-ended survey responses. Affinity diagrams are a method to perform the thematic analysis. Krause and Pernice accurately defines affinity diagram as *"Affinity diagramming refers to organizing related observations, ideas, concepts, or findings into distinct clusters."* [12]. Given that affinity diagrams are used to organizing concepts and ideas, it is often used in conjunction with idea generating activities such as brainstorming, to organize all of the thoughts that has been shared. Usually there are no rules on how to categorize the ideas, but rather it is encouraged to figure out relevant categories based on the ideas at hand.

2.2.3 Brainstorm

Having a brainstorming session is a great way to come up with, explore, and discuss different ideas when starting a new project [10]. A typical set up is having a timer on a few minutes and having each participant writing their ideas, as many as they can think of, on post-it notes and placing them on a table, each idea gets its own post-it notes. When the time is up, all notes are read out loud and discussed. No idea is called stupid or neglected, which often makes this method effective and generates many ideas.

2.2.4 MoSCoW Method of Prioritization

MoSCoW method is a prioritization technique for helping a team to understand and manage priorities [6]. The method consists of ranking potential features into one of the four categories, whereas the letters stand for:

- Must Have
- Should Have
- Could Have
- Won't Have

"Must haves" indicates that the final project will absolutely have these features and there is almost no point of finishing the project without it. 'Should haves' is important but not vital and 'Could haves' would do less impact if left out compared to the latter. Finally the 'Won't haves' has been decided to not be delivered at all.

2.2.5 Literature review

A literature review is a review of the published literature regarding a specific field [18]. Literature reviews are therefore typically done at the start of research projects to establish what is known about the field, and where new research can continue. It is commonly done by searching for articles and scientific papers related to the field. Skimming through a lot of them, as well as doing a full review of the ones that are deemed the most relevant.

2.3 Technology

This section describes the technology used in this master thesis to build and run a VR environment.

2.3.1 Meta Quest Pro

The meta quest pro is the VR headset used for this project. It was released by Meta in October 2022 with features such as a new advanced display and high-resolution sensors, which made it their most advanced headset at the time [14].

2.3.2 Unity

Unity was created in 2005 and was created as a game engine that developers could use instead of having to create their own engines. Unity has grown to become the most popular game engine, especially for smaller studios. Out of all the games that were released in 2024, 51% were made in Unity [24]. Unity now claims they are a

’real-time engine’, meaning that the engine is capable of being used for much more than to create games [23]. Unity has lots of support for development in VR.

2.3.3 Meta Horizon Link

Meta Horizon Link, previously Meta Quest Link, is a software that allows the user to connect their headset to the computer, either via USB cable or Wi-Fi. This has many benefits such as the ability to cast from your headset to your computer and overall better performance. When developing VR projects with Meta’s products, Link also speeds up the development process by eliminating the need to build a complete version of the project for each test [15].

3 Phase One

This chapter presents the methods used to structure the first phase of the project and the result from them. It also explains design choices and how they were made. Finally it discusses how the first test was conducted including its result and analysis.

3.1 Evaluating Prior Projects

At the start of this project we choose to evaluate the previous versions. Our assigned supervisor at the company was knowledgeable about them and helped by guiding us through the previous version of the project.

To understand how we could improve upon this project we firstly had to familiarize ourselves with the current version of the program. This was done through thoroughly exploring the environment as well as reviewing major code components.

3.1.1 State of the VR Environment

3.1.1.1 Start Menu

The previous project had a menu scene for setting up the environment and choosing simulation settings. In the start scene the synchronization between the virtual IFE and the real physical one is done. This is done through mapping and marking each corner of the physical screen with the corresponding corner in the virtual world. In the settings the user can also modify some aspects of the setup of the virtual airplane environment. For example on which side of the plane the user will sit on, if there should be other passengers on the plane and whether it is day or night see Figure 3.1.

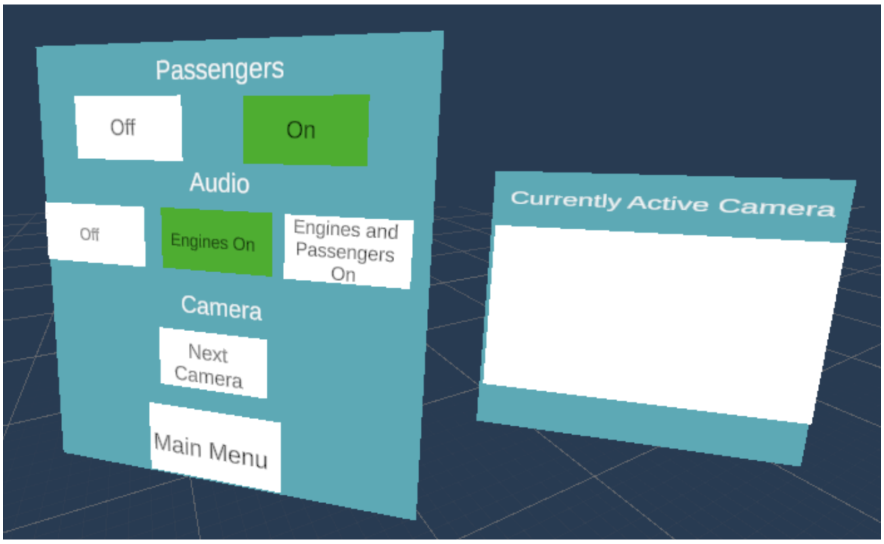


Figure 3.1 Part of the set-up menu

3.1.1.2 Airplane Cabin

When the setup is done the user is loaded into the airplane. The user is now inside of the airplane cabin. The cabin is filled with three rows of two chairs, see Figure 3.2. Each chair has a black screen on the back of it, apart from the users which has the IFE running. The user's seat is also supplied with an interactable safety card, although the card is not affected by gravity. If the passenger option was turned on in the settings, every other seat, apart from the user's, is occupied by an immobile robot or dummy, see Figure 3.3. Through the windows the user can see the clouds below them. Depending on the time of day chosen in the set up, there will be a brighter light during the day and a not as intense and dimmed light during nighttime. For interior design there are seatbelts signs and lamps in the ceiling. The lamps have accompanying interactable buttons which turns the lights on and off.

3.1.1.3 Eye Tracking

The program also features eyetracking functionality. With eyetracking turned on there is the possibility to, after the simulation is over, produce a heatmap based on where the user was looking during the test.

Figure 3.2 View of the airplane cabin from an earlier version of the project.

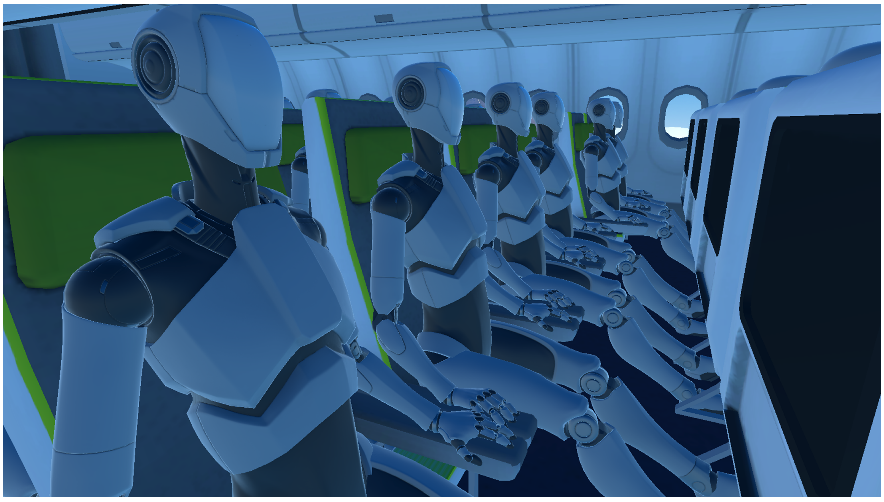


Figure 3.3 Robots in their seats.

3.1.2 Changes to be Made

3.1.2.1 The Environment

Although the previous VR environment worked well, there were aspects of all parts of the program that could be improved upon. However, given that this projects main focus is to research the impact of improved presence in VR, the scope of the project had to be focused on only improving upon the airplane cabin and leave the eye tracking and starting menu as is.

3.1.2.2 Character Design

In this phase there is also a huge emphasis on the human NPCs that will populate the plane and what they will look like. Choosing a design for our human NPCs, two alternatives were discussed. Either we develop and design the characters ourselves. The benefit of this approach is that we can design every little detail exactly how we like it, only limited by our own skill. With the main drawback being that it would take a lot more time than downloading already accessible character designs. When discussing the alternative of working with pre-made character designs, the pros and cons switch in comparison to making our own. It is significantly more time efficient to find assets online than to make them from scratch, but we cannot make any precise decisions regarding the characters' looks. The same goes for character animation. Making them ourselves could be more precise, but it is more time efficient and simple to download pre-existing ones.

Given that we had limited experience working with animation, it was a rather straightforward decision to work with pre-existing animations instead of spending time on learning animation ourselves. We do have some experience working with 3D computer character design. However, after an analysis of the assets available online, it was concluded that those designs will suffice. They are also most likely better than what we can create in this time period and will save a lot of time that can be spent on other design aspects of the environment.

3.2 Development Process

During the first development process, ideas of what to implement was generated and prioritized in order o create a joint understanding of where to lay our focus. Later on, characters and animations along with lights and sound were implemented in order to conduct a first small user test.

3.2.1 Literature Review

A literature review was conducted in the early stage of the thesis work. The study was comprehensive and focused on finding relevant information about VR, presence and how to measure the usability of these components.

3.2.2 Brainstorming

After interacting with and analyzing the current environment, it formed the basis for future work. We wanted to generate ideas for what to modify and add to make the setting feel even more realistic and provide higher presence. The best way to generate many ideas in a short period of time was to perform a brainstorming session. During ten minutes we wrote down potential ideas of what could be added or modified on post-it notes and placed on a table. After the time had passed, all the ideas were read out loud, discussed and sorted in smaller groups together with similar topics as well as given a overlaying category, see Figure 3.4.

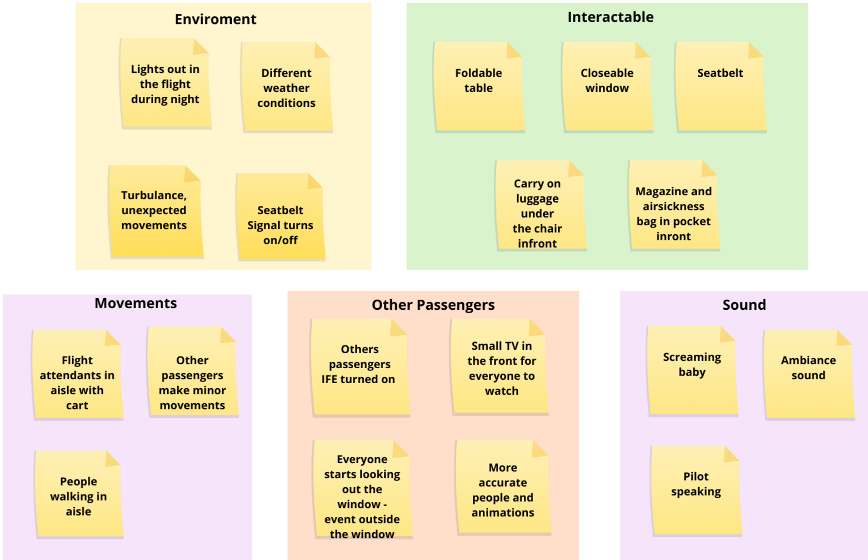


Figure 3.4 Brainstorming session with post-its

3.2.3 Prioritization

The implementation ideas were sorted according to the MoSCoW method of prioritization, thus ranked from 'Must haves' to 'Won't have', see Figure 3.5. From this prioritization it was established that the most important component that had to be implemented, in the first phase, was adding human characters into the scene. The robots and dummy passengers of the previous version of the project made the authors feel uneasy and it felt as an obvious improvement to include more human like passengers. Animations and movement was also highly prioritized. Furthermore in the previous version, apart from the user itself and their screen, there is no movement in the simulation. Movement was therefore one of the bigger disparities between real life and this simulation. Andersson and Voigt discussed adding movement into their version of the project, however it was decided it took to much resources and was pushed to 'future work' [1]. Therefore, animation felt like a natural continuation of the project.

Most of the items in the 'Should have' section were planned for implementation during phase two, such as 'Pilot speaking' and 'Foldable table'. All of the ideas in the 'Won't have' section were scrapped together with most of the items in 'Could have', with some exceptions being planned as extra features for phase two, only to be implemented if everything else was done and there was time left.

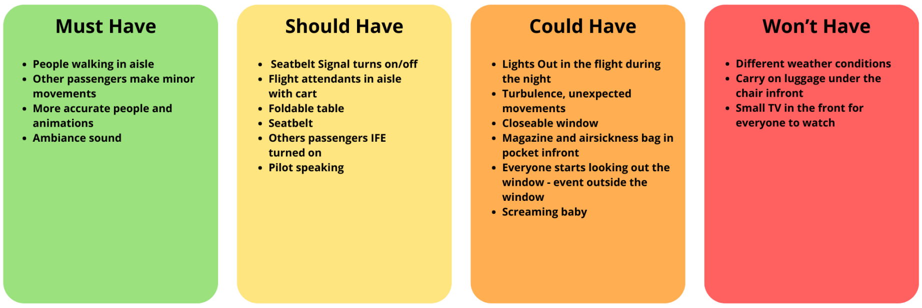


Figure 3.5 Prioritizing implementations according to MoSCoW

3.2.4 Storyboard

A simple storyboard was drawn to visualize a narrative of a user experience in a hypothetical scenario of using the VR environment, see Figure 3.6. This approach also provided a clear representation of the narrative flow and served as an efficient way of establishing a shared understanding of the project between the authors.

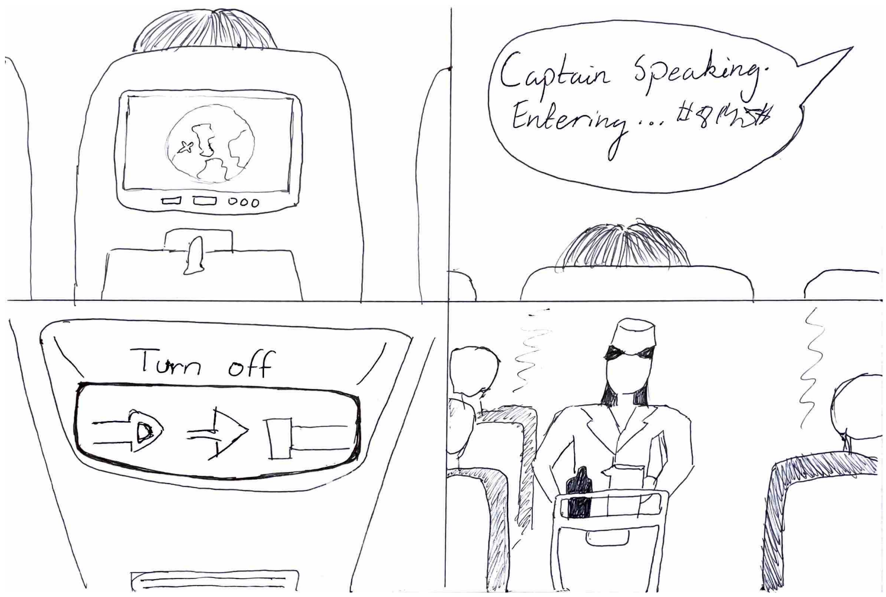


Figure 3.6 Storyboard

3.3 Design Process

The design of a VR environment is important, especially in this case where there is a large emphasis on the participants perceiving high presence. This section details the major design aspects that were addressed during phase one and how they were handled in the context of this project.

3.3.1 Human Characters

For this phase the most important design choices to be made were those of the human characters in the scene. A big issue that was discussed was that of the uncanny valley effect. The uncanny valley effect was the reason for Andersson and Voigt to include robots and dummies in the airplane instead of people [1]. The scope of their project excluded in depth focus on designing the environment, meaning they would not be able to spend the necessary time to make the passengers feel realistic. Therefore they populated the plane with robots and dummies instead, to avoid the possible issue of users experiencing the uncanny valley effect if the plane instead were populated with poorly designed human-like characters.

3.3.2 Where to Get the Characters

There are many assets of human characters available online. First Unity asset store and Mixamo [1] was searched for assets that may fit the style of the environment. One specific asset in the Unity asset store containing nine npc characters was early pinpointed as an alternative. However, apart from the uncanny valley effect, there are other aspects that has to be taken into account when choosing what asset to use. Other websites providing more realistic characters than the ones found on Unity asset store and Mixamo were also considered, but those can cost a lot more. Also more realistic characters are often more intense for the computer to handle than simpler characters. There are also legal implications to take into consideration. Given that this project is done in collaboration with a company, we have to take into consideration some legal aspects of what they can and cannot use if they were to use this project. Together with input from our supervisor at the company we concluded that the free assets from Unity asset store as well as some characters from Mixamo would satisfy the requirements of this project. The asset pack and characters required no licensing, was free and even though not as realistic as some others found, was still realistic enough to work in this project, see Figure 3.7.

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

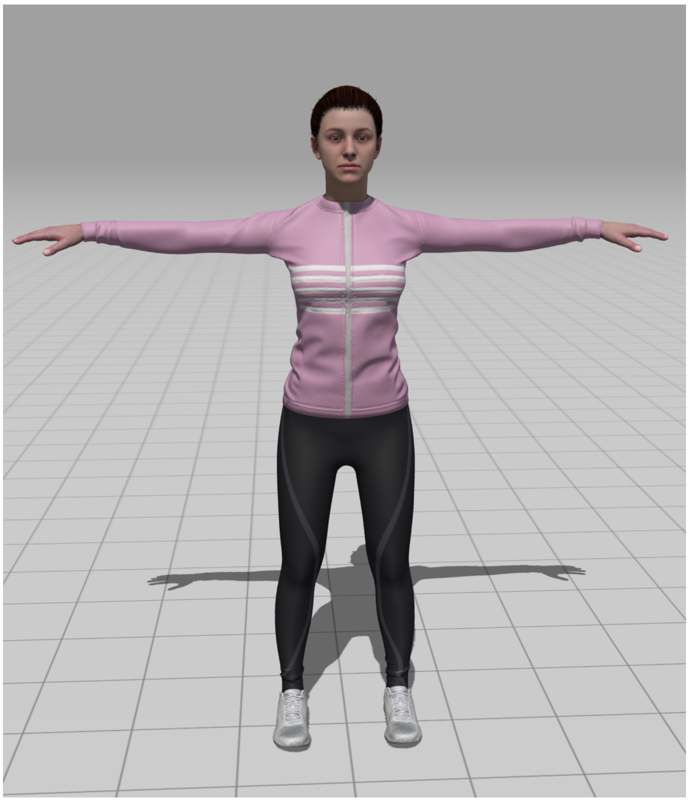


Figure 3.7 One of the characters used in the project, downloaded from Mixamo.

3.3.3 Lights in VR

After spending a significant amount of time in the VR environment, we began experimenting with lights and shadows to achieve more depth and realism. In Unity, objects' lightning can be rendered differently, one way is *baked lighting* which means the light is rendered once in the beginning, resulting in shadows being static, even though an object moves. This is unlike *real time lighting* where illumination is calculated every frame, resulting in dynamic light as objects move. The latter alternative appears more realistic but is also the most demanding for the GPU. Lastly there is incorporated mixed lighting which is a combination of baked and real time. We choose to use all three rendering options depending on whether the objects were static and dynamic, making sure the shadow followed them, but still saved processing power wherever possible.

3.3.4 Interactable Objects

It was decided to add some more interactable objects to the scene, with the intention of making it more realistic and fun for the user. Therefore, on the windows, an interactable airplane blinder was implemented. Due to realtime lighting, when the binder is up, light from outside the plane lights up the seating area. When the binder is down, the sunlight is blocked, just as it would be in the real world.

The exiting safety card was grabbable and easy to pick up from its pocket in the seat in front. However, it was not affected by gravity making it possible to place it stationary in the air resulting resulting in unrealistic behavior. Therefore we decided to add gravity so the user could place it back in the pocket or drop on the floor.

3.3.5 Sound in VR

Lastly, it was time to cover sound for the VR environment. The sound design was important to get right given that it is a whole other sense and has the possibility to add a lot of depth to the scene.

3.3.5.1 Spatial Sound

Having sound in VR can be a powerful tool to direct the user's attention and provide information to increase their understanding of an environment [2]. If a sound appears from the right, the players intuition will most likely be to look towards that direction. This can be done in VR since it uses spatial audio, which is an immersive 3D sound technology where audio elements can be placed in a 360-degree area. This creates the feeling that the sounds can come from behind or next to the player and changes when the head is moved. This technology differs from traditional computer and TV games which uses 2D sound and does not provide the same realistic experiences. This is how sound is not only used as a steering tool but also to create higher presence and degree of realism [7] highlights how sound is essential to how people cognitively perceive their surroundings. Background noise creates spatial awareness by communicating spatial information about the surrounding, even in case of perceived silence [7] also suggests that virtual environments without spatial sound lack the complex qualities required for a convincing immersive experience.

3.3.5.2 Sound in Airplane

For this reason selecting the appropriate sound in our virtual environment was crucial to achieve immersion. The goal was to maximize the user's sense of presence and to create the impression of being physically inside an actual airplane cabin. Common sounds to hear in an airplane environment includes the continuous low frequency hum of the engines, conversations or movements from other passengers, along with announcements from the flight crew. These layered sound components created both a recognizable experience as well as a realistic scenario. However it was important to avoid excessive volume levels or too many simultaneous auditory events sound as this could instead create a stressful or distracting experience for the user. Therefore we also chose to include sound in the first test to receive feedback regarding perceived realism, comfort, and overall immersion. The aim was to evaluate whether the audio chosen so far enhanced the sense of presence

3.4 User test

The main goal of the first test was to help ensure that the chosen characters were nicely implemented, realistic enough as well as provide an overall pleasant experience, but also as mentioned in previous section, testing the sound. We also wanted to receive valuable feedback regarding the general components of the environment and identify aspects that could be improved in the future. Choosing the correct characters and their animations was one of the most crucial parts in the projects, which is why it was important to include in the first test.

3.4.1 Prior to the user test

The test consisted of eight participants. Before the test, a pre-test screening, in the form of a small interview, was held in order to gathered background information about the participants, see Appendix A.1.1. This screening included questions about general information about the participants, prior VR experience and potential VR-related discomforts such as motion sickness or dizziness. The participants were also asked if they normally wear glasses or contact lenses, and if so if they felt comfortable using the VR headset with them. This was important to ensure both participant safety and comfort, but also to reduce the risk of discomfort affecting the test results. We also explained the term presence in a VR context, it was necessary that they understood this since the post-test included questions regarding presence. Finally, it was clearly stated that the participants could terminate the test at any time without consequences, and that they were anonymous and that the information gathered was to be used only in our study and could not be traced back to them.

3.4.2 During the user test

The test consisted of three parts:

- Testing the airplane environment in VR with robots together with passengers and engine background noise.
- Testing the airplane environment in VR with animated NPCs together with passengers and engine background noise along with other layered sounds.
- Post-test questioning about the experiences.

To avoid gathering biased information, half of the participants started in the robots environment and the other half in the NPCs environment.

During the test the participants were given different simple tasks and questions to answer or perform regarding the environment they were in, see Table 3.1 and Table 3.2.

Table 3.1 Tasks during NPC environment

There is a woman wearing a sweater with a cat on, what color is her sweater?
What colors are on the airplane wing?
What would you do if you had wanted to cover the sun for a while?
What color is the sign that says to fasten your seatbelt, are they all the same color?
How many passengers have been up walking during this test?
Try to turn on a light

Table 3.2 Tasks during robots environment

How many passengers would you estimate is on the plane?
What is the weather like outside the airplane today?
What color is the sign that says you are not allowed to smoke? Are they all the same color?
Which row do you think you are sitting in?
Where would you look for information about flight safety?
Try to turn on a light

The goal was not to measure how fast or easy they could perform the tasks but rather how they felt in the two different environments and in a natural way have them observe their surroundings. It was also important not to ask specifically about their opinions on the sound or characters but rather have them observe it. We did not include the actual in-air-entertainment of the actual system during this first test since the main focus laid on feedback of environment around.

3.4.3 Post-test

The questions asked after the test were carefully phrased to be open and not leading toward any specific direction in order not to affect the participants. The interview was semi-structured, meaning the questions were decided in advance but could have followed up questions in case something interesting was brought up and we wanted to have deeper insight in it. The questions asked were:

- Did you notice any difference between the two tests?

- Did you feel more or less presence in any of the scenarios?
  - In what way?
- Did the environment and its content feel coherent aesthetically and proportionally?
- What did you think about the other passengers on the flights?
- Did you notice any change in animations in the two environments?
- Did you notice different sounds between the two environments, and what did you think about it?

3.5 Result

The result of the test gave several important insights and takeaways to continue working on with the project. Since the test was based on observations and open interview questions regarding the participants' subjective experience, it generated qualitative data for analysis. The data was thematically analyzed by identifying recurring patterns and translating these into general codes. The codes were then grouped into different categories and organized in an affinity diagram, see Figure 3.8.

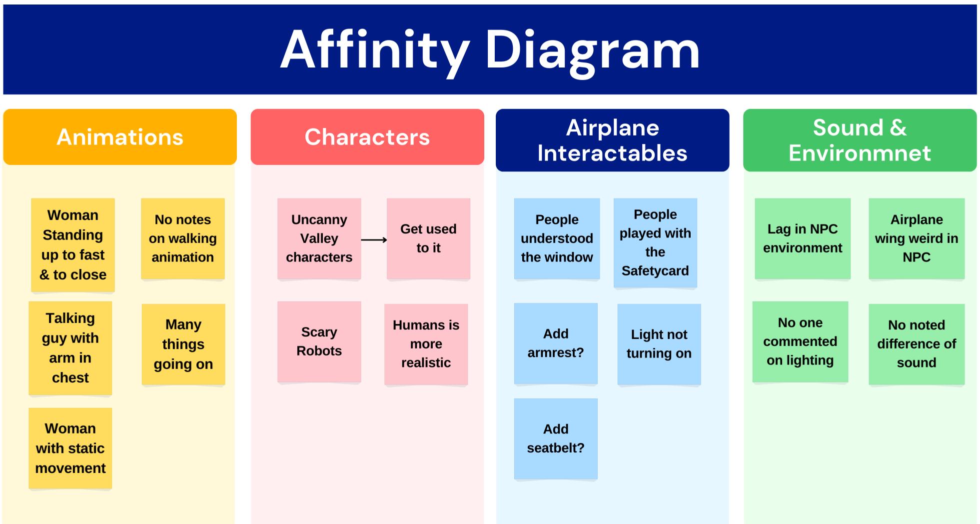


Figure 3.8 Affinity diagram from results of tests in Phase 1.

3.5.1 Prescreening

There were eight participants in total. Two of them were women and six were men. All of the participants were in the age between 22 to 26 years old and were studying as their main occupation. All participants, except for one, had experienced VR once or a few times in their lives, while the remaining participant had never tried it before. Those who had used VR had done so in various contexts, including playing games, development activities and watching movies, see Appendix A.1.1. All of the participants were acquaintances off the authors and therefore studying similar engineering programs. This was a consequence of conducting the tests on the premises of Lund University, which facilitated participant recruitment.

3.5.2 Sources of Error

During the first participants test, there was several lights sources in the NPC environment which did not work the way they should but still consumed a significant amount of processing power, thus resulted in quite a lot of lagging and poor quality. The render was slower when looking around and the animations appeared choppy. The participant frequently commented that due to the lag the experience and characters were not as immersive as the robot environment which had better quality. We noticed this during the participant's test and removed some unnecessary components which used processing power to prevent the same problem from affecting other tests. As a result, the data from the first test is considered somewhat unreliable.

During one test, the sound was not turned on. This was noticed after the first environment had already been experienced, and it was therefore decided to exclude sound from both environments to avoid unfairly influencing the participant's perception of either one. Therefore this participant was not asked the question regarding the difference in sound.

During all of the tests, the computer connected to the VR headset produced a loud fan noise, which many participants commented on. Some participants also mentioned afterward that they were unsure whether they heard sound from the headset itself or only the computer fan. This unfortunate inconvenience may have affected the participants' perception of the sound experience.

3.5.3 Observations and Post-test answers

3.5.3.1 Presence and Realism

The first two questions of the post-test interview were regarding the presence of the environment. This naturally also led the discussion to how lifelike the environments were.

Half of the participants described feeling more present in the NPC environment. Three participants described feeling no difference in regards to their presence and one person felt more present in the robot environment. Worth noting was that this was the very first participant, who experienced more lag due to poor set up of the test. Six out of the eight participants stated that the NPC environment felt more realistic and familiar, resembling a real-world airplane scenario. Participants expressed impressions such as:

“...different faces and clothes make it feel more real”

“... felt more realistic that they moved and behaved”

3.5.3.2 Characters and Animations

Three of the questions were specifically about the characters and animations. However, these topics were also mentioned by participants in response to other questions and during the observational part of the test. Three people stated they experienced the NPCs as a bit uncanny and unpleasant. However they also included that after a while they got used to it and it felt better. These are some quotes regarding characters:

“The first reaction was an uncanny valley feeling because you weren’t used to it, but then once you got used to it, it wasn’t something I gave much thought to”

“At first it was weird that everyone looked strange, but then the immersion became better”

However two participants thought the robot environment was somewhat scary due to the completely still robots and that it felt *“Matrix-like”*. Furthermore there were three participants who flinched and pulled back when the NPC next to them stood up and walked away. Six participants did not criticize the animations at all, but there were two negative comments about the specific stand-up animation:

“Wow! She flew up”

“The woman next to me standing up was jarring.”

One question during the test in the NPC environment was to estimate how many characters they think had been walking in the aisle. All participants answered between three to six people. No participants commented on the NPCs’ walking animations either during or after the test, unlike the standing animations, which received remarks.

By observing the participants’ behavior during the two experiences, a clear difference could be noticed. In the NPC environment, all participants looked around at the other passengers and their movements much more throughout the test. In the robot environment, once they realized that the characters remained seated and still, they appeared less curious and did not look around as much.

3.5.3.3 *Sounds and Interactions*

During the test, the participants were asked to interact with the different interactive components; the safety card, the light switch and the window blinder. All participants easily located each component at the correct place but had different difficulties performing the correct action to interact without guidance. The overall response when interacting with the components was positive. Half of the participants played around with the safety card by turning and tossing it. Five participants did not succeed turning on the lamp above them, some despite pressing the correct button and others due to pressing the light bulb instead. Everyone achieved pulling down the window blinders but two required a hint from the test-leader that it was possible to touch it.

In the post-test interview, participants were asked whether they noticed any differences in sound between the two environments and what their opinions were about it. Four out of the seven participants who experienced the environments with sound stated that they did not reflect on the sound and therefore did not notice any differences. The remaining participants remembered the sound but were unsure whether there was any difference between the environments. Two of them clearly stated that they enjoyed the seatbelt "ding" sound effect. But as mentioned in source error, the sound was affected by the loud fan noise right next to the participant.

3.5.4 Takeaways

From these observations three major takeaways were formed. The takeaways were that performance matters a lot, that some of the animations could be improved in design, and that participants felt more presence and experienced realism in the NPC environment.

3.5.4.1 *Major Takeaway - Lag*

The NPC environment had notable lag. The issue was that sometimes when moving their head, the screen would flicker. In addition, there were also occasional drop in FPS. This severely distracts the participant and damages the illusion of presence. This is a typical example of how immersion affects presence.

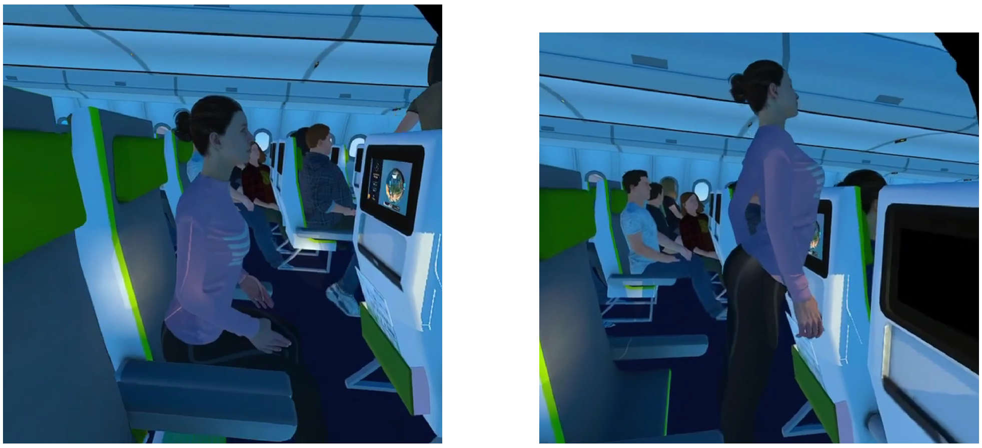
The solution for this is to inspect the environment and research what requires most processing power and find a balance between a realistic environment and its affect on performance. This is further discussed in phase 2.

3.5.4.2 *Major Takeaway - Animation*

The poor animation was a result of a handful of smaller issues. The most notable animation is the woman sitting next to the player standing up and walking away. The woman starts with an idle breathing animation on loop, just similar to all the other NPCs around her. After 15 seconds she rises, turns left out of the seat, takes a step into the aisle, turns right to align herself with the aisle and lastly starts walking down the aisle.

There are several reasons for why this animation sequence could be interpreted as jarring.

- **Close to participant** - Given that the character is so close to the participant, and that it is such a big movement, it is almost impossible for the player to miss it.
- **Aggressive movement** - The transition between her sitting pose and her standing pose was deemed aggressive. The time between the character starting the animation, showing the visual cues of a "standing up motion", and the point in which she was stood up was too short.
- **Physique** - The woman has an unnaturally arched back, which is only noted once she stands up see Figure 3.9a. There was also a general comment about the characters straights back which easily could be fixed, as well as goes for one passenger who's arm disappears in to its body.
- **Robotic animation** - Given that the animation sequence are four different animations coupled together, and not one long organic animation, the movement looks robotic and unnatural.



(a) Woman sitting down **(b) Woman standing up**

Figure 3.9 Comparison of sitting and standing posture

We believe that this animation improves the realism of the environment. However, given its flaws, the solution to the issue was to move the woman further away from the participant. By moving the animation it still contributes to realism but minor flaws will not be as easily spotted by the user. The placement of the woman was discussed beforehand, but it was ultimately decided to keep her next to the participant during the test to improve the chance of getting stronger reactions to her movement. An attempt at correcting her animation issues and arched back is also further addressed during phase two of the thesis.

3.5.4.3 Major Takeaway - Presence and Realism

As mentioned, almost all participants stated that the NPC environment was more realistic. In general all the extra sounds, people, movement and added interactable components led to the environment feeling more realistic. Even though the worse performance negatively affected the immersion, which in turn hurt the perceived presence, people still felt that the NPC environment was more realistic.

Another point of interest was that it was generally observed that around half of the participants felt no difference between the environments and those who did, felt a higher presence in the NPC environment. People who are unfamiliar with VR seem to have a low threshold for what enhances their presence in VR. Simple visual cues, such as the inside of an empty airplane, paired with the novelty of VR seems to give the person a high baseline for their presence. Meaning they already feel very present in a basic environment. Adding lots of realistic features adds to the realism and experience, but not necessarily the perceived presence of the participant. Therefore it was determined not to add any more visual features, apart from interactable components, but rather improving the performance and focusing on improving exiting animations.

3.5.4.4 Minor Takeaways

There were also a handful of minor takeaways that are important to take note of. They were categorized as minor because they were deemed as important for the project, easily fixable or simply something we are unable to affect.

Lamp not working - In the NPC environment, the lamp above the participants head had stopped working the way it was supposed to. It was a lot more difficult to turn on compared to the one in the robot environment.

Wing poorly rendered - The wing in the NPC environment appeared to have a rendering issue that gave it an unusual grainy appearance. Opposed to the wing in the robot environment which appeared as intended.

Window blinder - As mentioned, everyone managed to pull down the window blinder, although two people needed hints. No one had negative remarks regarding this feature and those who had things to say were gladly surprised about the feature. This further reinforces that interactable features enhances the overall experience of the virtual environment.

Participant behavior - In general participants were better behaved in the NPC environment compared to the robot environment. In the NPC environment participants were reluctant to disturb the NPC whereas in the robot environment, several participants grabbed or poked the robots. This hints at the participants thinking of the NPC more as actual humans compared to the robots.

Effect of gaming - Although the participants were not asked any questions regarding their gaming habits, given that they were all people who we associate with, we have a general understanding of their gaming habits. People with more gaming experienced seemed more eager to explore the environment, meaning they spent more time looking around and playing with the interactable components compared to those with no gaming experience. Out of the three people who stated experiencing feelings of uncanny valley, the two of them who after a couple of seconds stated the effect fading were people who frequently game. Whereas the one who felt an unnerving feeling from the NPCs throughout the test was a person who never played games. This suggests that people who regularly game are more susceptible to the experience, and more specifically the NPC environment.

Novelty of VR - During the tests, participants spent a lot of their energy and focus on grasping the fact that they were in a VR environment. The authors have worked a lot with VR during this project and in prior projects, therefore it was important to get the perspective of people less versed with VR. It was observed that the general person has rarely used VR and is still in awe of the novelty that is a virtual environment. This in turn distracts from the actual test.

There is not much to be done about the novelty itself. However, it is discussed further in phase 2 whether, before the final test, some sort of warm up session is needed to get the participants comfortable in a VR environment.

4 Phase Two

This chapter presents the adjustments made to the environment after the first user test. It also discusses the plans for the next user tests and how they were conducted. Lastly the results of the user tests are analyzed and discussed.

4.1 New Implementations

Based of the results and analysis from the user test conducted in phase one, several important insights were gathered regarding the current environment. New implementations and changes were made to the animations, sound and performance based of these results.

4.1.1 Animations and Sound

The first modification to be modified was the animation of the NPC located next to the user. The character stands up, turns left and takes one step forward and then turns right and walks away. The standing animations was replaced with an alternative one which appeared more natural and so she stands up for a shorter period of time resulting in an overall more smooth animations. Additionally, the character was moved to a seat positioned in front of the user instead, which helped reduce the visibility of minor animation inaccuracies while still maintaining the movement necessary to create a realistic scenario.

Other minor adjustments to the animations and environment were also implemented. The floor mesh was lowered, creating the impression that the floor was positioned further down. This was in response to a comment during the test from one participant, noting that the safety card appeared to be positioned too high above the floor. Additionally the signs with *no-smoking* and *fasten seatbelt*, had a transparent background making it look like they were floating in the air. Therefore another background was added to avoid resulting in a more realistic scenery.

Furthermore the talking character, was switched to a different character, along with its animation. This modification addressed feedback indicating that the character's arm at times looked like it was inside the stomach during the animation which looked unnatural. Also more characters were added, including both walking and standing individuals. These were distributed more evenly across the environment and time to avoid overcrowding and to create a smoother, more natural flow of activity.

We decided to keep the sound as it was from phase one without any modifications. Even though a majority of the participants did not recall noticing any difference in the sound, there was no negative remarks regarding it. Also considering the majority of participants perceived the NPC environment as more realistic, the potential contribution of audio to this perception cannot be ruled out. Well-designed audio does not necessarily need to be noticed to be effective, however, it can still play an important role in affecting the overall impression of an environment.

4.1.2 Performance

As mentioned in the design part, many different light sources and rendering methods were experimented with during the development phase, in order to reach natural lighting. However after conducting the first user test, as mentioned in subsubsection 3.5.4.1 some responses were regarding lag. To resolve this issue we experimented with older versions of the project which did not experience the same amount of lag. The lag issue was finally tackled by doing these two major things.

4.1.2.1 Two Airplanes

After excessive troubleshooting, and comparing with older versions of the project, it was discovered that, in the version of the project which the user tests were conducted, the airplane is rendered twice in the exact same spot. The airplane prefab and some of its corresponding prefabs had been copied, meaning there were duplicates of a lot of components that were not supposed to be duplicated. This explained some of the issues that test participants had noted during the user tests. The wing looking unnatural and grainy was due to the fact that two identical wings were placed in the exact same place, and the game engine therefore displayed both of them at the same time. This phenomenon is called Z-fighting, and is a common issue in 3D-graphics [20].

This also explained why the lamps were difficult to turn on. When test participants failed to turn on the lamp, even though they pressed where they should have, the game did not register that the lamp had been pressed due to the other airplane cabin being in the way of the button to turn on the lamp.

Lastly, the duplicate airplane also explained some of the performance issues. Z-fighting does generally not result in major performance issues. In this case, it was noted that some of the lag was resolved when the second airplane was removed. However, there were still notable performance issues.

4.1.2.2 Changing Scene

During the user tests in phase one the users were seated on the right side of the plane. When the lag issue was not managed to be resolved it was decided to move the user to the left side of the plane and redo the necessary parts. The reason this worked to improve performance was that each side of the plane is a different unity scene. It was most likely some animation and lightning that had caused a majority of the performance issues during the user tests in phase one. However, the exact culprit was never discovered. Therefore, when the user test for phase two was redone on the left side of the plane, lighting and animation was set up more carefully, which helped improve performance.

4.1.3 Interactable Objects

Due to the positive feedback of the window blinder from the first user test it was decided to add another interactable object to the environment. On the 'Should Have' section on the MoSCoW prioritization from Phase one there were two interactable elements, a seatbelt and a foldable tray table, see Figure 3.5. These options along with a foldable armrest, which was also mentioned by two of the test participants, were discussed for implementation.

The armrest was quickly disregarded due to the fact that the chair models in the environment already had armrests, but they were non-modular. Meaning that if foldable armrests were to be implemented, the original model of the chair would have to be remade.

4.1.3.1 Seatbelt or Tray table

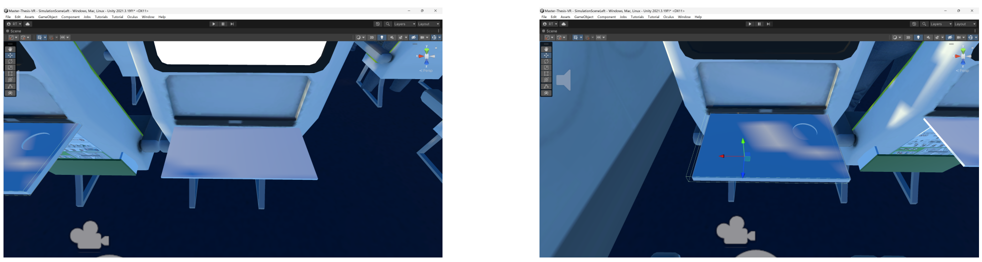
The prioritization regarding the implementation of the seatbelt and the tray table led to extensive discussion. They were estimated to be comparable in terms of effort of implementation. Rather the discussion was focused on the effect the two options would have on the users presence.

The main supporting argument for the seatbelt was that the seatbelt is something that everyone has to interact with at least twice per flight. Once when they are seated and once when they leave their seat. Along with this the seatbelt is also given a lot of attention during the flight. Both the safety demonstration and the overhead seatbelt sign reference the seatbelt, and both of them are strongly affiliated with the typical flight experience. All of this would presumably enhance the users presence, at the very least not lower it. The main drawback of the seatbelt has to do with how VR works and how this specific environment is setup. Generally in VR there are two common alternatives for how the environment is set up. Either only hands or full body experience. Meaning that in the VR environment the user either sees a full virtual body, virtual legs, torso, arms and hands which replicate their real body parts.

Alternatively, the program only displays the hands or controllers in the VR environment and nothing else of the users body. Both alternatives come with advantages and disadvantages, which are beyond the scope of this project. This project is set up to only display the users hands with no representation of the users body in the environment. This is relevant to the discussion of whether or not to implement the seatbelt. Because of how the environment is set up, if the user were to have a seatbelt that they could fasten themselves with, the seatbelt would visibly secure nothing due to the users body not being virtually represented. This could possibly lead to the users focus shifting to the fact that they do not have a representation of their body, which in turn could negatively affect their presence. It was also predicted that it would probably look unnatural that the seatbelt was attached to the chair and nothing else.

The advantages of the foldable tray tables were similar to those of the seatbelt, but not to the same extent. The tray table is present on many flights, but not all of them. When they are present on flights people do use them, but it is not mandatory as with the seatbelt. The tray tables are also explicably given focus when food or snacks is served on the plane. The main drawback was, which has already been mentioned, that it is not as important of an object as the seatbelt. There were also concerns that the tray table could possibly block or directly distract from the users experience with the IFE. Concerns were also raised to whether the physical body of the back of the chair would disturb the movement of the table, with the table and the chair possibly colliding in an unrealistic way. Similar issues arose while implementing the window blinder.

The conclusion was to try and implement the tray table, see Figure 4.1. Mainly because interacting with a seatbelt without a visual torso was deemed probable to negatively affect the users presence, and the alternative of implementing a virtual body to match the seatbelt was determined too time consuming.



(a) Table without a cupholder. **(b) Table with cupholder.**

Figure 4.1 Comparison of two different versions of the tray tables.

4.2 User Test

This section discusses the planning and execution of the user test of phase two.

4.2.1 Measuring Presence

For the previous smaller user test it was sufficient to ask a few general questions regarding the participants perceived presence to get a general understanding of their presence, to improve the development of the project. During this user test it was important to get a more accurate representation of the users presence. Otherwise, the final results would not yield valuable information in regards to the first research question which states *"In what way does the level of presence affect how a test person performs task in VR?"*. Therefore, after the user tests the participants had to answer the IPQ to ensure one environment had higher presence.

4.2.2 Test Environments

Similarly to the first user test, the participant experienced two different environments. This was done so that there could be an analysis comparing two environments with differing levels of presence.

The robot environment was similar to the robot environment during the first user test. Two aspects of the tests were adjusted. A touch screen was added. This was done so that the participants could interact with the IFE in both environments to be able to perform similar tasks in both of them. The background noise was removed to potentially reduce the perceived presence. This was due to the risk of not achieving a substantial difference of presence in the two environments, which would have ruined the test.

The NPC environment was adjusted according to the changes described in section 4.1.

4.2.3 Choice of Touchscreen

An important decision had to be made regarding what screen to use during the test. During one of the prior versions of the project, a 17" touchscreen was used, and during the other, an 11" android tablet was used. The main argument for using the bigger screen is the advantage of the text being bigger. Reading inside the VR environment is more difficult than in real life. The text is blurry and flickers a bit. An easy work around is to make the text bigger. A bigger screen results in bigger text, meaning it is easier to read on the bigger screen, and therefore it is more effortless to use the IFE system.

However, one of the most important aspects of this user test and this master thesis is achieving good presence. In the virtual environment the screen is attached to the back of a chair, meaning that the size of the screen is limited to the size of the chair. A way to solve this would be to remodel the plane using bigger chairs, but that was not feasible given how far along the process we were and the time constraints of the project. Another way of using a bigger screen is simply to ignore the given constraint that the screen has to fit on the back of the chair and make the screens edges hang over the edges of the chair. There is also the possibility of making the entire airplane larger, which in turn makes the chair larger and gives more room for the big screen. To use the 17" screen in this environment we both had to make the screen hang of the edges of the chair and make the entire plane larger. This is fine when participants are mainly focused on testing the IDE itself. However, when focus is on determining presence, having the user seated in an abnormally large airplane and interacting with a screen that is wider than the chair it is attached to, this approach was ultimately deemed too unrealistic to be used during the user tests. Finally a feature had been added to the IFE so that the user can manually adjust the text size, which did not exist during the two previous theses. This meant that our current version using the 11" screen is more readable than the previous one.

4.2.4 Choice of Test Participants

Considering that this project was made in collaboration with a company, all code, models, hardware and equipment used were owned by them. This naturally made it easier performing the tests at their office. Consequently, participants were primarily recruited internally within the company, as involving external individuals would have introduced complications related to confidentiality and the protection of company-sensitive information. Due to the time limitations we choose to limit the number of test participants to 16 people to still ensure reliable data.

4.2.5 Set Up

The set up was deliberately placed using a height-adjustable desktop located next to a wall. This added presence since the airplane wall and window in VR aligned with the real world wall, ensuring the user could not reach their hand outside the aircraft. The tablet was placed on the table which had adjusted height, to match where the entertainment system would be placed on an actual flight. Furthermore the tablet was secured using duct tape so the placement would be secure even though user pointed at it with a lot of force, see set up Figure 4.2. The chair used for the test was a traditional office chair with high back support to imitate a airplane seat.

Figure 4.2 View of the test set up.

4.2.6 Pre-test Screening

As done in the first user test in phase one, a pre-test screening was conducted to gather information regarding the test participants. All previous questions were kept, however, two additional questions were introduced to gain further insight. One question addressed the participants previous experience with video games, as observations from the first test indicated that interaction with the environment varied significantly between experienced gamers and non-gamers. The second additional question asked whether participants had tried any of the previous versions of this project, as previous exposure could affect their behavior and thus the result. The test was conducted at the company's office weere some employees who participated in our test, worked and also tested the previous versions of this project. Even thought this was over one and three years ago it was still considered relevant information to collect. See all questions in Appendix A.3.1.

4.2.7 Conducting the User Test

Before conducting the test on actual participants, a table of the test-plan was created to establish a mutual strategy between the test leaders.

Prior to conducting the study, two distinct roles were pre-decided between test leaders in order to ensure a structured and consistent procedure. These roles designated

Table 4.1 Table of user test workflow

Phase	Activities	Materials needed	Time
Briefing	Fill in pre-screening Provide information	Introduction Script Informed consent Pre-test questions	4 min
Test 1	Introduce headset Perform task	Meta Quest Pro Task descriptions	5 min
Test 2	Switch Environment Perform task	Meta Quest Pro Task descriptions	5 min
Debriefing	User fills in PQI Post test interview	PQI form Post test questions	10 min

to the same person for every user test and were defined as follows:

- **Observer**
  The observer was responsible for managing the technical and documentation aspects of the session. This included operating the mobile phone used to record both the participant's interaction with the screen and the screen. Additionally, the observer managed the stopwatch to measure the time to complete a task and counted errors made during the tasks. The monitor also took notes throughout the session, including participants' comments, and responses from the post-test interview were documented.

- **Moderator**
  The moderator was responsible for guiding the participant through the session. This included reading instructions and informed consent information at the beginning of the test. The moderator helped the participants with the VR setup and ensured they understood the tasks. During the debriefing phase, the moderator read the post-test questions and added follow-up questions when necessary to clarify or deepen the responses. They were also in charge of screen recording the android tablet.

Since this test, just as the first one, included two different environments it was essential that half of the participants started with the NPC environments and the other half started with the robot environment. Similarly, another precaution that was made was that half of the participants stared with 'scenario 1' and the other half with 'scenario 2', the results can be seen in Table 4.2. The tasks and questions asked to perform can been seen in Table 4.3 and Table 4.4

Table 4.2 Numbers of participants starting with each scenario/setting

Start with:	Scenario 1	Scenario 2
NPC	4	4
Robots	4	4

4.2.8 Choice of Tasks and Questions

The purpose of the user performing the tasks is to be able to get measurable data to then compare results between the two environments. Tasks 1, 3 and 5 are all meant to measure speed and accuracy. The fourth task is meant to test their cognitive skills and their speed. The second one is simply meant to return them to the start screen, this is therefore the only task that is the same in both scenarios.

This way we can achieve an understanding of their overall usage of the IFE by measuring speed, accuracy and cognitive load with only five tasks.

Table 4.3 Tasks in Scenario 1

1	Find how many routes there are and to how many countries they fly
2	Go to flight info
3	Read about a tourist attraction in New York City
4	How many passengers would you estimate is the maximum capacity of this flight
5	Change the settings from imperial to metric measurements

Table 4.4 Tasks in Scenario 2

1	Find how many aircraft that are currently in the air
2	Go to flight info
3	Read about a tourist attraction in Tokyo, Japan
4	What row would you estimate that you are sitting on
5	Change the settings from 12 Hours clock instead of 24 Hours

4.2.9 Post-test Questionnaire

After conducting the test a questionnaire was filled in order to measure presence, which was done using the 14 question IPQ, see Appendix A.3.2. The participants filled it in by themselves using Google Forms. This was done in order to support the claim that the two environments differ in perceived presence.

There was two alternatives on how to perform this questionnaire; either answering the questions immediately after each environment, or experience both first and then afterwards answer the questionnaire joint for them. The downside of answering for both at the same time was that it could lead to a comparison rather than the actual opinion. The IPQ is specifically developed to measure presence not compare two environments.

However by answering in between the environments, the participants will be exposed to the questions before experiencing the second environment. Having seen the questions in advance could introduce bias, as participants might consciously or unconsciously focus on specific aspects they know will be evaluated.

Another argument for having the questionnaire after both environments was the practical consideration that the participant did not require to remove the headset in between each setting. When removing the headset, the anchors which positions the screen needs to be placed again out by the user. This can be a bit complicated and time consuming and was preferably only done once to not extend the test even further.

Additionally, since VR technology is unfamiliar or entirely new to a large precentage of participants, it was considered reasonable to let them experience both environments before responding. This approach allows them to answer with a more informed judgment. Since the questionnaire uses a scale from -3 to 3, the responses could be misleading if someone would answer the best possible score on e.g. the NPC version, and then cannot give a higher score to the other e.g., the robot version, even if they actually preferred the latter.

4.2.10 Post-test Interview

Lastly the final thing the participant would do to complete the user test is to answer a couple of interview questions. The purpose of the questionnaire was to measure the participants presence. The purpose of the tasks was to measure possible difference in user performance. The purpose of the interview question was to give the users a chance to speak freely of their experience, to give them a chance to voice their opinions and thoughts that otherwise would not be shared. See Table 4.5.

4.2.11 Pilot Test

When everything was planned and setup, a pilot test was performed. The project supervisor, was the test participant. We choose our supervisor to perform the pilot test because he had performed similar tests a couple of years back and, therefore, could

Table 4.5 Interview Questions

Q1	Did you experience any difference between the two environments?
Q2	Did using the IFE feel different between the two environment?
Q3	Did you prefer one of the environments? If so, why?
Q4	Did you feel more distracted in one of the environments?
Q5	Is there anything else you want to share?

hopefully recognize common errors with our test. After the pilot test was conducted, we had garnered a lot of useful insights from our supervisor and through our own conclusions. The following is a sample of the changes that were made following the pilot test.

4.2.11.1 Removing the Foldable Tray Table

Going into the pilot test we were aware that the foldable tray table was not functioning exactly as we had liked. A lot of time had been spent on figuring out the mechanics of attaching the table to the chair, correcting the hinge mechanism and making it grabbable. During the pilot test the table worked almost as intended, apart from one thing. If the user were to slowly push the table downwards with their hand, the table would follow and could then be pushed out of its original position. When the user lets go, the table would snap back to its original position and work as intended. The table was actively kept in the environment during the pilot test so the supervisor could voice his opinion regarding the table. It was ultimately deemed that this flawed table could possibly hurt the presence more than it would add. Therefore, due to it not working completely as intended, the table was removed from the scene.

4.2.11.2 Fewer Tasks

During the pilot test the participant performed eight tasks in both environments. Some of the tasks were removed due them being to similar to others, or not adding anything relevant to the test. One removed task was *"Navigate to 'Time zones'"*. This was an easy task, the participant would only have to move the menu bar slightly, and it did not test anything specific, meaning it would not add any meaningful information. During the pilot test the whole test took about 35 minutes and the real user test took about 25 minutes. A strong argument for not having a too long user test, was to maintain the participants attention and engagement throughout the entire session.

4.3 Result

4.3.1 Result of Pre-test Questionnaire

From the pre test questionnaire we got the information that 10 out of 16 participants were male and the remaining 6 were female. The mean age of the participants was 36 years (youngest 22 years, oldest 50 years). A majority, 10 participants of 16 had experienced VR once or a few times in their life, while the rest had used it more frequently, consequently no one were completely new to VR. Another interesting finding was that 7 of 16 participants had tried at least one of the previous versions of this project and 9 had not. All collected data and corresponding graphs from the questionnaire can be found in Appendix, see A.2.2.

4.3.2 Result from IPQ

The 14 IPQ questions is, as mentioned in subsection 2.1.3, divided in 4 different sub categories; spatial presence (SP), general presence (GP), involvement (INV) and experienced realism (REAL). Three questions, number 2, 11 and 13, were negatively phrased and their answers had to be reverted before the mean value could be calculated:

- **REAL1**: How real did the virtual world seem to you?
- **INV11**:I still paid attention to the real environment.
- **SP13**: I felt like I was just perceiving pictures.

After reverting the correct answers, the mean values for each questionnaire item were calculated separately for the NPC environment and the robot environment, see exact score for each item in Appendix A.3.2. Followed by calculating the mean value for each sub category, also for each environment, creating a comparison of presence across the different conditions.

Table 4.6 Mean values for NPC environment

GP:	1.4
SP:	1.7
INV:	0.2
REAL:	-0.2

By comparing Table 4.6 and Table 4.7 it shows that the NPC environment provides higher scores across all four categories resulting in higher presence. The largest difference was observed in General Presence, where the scores differed by 1.0. In

Table 4.7 Mean values for robot environment

GP	0.4
SP	1.1
INV	0
REAL	-1.1

contrast, the Involvement category shows the smallest difference between the two environments, only 0.2. The Spatial presence differed 0.6 and the Realism category had a difference of 0.9. Notable, was that both environments received their lowest scores in the Realism category, with values of -0.2 and -1.1, respectively, indicating that participants perceived both environments as relatively less realistic compared to the real world. The overall mean score for the NPC environment was then calculated to 0.78 and for the Robot was 0.1. In the report [25], 243 IPQ studies and results have been analyzed to present ranking class depending on the score. Five ranking classes names are proposed depending on the percentile threshold of the corresponding score, see Table 4.8. Based on presence scores the NPC environment can be classified as *High*, whereas the robot environment falls within the *Low* category. This indicates in stronger sense of presence in the NPC environment compared to the robot environment.

Table 4.8 Ranking class depending in the score

Ranking Class	Percentile	Score Range
Exceptional	Above the 95th	1.3, 3
Very High	Above the 90th	1.07, 1.3
High	Above the 75th	0.73, 1.07
Moderate	Above the 50th	0.28, 0.73
Low	Below the 50th	-3, 0.28

4.3.3 Analyzing the Interviews

A thematic analysis was performed in order to analyze the answers of the interviews. It was done by going through each questions and its answer from every participant to spot any potential similarities and patterns. Given that these are open-ended questions, answers can differ by a lot. Therefore, there will not always be a natural way to categorize them. This analysis of the questions is an attempt to give a fair and unbiased representation of the data. This means answers to one question does not necessarily represent the feelings of the participants as a whole. This is simply their answers to a question in isolation. Another important thing to note is that the total value of all the combined categorized answers does go above the total number of participants in some cases. This is again due to the nature of trying to categorize answers to open-ended questions. For example one such instance, in Q1, participant with id #2 both mentioned that they think the NPC environment is more realistic, but also think the the NPC environment is distracting, including them in both categories. However, in Q3 for example, such an answer could not exist, given the nature of the question. Unique and specific answers are further discussed in subsection 4.3.7.

Q1: Did you experience any difference between the two environments?

All participants noticed the difference between the other passengers in the aircraft, that one had robots and the other human characters. Otherwise what was noticed differed a lot, with contributions from the answers from question one along with rest of the interview, a thematic analysis could be conducted. These results are presented in Figure 4.5.

Q2: Did using the IFE feel different between the two environments?

All 16 of the user test participants stated that using the IFE did not feel different between the two scenarios.

Q3: Did you prefer one of the environments? If so, why?

The NPC environment was preferable to 12 out of the 16 test participants, 2 preferred the robot environment and 2 did not have a preference. See Figure 4.3.

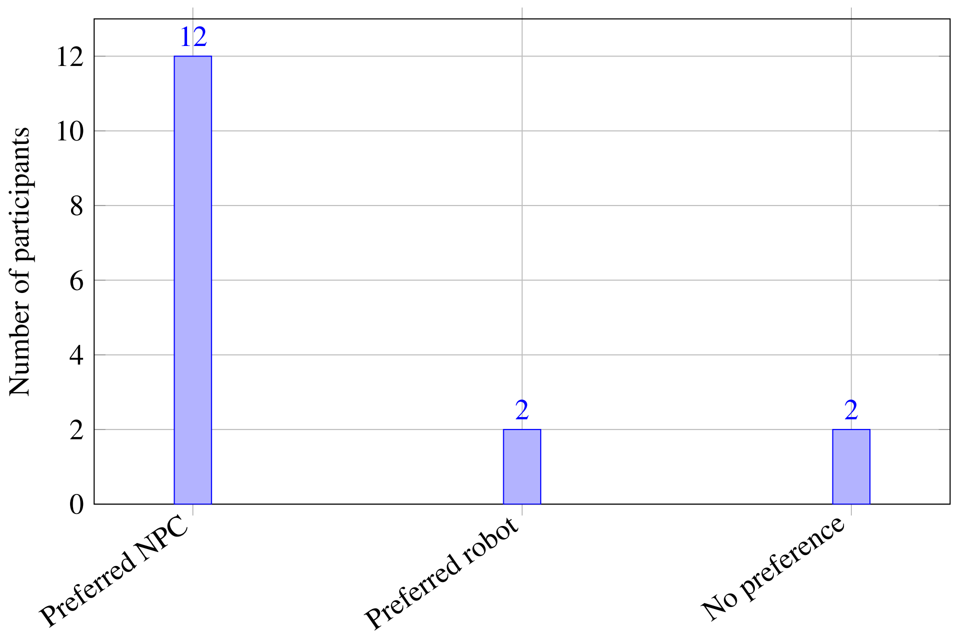


Figure 4.3 Summary of participant answers to open-ended question Q3.

1 out of the 12 participant who preferred the NPC environment did not complete this part of the interview due to technical issues that put us over time. However, it can

be concluded from their answer to question #1 that they did indeed prefer the NPC environment.

> *"I thought that the one with people was funnier and more realistic. It was fun that they were doing things, and fun that someone came out of the bathroom. I don't know if it was the sound or something,* ***but I preferred that one.****"*

– Participant #10

Amongst the participants who preferred the NPC environment common quotes where such that mentioned it feeling more realistic and less scary in the NPC environment.

Q4: Did you feel more distracted in one of the environments?

9 out of 16 participants did not feel any difference in their level of distraction in any of the environments, five participants stated feeling more distracted in the NPC environment, one participant stated feeling more distracted in the robot environment and lastly, one participant did not answer, see Figure 4.4. Out of the five participants who mentioned feeling more distracted in the NPC environment, two participants explicitly stated feeling distracted in the beginning of the scene, but not while performing the tasks and the other three participants mentioned in someway that the distraction was something positive, that it added to the realness of the scene. The person who was distracted by the robots mentioned that they were scared of them, and that is why they were distracted.

> *"I thought the sound added to the realism and did not bother me."*

– Participant #8

> *"No [I did not feel more distracted], I forgot the others quickly because they didn't interact with me anyways."*

– Participant #14

Q5: Is there anything else you want to share?

For this question, the answers are so varied that it is difficult to properly display the results. Therefore, this data will only be used as supporting material for other claims, rather than as the basis for any independent claims.

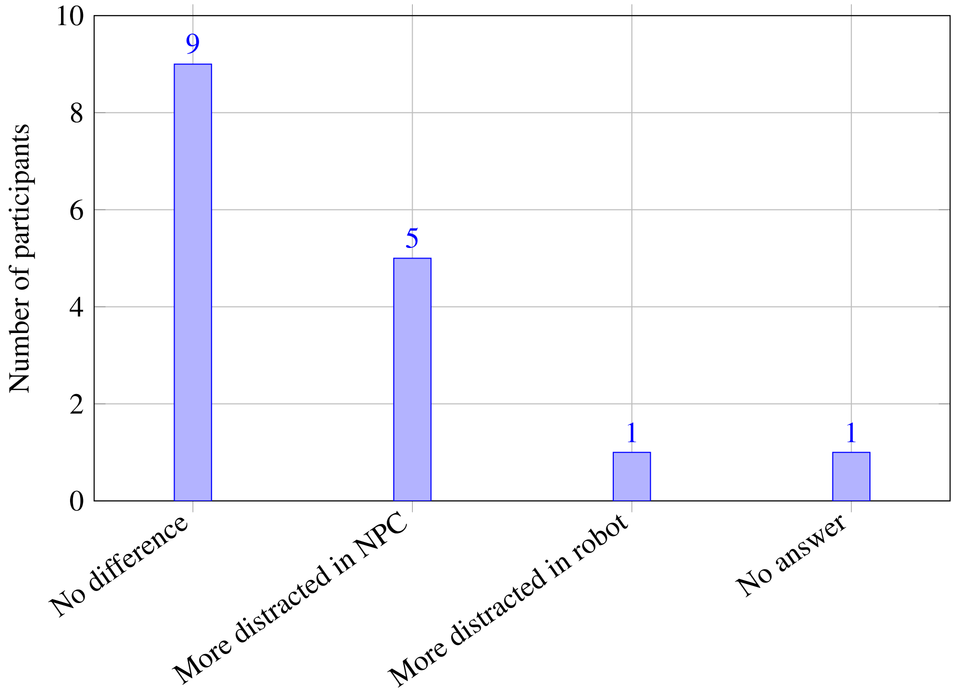


Figure 4.4 Summary of participant answers to open-ended question Q4.

4.3.3.1 Overall Insights

Overall many insights were gathered regarding the different environments throughout the interviews. 8 out of 16 participants mentioning something in regards to the NPC environment feeling realistic. 5 out of 16 mentioned that either the NPC environment was more fun and engaging or that the robot environment was more boring. 6 out of 16 participants mentioned lighting or sound as something positive while 2 talked about liking interactable objects. Four participants mentioned being distracted in the NPC environment. Five participants mentioned being scared in the robot environment. See Figure 4.5.

4.3.4 Analyzing User Groups

There were 6 participants out of 18, exactly 1/3 of the participants, who play video games at least once a week. These are the results of the IPQ when analyzing the answers of those participants.

4.3.5 Task Completion Time

During all the task performed, a timer was started when the task had been read and understood, and stopped when it was finished. As mentioned, this was done to evalu-

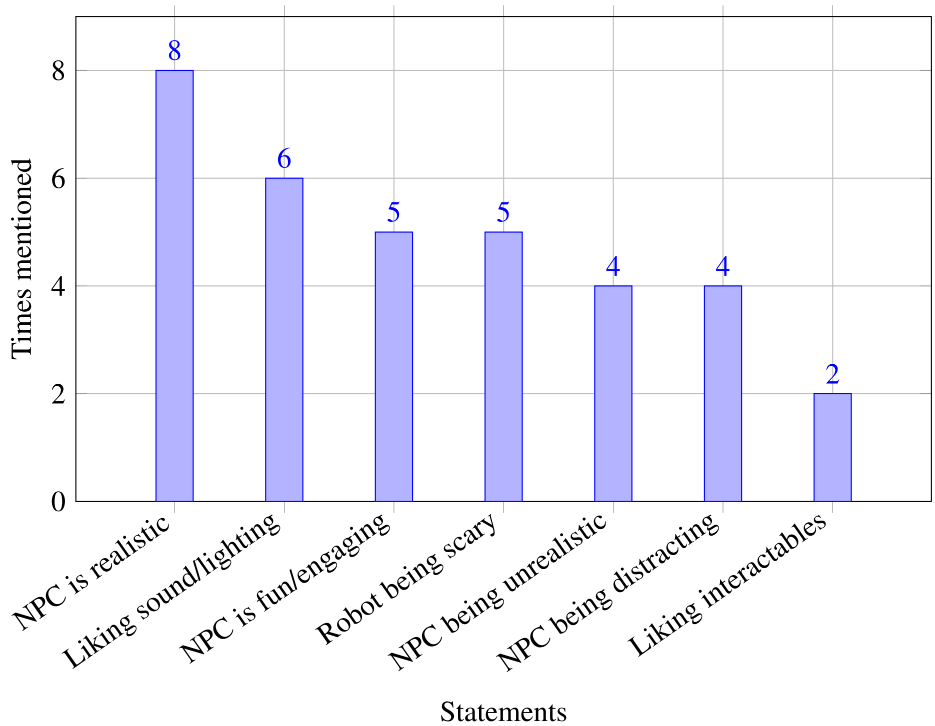


Figure 4.5 Summary of mentioned statements throughout the interviews.

Table 4.9 Mean values for NPC environment for participants who play video games at least once a week.

GP	1.7
SP	2.2
INV	1
REAL	-0.2

Table 4.10 Mean values for robot environment for participants who play video games at least once a week.

GP	0.8
SP	1.5
INV	0.6
REAL	-0.6

ate whether, and how, user performance differed between the two environments with varying levels of presence.

To analyze task performance, the median time was used instead of the mean in order to exclude extreme values and gain a more reliable results. Three main calculations were performed:

1. The median completion time for each task in Scenario 1 and Scenario 2.
2. The median completion time for each task in the NPC environment and the robot environment.
3. The mean completion time for each task based on the order the participants did the scenarios, i.e the scenario they started with versus the one they finished with.

The result of the first calculation can be seen in Table 4.11. Overall no clear pattern can be seen regarding which scenario resulted in longer task completion times. Two tasks in scene 1 required more time, especially task 4, while two tasks in Scenario 2 took longer to complete. Task 2 showed approximately equal completion times in both scenarios.

Table 4.11 Median values of task completion time based on scenario, (sec)

Task	Scene 1	Scene 2
1	13,5	18,5
2	3,5	4
3	35	29,5
4	20	10
5	10	12

As seen in Table 4.12, the time to complete a task in each environment is presented. Similar to the previous results, no clear pattern can be identified. Tasks 1 and 4 took longer to complete in the robot environment, while Task 3 required more time in the NPC environment. However, these differences are relatively small and can be considered marginal. Tasks 2 and 5 showed approximately the same completion times across both environments.

Table 4.12 Median values of task completion time based on Environments, (sec)

Task	Robot	NPC
1	15,5	13
2	5	4
3	29,5	33,5
4	19	15,5
5	11	11

The last comparison, to see difference between which scene the user started versus finished with is presented in Table 4.13. Unlike the two previous tables, here a clear trend can be noticed. Every single task is completed in a shorter amount of time in the scene the user finished with. Task 2, 4 and 5 is marginally shorter, but still shorter. Indicating that this is the most important factor for time consumption to complete a task.

Table 4.13 Mean values of task completion time based on starting scene, (sec)

Task	Scene started with	Scene finished with
1	19,9	12,4
2	7	5
3	41,5	29,8
4	19,8	16,7
5	13,2	12,1

4.3.6 Errors

There were some faults with the user tests. Most of them were minor ones, such as the virtual screen not always matching well with the real screen. This only slowed down the participants a bit in the beginning. However, they were able to rather quickly adjust where they would place their finger due to the screen showing feedback of where the person touched the screen. Meaning the participant could see and feel the difference between their touch and were the screen displayed that the touch happened. Another minor issue was that the moderator forgot to turn on the screen recording for a couple of the test participants. This occurred some times in the beginning due to inexperience with performing the user tests, and did also occur when we experienced major issues with the user test, which distracted the observer. Due to the fact that no participant reported a perceived difference interacting with the screen in the different environments and that there were no observed difference of time and error rate between the two environments, apart from that a majority of participants were quicker during whichever scenario they performed second, this issue did not affect any meaningful data.

A major issue that unfortunately happened quite often was the Link crash. This had happened earlier during development and the issue is further discussed in section 5.3. In short, the Link crash can happen at every stage of the test and forces us to restart the headset, restart the computer and redo the setup. This process takes about five to ten minutes. This issue was believed to be resolved about two weeks prior to the user test, unfortunately the issue resurfaced on the day of the first scheduled test. Luckily, the crash only occurred at the beginning, in between, and at the very end of the user tests, meaning no participant had to redo any part of their test due to the crash. The crash would mainly prolong the total test time and briefly confuse the participant.

4.3.7 Takeaways

The major takeaways from the second user test is presented and discussed under this section.

4.3.7.1 Realism Equals Higher Presence

The result of IPQ scores for both environments is presented in subsection 2.1.3. According to the result, the NPC environment had higher level of presence across all sub categories compared to the robot environment. Furthermore based on the aggregated scores, it was also classified as High while the Robot environment fell within the Low category. This suggests that the environment with more real-world features, was experienced with higher presence.

This conclusion is additionally supported by the thematic analysis from the interviews, where the majority of participants stated that the NPC environment felt more realistic. In particular, participants highlighted that the presence of moving human characters contributed to a greater sense of realism compared to the still robots.

These findings indicates that by adding more features that resemble real-world characteristics in a VR scenario, can provide more realism and, consequently higher presence.

4.3.7.2 Presence does not affect performance

The presence does not affect performance, at least not in this test environment. Every single test participant said that they felt no difference in their performance on the IFE between the two test environments. This is also confirmed by the time taken to complete each task not measuring a sufficient discrepancy between the environments. Some participants mentioned being more distracted in the NPC environment, but that still did not affect their result since they were all able to focus on the assignment at hand when the moderator started reading them their tasks.

In this particular test environment, all but one task had the user focus solely on the screen in front of them, meaning they were not actively interacting with the virtual world surrounding them. This could suggest that this was why the participants were able to focus and perform their tasks as usual, even with an interesting and distracting world surrounding them. This is further discussed in section 5.5.

4.3.7.3 Presence Affects Users Experience

Even though the time to complete a task as well as users performance appears not be affected by the level of presence, other aspects do. One question in the post test interview was regarding which environment the participants preferred. From the analysis of the interviews, twelve preferred the NPC environment and only two preferred the robot environment, and as mentioned a common opinion was that the robots felt scary. Additionally, four participants reported that the NPC environment was more enjoyable and fun, while another four stated they felt more curious due to the presence of moving human characters with varying appearances.

This suggests that including elements with higher realism and variability may enhance user preference and overall experience, even if it does not directly impact task performance.

4.3.7.4 Gamers are More Susceptible to Non-Realistic Events

From the analysis of the IPQ results there is a clear trend that those who play more video games had a higher perceived presence in both environments, especially in the Robot environment. This was something we suspected by the end of the user tests in Phase one and could confirm with the user tests in Phase two.

4.3.7.5 Natural in Higher Presence

The participants generally behaved more natural in the NPC environment compared to the Robot environment. This was displayed in two particular ways. Firstly, they explored the plane more. Noticing other humans, listening to what they had to say, and they were generally more eager to have a look around in the plane. Compared to the robot environment where it was more common to sit in your chair, have a look at the robot next to you, then focus on the IFE screen in front of you. Another unnatural behavior in the Robot environment was that many of the participants poked or touched the Robot sitting next to them. This was already noted during the user tests in Phase one and was more properly tracked during the user test in Phase two. To poke or touch your unknown neighbor is something extremly rare. None of the participants did this in the NPC environment.

5 Discussion

This chapter discusses the results of this paper, and the implications of those results. This is done in relation to the papers research questions. It also covers the projects technical issues, limitations, and lastly the potential future work.

5.1 Research Questions

5.1.1 In what way does the level of presence affect how a test person performs a task in VR?

Through the results of this paper it can be concluded that presence does not affect performance when conducting user tests in VR, see subsection 4.3.7. However, that conclusion was made with the specific circumstances of this paper and its user tests. The user tests were designed so that the participant spent a lot of time and focus on the screen in front of them, not interacting much with the environment surrounding them. What would the result have looked like if that was not the case?

Consider the case where a virtual operating room is built and a medical student is presented with the task of practicing for a surgery in VR. The participant would have to interact with a lot more in the room, such as picking up tools around them from trays and medical assistants. The distracting elements are also more important here such as the sound and the lighting. The participant has to recognize what sounds are important right now, and which ones to filter out. Then also make the participant practice another surgery, but from a virtual operating room which does not give them the same perceived presence, similar to the setup of this master thesis. Will the results be the same as in this master thesis, that higher perceived presence does not affect results, or will it in fact have a bigger effect on result when the test forces the participant to be more involved in the environment? Such a scenario would make up for the limitations of this master thesis and give more insights to whether or not presence actually matters in VR testing scenarios.

The conclusion that presence does not affect performance in subsection 4.3.7 is based on the premise that performance correlates with success rate and speed of completing the tasks. If performance were to be measured more broadly and not only measure success rate and speed, the conclusion would most likely differ. It would be reasonable for the performance of a test participant to account for the 'how' in "*In* *what way does the level of presence affect* ***how*** *a test person performs a task in VR?*, not only success rate. This research question also specifically asks "*In what way...*", which is also asking for more detail of the users general performance as opposed to simply measuring speed and accuracy.

Measuring how a test person performs a task in VR accounts for their general behavior in the environment, which is indeed concluded to differ depending on perceived presence. When perceiving higher presence, participants engaged more with their environment. They were generally more curious with the environment surrounding them. This was noted through the participants small remarks regarding the NPCs, looking around the plane, and their answers to the interview questions. This is all parts of their behavior that changed during higher presence. Therefore how they performed a task, when including how they behaved and how they were thinking, was in fact affected by their level of presence.

Additionally, some participants stated that they wanted to look around more and that they felt slightly more distracted when human characters were moving around them, but that they were distracted in a positive way. This indicates that in this environment, some participants had to sharpen their focus and block out background distractions. Similar to how one would have behaved in a real-life setting when activities are happening all around. This could affect a test persons performance in cognitive load when having to actively focus even more or filter out irrelevant activity when completing a task. In more demanding tests or tests which takes up a longer period of time, this effect could negatively impact performance and, consequently, the overall test results.

Therefore, it can be stated that, when performance of a task is calculated solely by speed and accuracy, not general behavior, this paper can conclude that a higher level of presence does not have to affect how a user performs specific tasks. However, it is important to note that this does not necessarily indicate that higher presence is not able to affect how a test person performs a task in VR.

5.1.2 When researching in VR, how much effort should be put into making the environment realistic?

As previously discussed, there are several strategies to enhance realism in VR. This project implemented moving humans characters, a few extra sound effects as well as natural light from the window and finally more interactable objects, while static robots were removed. The result showed that a majority thought that the new environment was more realistic.

However, it is difficult to determine which specific component influenced this outcome the most. In research there is often time and resource limitations resulting in having to prioritize what to implement. Meaning it would be preferable to attempt to make a clear prioritization of when and what feature to implement.

Even though certain attributes were not explicitly identified by participants it may still have contributed to the overall experience. For example, one participant stated that something was different with the colors between the two environments but was unable to specify exactly what had changed. This is likely due to the different lightening which can subtly alter color, and therefore subconsciously affects the surroundings. A similar effect can happen in a real-world scenario: people often tend to not notice dim lighting until a lamp is turned on, making the change so rapid it is easy to notice it. The same effect can be stated about sound, it is mainly noticed once it does not fit in the context or does not exist at all. Some participants commented that they enjoyed the sound whilst other did not notice any difference, even though one scenario was completely silent. Most participant did not notice or interact with the interactable objects, but the ones who did stated that they enjoyed them. The conclusion for this project is that perhaps not all smaller components are necessarily needed but they all can contribute to an overall more realistic environment. Consequentially meaning that, in a project with very limited resources, these components realistically are not the most crucial to implement.

In contrast, the difference which all participants noticed, was that one environment had moving human characters and the other had static robots. This was also pointed out as the most enjoyable difference as well as the biggest factor for increasing the engagement of participants' surroundings. Therefore, when discussing *'...how much effort should be put into making the environment realistic...'*, it has to be mentioned that the visual aspects, including moving objects, will most likely have the biggest impact. From the observation during the user test, all participants looked around at least once, but not everyone noticed the smaller changes. This will also, as previously mentioned, depend on the context of the environment and the goal.

However, in a scenario where basic funding is secured, but a larger resource grant could be argued for, whether it be in the public sector, academic world or general research and development, the question *'...how much effort...'* changes. The focus shifts from meeting the bare minimum threshold of valuable realism, to determine the cost–benefit of investing additional effort to achieve an even higher level of realism in VR.

Consider a city planner who can in greater detail visualize their city district using VR. In this case, realistic lightning, along with visual elements, suddenly becomes crucial in order to identify major flaws before construction begins. For instance drivers are blinded by the lights at a stop sign due to the orientation of two buildings and the poor placements of their windows. Or perhaps the city is partially run on renewable energy, but their solar panels are blocked from the sun more than anticipated. It becomes considerably less expensive to tackle and notice these issues in a costly, but realistic VR environment, than to handle them midway through construction.

Or if a museum were to have a VR environment of a historical setting to visualize for visitors, their main goal is to create engagement as well as educate correct facts by making it realistic. As stated, not all users notice every component but some do, therefore to make as many as possible enjoy it, incorporating sound, light and interactable objects increase the possibility to do so as well as creating realism.

Therefore arguing that increased resources to be able to put more effort in enhancing realism can yield valuable results. This will of course vary depending on the specific environment of what is important to include; lights, sounds, level of details or movement. Adding a feasible light from a realistic angle or adding appropriate background sound to the specific environment, can, in many cases, be sufficient to meet the requirements of a project and justify the effort invested in increasing realism.

Finally, to determine how much effort should be put into making a VR environment realistic, it is important to determine some key aspects. As a baseline, a low effort environment, with focus on the visual elements, goes a long way. This is especially true when the research does not require much interaction with the virtual world, or where the tasks at hand are not VR specific tasks, such as the IFE focused tasks of this paper. If the research seems to require a more advanced environment, it is important to beforehand determine why that is and what specific elements needs to be improved and lastly, what the desired outcome of those extra efforts would be.

5.2 Limitations

This section discusses some limitations of this project and why certain choices were made from it. The limitations were out of our control but have not affected the greater outcome of the study.

5.2.1 System Used in User Test

As this project was made in collaboration with a company working with developing IFE systems, it was natural to preform the test using their system. However, the system used was a prototype, meaning that not all features was fully implemented limiting tasks to perform in it. Additionally the content available was very limited, it primarily included information regarding flight routes, destinations and settings. An alternative approach could have been to use a another broader program with more features and variations to gain even deeper insight.

5.2.2 Participant Selection

Another limitation is that all participants were employees of, or otherwise associated with, the company. Therefore many participants had worked with or at least tested the system before which may have affected the outcome compared to using people who were completely unfamiliar with it. The reason for this was largely due to company policies, which made it difficult to invite external participants to the office. This in combination with the nature of the test's stationary set-up, making it difficult conducting the test outside of the office in an alternative location.

5.2.3 Hardware

The hardware that was provided by the collaborating company were for the most part sufficient for testing and developing in VR. However, the limits of the hardware were reached during development, which in turn limited the development of the virtual environment. Given that it was important to distance the NPC environment from the Robot environment, in terms of realness and perceived presence, demanding features such as realistic lightning and shadows were tested. *'Baked', 'Mixed' and 'Real time'* lightning were all tested during development of the NPC environment. This is further discussed in subsection 3.3.3, but in short, Real time looks better, but is more demanding for the computer, compared with Baked, which is less demanding for the computer, but does not generally look as good. There were instances were the desired lighting setting could not be used without causing the system to lag, due to the limits of the hardware. Although the same issues arose with other elements, such as animation, the main development issues regarding hardware limitations were those concerning lightning.

Another limitation regarding the hardware was during conducting the user test, a VR headset was connected to the computer, simultaneously as multiple heavy programs were running. The computer used was sufficient to handle this for the most part, but froze at times. During a few user tests the program crashed right before the entering the environment, forcing to do a restart, and in one case towards the end with two tasks remaining. These crashes resulted in prolonging some tests and may have caused inconvenience for participants, issues when they had to redo the entire set-up procedure. In one instance the session had to be terminated abruptly without completing the last tasks. Using more powerful hardware could likely eliminated these problems.

5.3 Technical difficulties

This section highlights some specific technical difficulties which mainly affected the time consumption of the development process.

5.3.1 Meta Horizon Link

During this project it has been crucial using Meta Link, mainly for the purpose of the final test but also for developing in the environment. It was necessary for the test because while using the VR environment, the participants had to interact with the IFE system at the same time on an actual tablet. This is only possible to do while using Link.

However, since the very start of this project, we have encountered major complications using Meta Link. During the first week of development the system worked exactly as expected. However, after an update of the Meta Horizon Link app, the system stopped working properly. Each time Meta Horizon Link started, regardless of whether our project was running or not, it crashed after approximately one to five minutes. Throughout the major part of the development process, many attempts were made to identify and resolve the issue.

Extensive troubleshooting was performed online through reviewing reports from other users experiencing similar problems. Several proposed solutions were tested, including updating the drivers for both the Dell and NVIDIA graphics cards but without any result.

After this, a support ticket was submitted to Meta Horizon Link customer support, explaining the current situation in depth along with providing technical information about the computer being used. The support team suggested several additional troubleshooting steps, including setting high-performance priority for specific Meta-related files, disabling GPU scheduling, and removing older files from the Oculus package. However after implementing all the suggested changes, when testing Meta Horizon Link again it still did not work the way it should.

Initially, we suspected that the headset or cable might be damaged and causing the technical issues. To test this hypothesis, another Meta Quest headset and Link cable were borrowed from Lund University. However, when using the alternative hardware with the same computer, Meta Horizon Link continued to crash in the same manner as before.

We then decided to transfer our project to an alternative computer and tested using the original headset and Link cable. It was tested both on a computer with another version of Meta Horizon Link as well as a computer with the exact same version of the application. This constellation worked without anything crashing, which clearly provides the conclusion that problems lays on the initial computer. After establishing this, a final attempt was made to resolve the issue by completely resetting the original computer and reinstalling the necessary software before testing Meta Horizon Link again. This ended up solving the issue. We suspect that the underlying issue could be a misplaced old unused Meta file or folder that somehow interrupted the Link software.

There were still issues with Link such as Link software not loading, it would still crash, although not as dramatically, and sometimes the headset would fail to connect to the computer. However, the overlying issue of Link always crashing was solved for a couple of weeks after the factory reset of the computer. Unfortunately, the issue resurfaced on the day of the first user test, possibly due to a software update. Luckily, the issue only had a minor affect on the final result.

5.3.2 Rebuilding the Environment

Due to the issues with Meta Horizion Link on the computer at the technology consultancy, some of the development was done in the VR-lab at Lund University. To be able to work remotely from both the school and the office, the version control system 'Git' and its corresponding platform 'Github' were used. Moving the code and working on the project from two different places worked great, except for an issue with one specific unity asset.

Some times when we were required to pull the latest version of the project to a computer, the interactable windowshader got corrupted. In theory it should not be that difficult to rebuild the element and copy it into the correct positions, which was the case the first time. However, when pulling the code to the technology consultancy's computer after the factory reset, the windowshader both got corrupted and did not work as intended when the file was replaced. This issue took a couple of days to resolve, efforts were also spent elsewhere during this time, however still a substantial amount of time spent. The issue turned out to be a combination of the windowshader file format and the reset of the computer. The windowshader was made as a .blend file. The .blend format is the format used when modeling the asset in Blender. Usually when a model is complete, it is exported from blender as a .fbx, which is the file format used when working with the asset in Unity. However, the windowshader made was never exported as a .fbx but was directly imported into the Unity project as a .blend file. This only worked because Blender was installed on the computer, which we at the time did not realize. Later when the computer got factory reset, Blender was, along with everything else, removed from the computer. Meaning the windowshader.blend file did not work, and when we tried exporting it as a .fbx file, the asset functioned differently in the scene. This was finally resolved by once again downloading Blender to the computer.

5.4 Reflections

When developing in an already existing project, there will always be certain difficulties due to the given hardware components and software systems, which are predetermined. Even though Meta Horizon Link provide huge complications early on in the project, it was not an option to switch VR headset brand since the previous versions had been developed with Meta. In spite of that, if this project were not made in collaboration with the company, we would most likely have developed all of our environments by ourselves from scratch. Then it would also have been easier to navigate and gradually edit the code when changes were required.

Reflecting back, the best solution to avoid a majority of the technical problems faced, would have been to start developing at the computers at Lund University at an earlier stage. This would have saved a lot of time from thoroughly analyzing code, troubleshooting, and finally factory resetting the computer and having to reinstall many programs. However, taking into account our knowledge of VR, prior to this master thesis, if given the chance to redo the project with the same knowledge we had then, we would probably have troubleshooted the same way as now. Only by having experienced issues with Link, as during this project, could we have been aware of the issues, and therefore make the troubleshooting more effective.

5.5 Future Work

The following section discusses the potential future work of research related to this master thesis. Research that the authors deem to be close in nature to the research of this master thesis, as well as general research focused on realism and presence in VR.

5.5.1 Research with Different Task and Environment

Given that the collaborating company is interested in the mechanics of the IFE system on an airplane, it was natural for this master thesis to test their system. Referring to the user tests being very focused on the IFE screen in front of the participant. However, to gain more perspective on whether presence affects performance in a virtual environment, further research in different and more demanding environments could be compelling. This is further discussed in section 5.1. Having participants perform tasks in an environment or using VR objects instead of a static screen in front of them could have different impact on task performance and completion time. Many participants stated that is was easy to filter out background actives since it they focused on the screen. If the tasks instead required to move around and interact with the environment would increased realism and higher present affect the result?

5.5.2 Broader Group of Participants

As mentioned in Limitations, the group of participants were all employees or associated with the company of this collaboration. A majority also has similar background working with technology and development. In a future project similar to this it would be interesting to include a wider group of people with different backgrounds and main occupation.

5.5.3 Isolation of Realism Factors

This project compared a more realistic environment with a less realistic one to investigate whether increased realism leads to a higher level of presence. Multiple elements were changed or added simultaneously, including lighting, sound, interactive objects as well as visual aspects, which all contributed to an overall increase in perceived realism. While the visual differences were clearly noticed by all participants, the impact of the other factors was not explicitly mentioned as much.

Future work could focus on isolating these individual components to better understand their specific contributions to perceived realism and presence. For instance, separate user studies could be conducted where only one factor is modified at a time, such as sound, natural lighting, or having interactive objects. This would make it easier to evaluate which elements have the biggest influence on user perception and experience. Additionally, future studies could also investigate having combinations of these factors to see whether certain elements can reinforce each other in increasing presence.

6 Conclusions

Will spending more effort to increase realism in a Virtual Reality environment affect user performance and is it worth it?

The results from the user tests of this master thesis concludes that a heightened presence does not affect user performance, in the regards to speed and accuracy. In an environment where the user is made to interact more with their surroundings, this study suggests that presence has a greater effect on performance.

Two environments with identical templates where one includes a few realistic components resulted significantly higher presence. Indicating that small changes has a big impact on the users perceived presence. Higher presence results in greater engagement by participants, which suggests that adding small changes is beneficial.

The conclusion is therefore that, for VR research with simpler tasks, which are not heavily dependent on interacting with the VR environment, a simple environment will satisfy. If the research is more complex, and perhaps more demanding for the user, the research could benefit from allocating resources to the design of the environment. This is also true for research focusing on behavioral and emotional responses.

References

[1] R. Andersson and F. Voigt. *Evaluating virtual reality as a medium for usability testing on inflight entertainment applications*. English. Student paper. 2023. URL: http://lup.lub.lu.se/student-papers/record/9125491.

[2] I. Bosman, O. Buruk, K. Jørgensen, and J. Hamari. "The effect of audio on the experience in virtual reality: a scoping review". *Behaviour & Information Technology*, 2024. DOI: 10.1080/0144929X.2022.2158371. URL: https://doi.org/10.1080/0144929X.2022.2158371.

[3] S. Chang and J. Suh. *The impact of VR exhibition experiences on presence, interaction, immersion, and satisfaction: focusing on the experience economy theory*. 2025. URL: https://doi.org/10.3390/systems13010055 (visited on 2026-01-26).

[4] M. DeJonckheere and L. M. Vaughn. "Semistructured interviewing in primary care research: a balance of relationship and rigour". *Family Medicine and Community Health* **7**:2, 2019. DOI: 10.1136/fmch-2018-000057.

[5] W. Do and V. Phung Nguyen. *Investigating how VR eye tracking can be used in usability testing of inflight entertainment systems*. English. Student paper. 2025. URL: http://lup.lub.lu.se/student-papers/record/9183663.

[6] DSDM Consortium. *DSDM agile project framework*. 2014. URL: https://www.agilebusiness.org/dsdm-project-framework/moscow-prioririsation.html (visited on 2026-01-26).

[7] M. Grimshaw. "Game sound technology and player interaction: concepts and developments". In: *Game Sound Technology and Player Interaction: Concepts and Developments*. Information Science Reference, 2011. ISBN: 9781616928308. DOI: 10.4018/978-1-61692-828-5.ch014.

[8] M. Hamash and P. Tiernan. "Trends in virtual reality research for post-primary education (2013–2024)". In: *International Conference on Engineering Education*. 2024. URL: https://ieeexplore-ieee-org.ludwig.lub.lu.se/stamp/stamp.jsp?tp=&arnumber=10923821.

[9] igroup. *Igroup presence questionnaire (IPQ) overview*. 2025. URL: https://www.igroup.org/pq/ipq/index.php (visited on 2026-04-16).

[10] Interaction Design Foundation. *What is brainstorming?* 2016. URL: https://www.interaction-design.org/literature/topics/brainstorming (visited on 2026-01-28).

[11] J. P. A. Ioannidis, D. Fanelli, D. D. Dunne, and S. N. Goodman. "Meta-research: evaluation and improvement of research methods and practices". *PLOS Biology* **13**:10, 2015, e1002264. DOI: 10.1371/journal.pbio.1002264. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC4592065/.

[12] R. Krause and K. Pernice. *Affinity diagramming for collaboratively sorting UX findings and design ideas*. 2024. URL: https://www.nngroup.com/articles/affinity-diagram/ (visited on 2026-01-27).

[13] C. Ma and T. Han. *Combining virtual reality (VR) technology with physical models: a new way for human-vehicle interaction simulation and usability evaluation*. 2019. URL: https://link.springer.com/chapter/10.1007/978-3-030-22666-4_11 (visited on 2026-01-14).

[14] Meta. *Meta quest pro is now available*. 2022. URL: https://about.fb.com/news/2022/10/meta-quest-pro-is-now-available/ (visited on 2026-01-23).

[15] Meta. *Set up and connect meta horizon link and air link*. 2026. URL: https://www.meta.com/help/quest/509273027107091/ (visited on 2026-04-19).

[16] E. J. Ramirez and S. LaBarge. "Real moral problems in the use of virtual reality". *Ethics and Information Technology* **20**:4, 2018, pp. 249–263. DOI: 10.1007/s10676-018-9473-5.

[17] T. Russell-Rose and T. Tate. "Social search". In: *Designing the Search Experience: The Information Architecture of Discovery*. Morgan Kaufmann, 2013. Chap. 9, pp. 253–275. ISBN: 978-0-12-396981-1. DOI: 10.1016/C2011-0-07401-X.

[18] H. Snyder. "Literature review as a research methodology: an overview and guidelines". *Journal of Business Research* **104**, 2019, pp. 333–339. ISSN: 0148-2963. DOI: 10.1016/j.jbusres.2019.07.039. URL: https://www.sciencedirect.com/science/article/pii/S0148296319304564.

[19] V. Souza, A. Maciel, L. Nedel, and R. Kopper. "Measuring presence in virtual environments: a survey". *ACM Computing Surveys*, 2021. DOI: 10.1145/3466817. URL: https://dl.acm.org/doi/epdf/10.1145/3466817 (visited on 2026-01-27).

[20] J. Supanat and A. Anucha. "Image-based z-fighting detection in 3d renders using deep learning approaches". In: *2025 17th International Conference on Knowledge and Smart Technology (KST)*. 2025, pp. 1–6. DOI: 10.1109/KST65016.2025.11003339.

[21] United Nations. *Global indicator framework for the sustainable development goals and targets of the 2030 agenda for sustainable development*. 2025. URL: https://unstats.un.org/sdgs/indicators/Global-Indicator-Framework-after-2025-review-English.pdf (visited on 2026-01-20).

[22] United Nations. *Transforming our world: the 2030 agenda for sustainable development*. 2015. URL: https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981 (visited on 2026-01-19).

[23] Unity Technologies. *What is unity?* URL: https://learn.unity.com/tutorial/what-is-unity?version=undefined (visited on 2026-01-28).

[24] Video Game Insights. *The Big Game Engines Report of 2025*. Report. Video Game Insights / Sensor Tower, 2025. URL: https://app.sensortower.com/vgi/assets/reports/The_Big_Game_Engines_Report_of_2025.pdf (visited on 2026-01-28).

[25] J. Young, T. Langlotz, T. Tran, T. Schuber, and H. Regenbrecht. “Classifying presence scores: insights and analysis from two decades of the igroup presence questionnaire (IPQ)”. *ACM Transactions on Computer-Human Interaction*, 2024. URL: https://www.hci.otago.ac.nz/papers/TranACMTOCHI2024.pdf (visited on 2026-05-03).

A Appendix

A.1 Pre Screening phase 1

A.1.1 Pre-test Questions

- What is your Age?
- What gender do you identify with?
- How often do you use VR?
  - Never
  - Tried once/few time (not frequently)
  - Used to use frequently (not frequently anymore)
  - Couple time every year
  - 1-3 per month
  - 1-2 per week
  - 3+ per week
- Have you ever used VR before?
- In what way have you used VR?
  - I have never used VR regularly
  - Developed in VR
  - worked with VR (not developed)
  - Played games in VR
- Do you easily get motion sickness, feel dizzy or nausea?

A.1.2 Pre-test Answers

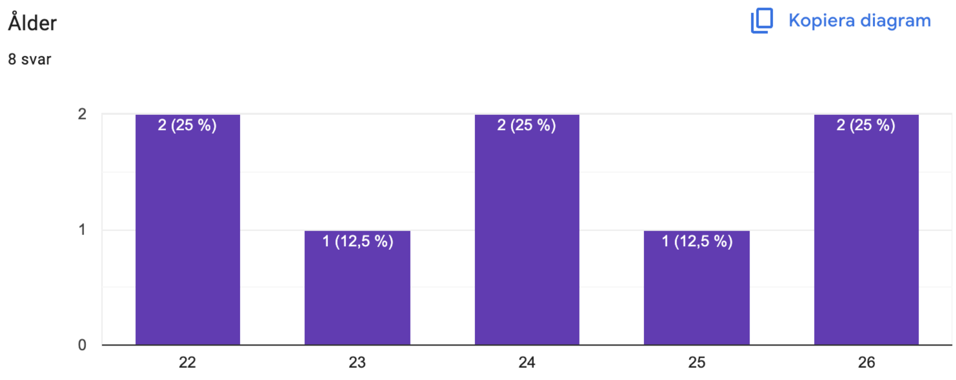


Figure A.1 Age of participants

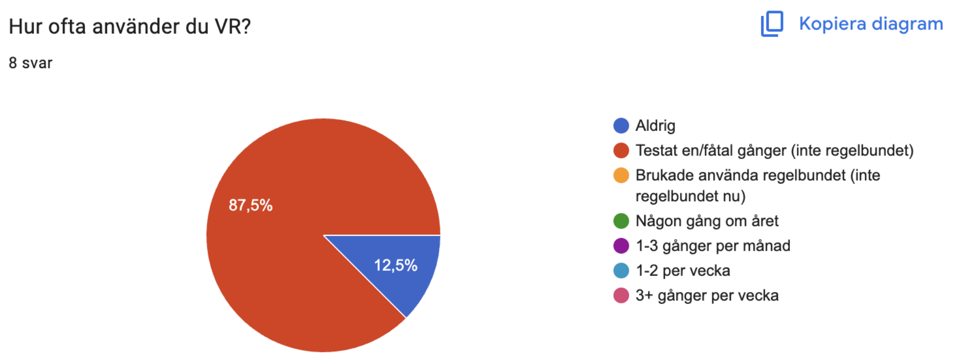


Figure A.2 At what extent have you used VR

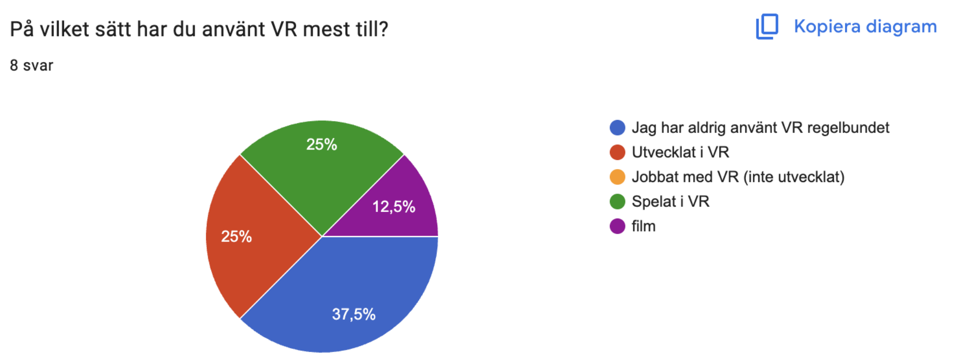


Figure A.3 How have you used VR

A.2 Pre Screening phase 2

A.2.1 Pre-test Questions

- What is your Age?
- What is your main occupation?
- What gender do you identify with?
- How often do you use VR?
  - Never
  - Tried once/few time (not frequently)
  - Used to use frequently (not frequently anymore)
  - Couple time every year
  - 1-3 per month
  - 1-2 per week
  - 3+ per week
- Have you ever used VR before?
- In what way have you used VR?
  - I have never used VR regularly
  - Developed in VR
  - worked with VR (not developed)
  - Played games in VR
- have you tried any of the previous version of this project from prior other master thesis?
- How often do you play computer or video games?
  - Less than a few times a year
  - Couple time every year
  - 1 per month
  - 2-3 per month
  - 1 per week
  - 2-3 per week

- 3+ per week

- Do you easily get motion sickness, feel dizzy or nausea?

- Do you use glasses or contact lenses?

A.2.2 Pre-test questionnaire results

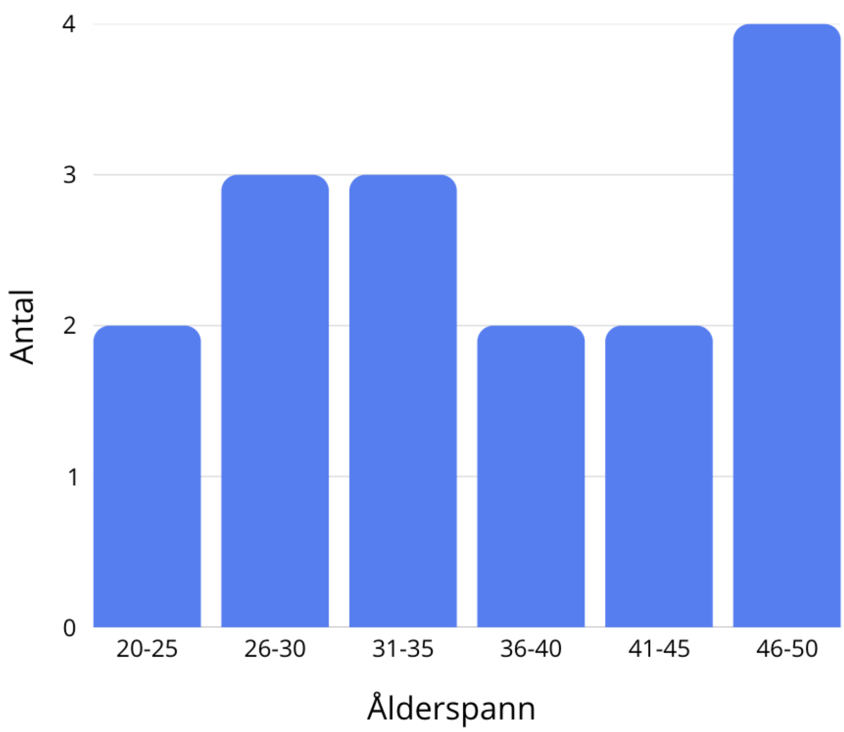


Figure A.4 Age distribution of participants

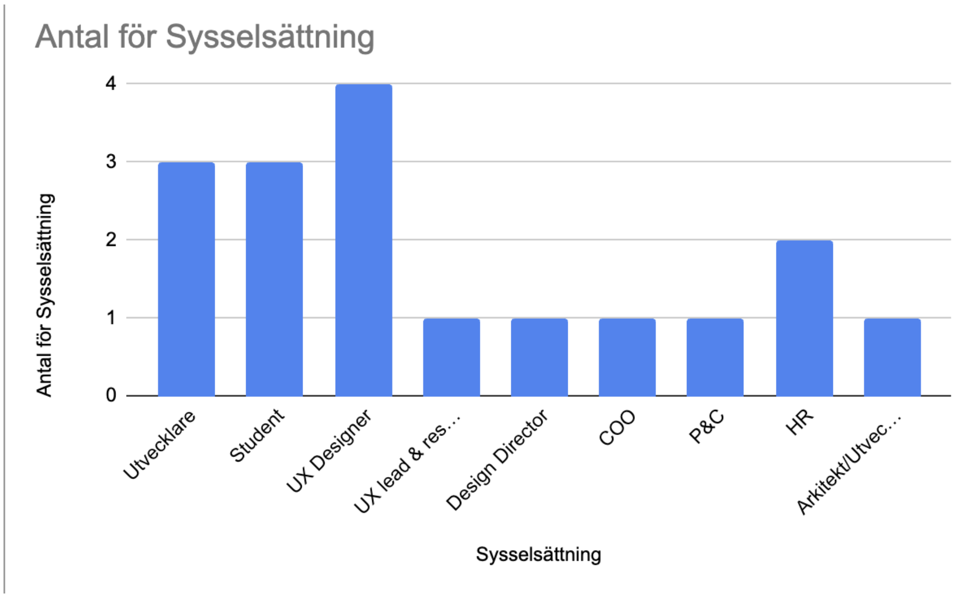


Figure A.5 What is your main occupation

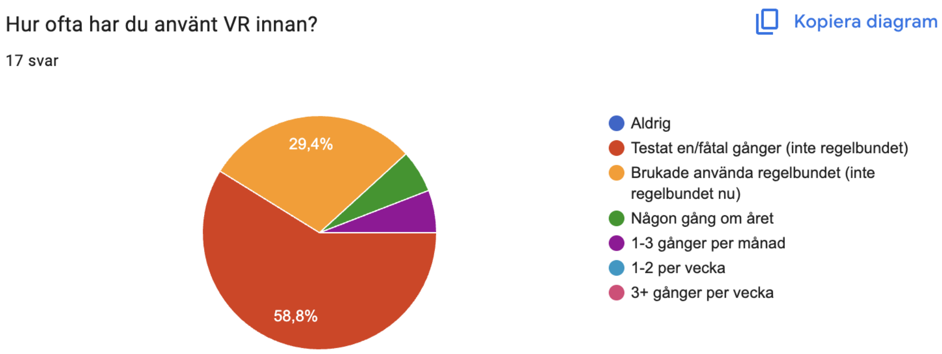


Figure A.6 How often have you used VR before

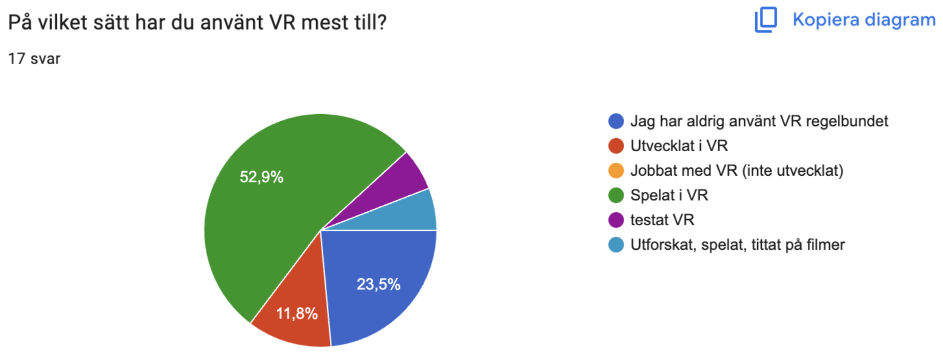


Figure A.7 How have you used VR before

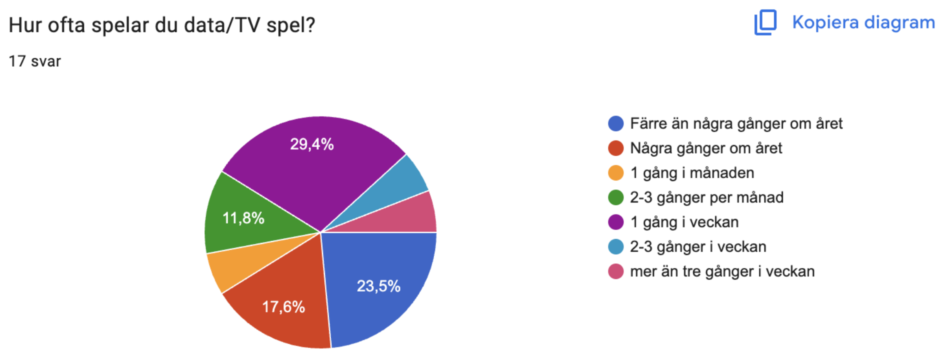


Figure A.8 How often do you play computer or video games?

A.3 IPQ

A.3.1 Items in the IPQ

1. How aware were you of the real world surrounding you while navigating in the virtual world (e.g., sounds, room temperature, other people)?

2. How real did the virtual world seem to you?

3. I had a sense of acting in the virtual space, rather than operating something from outside.

4. How much did your experience in the virtual environment seem consistent with your real-world experience?

5. How real did the virtual world seem to you?
6. I did not feel present in the virtual space.
7. I was not aware of my real environment.
8. In the computer-generated world, I had a sense of “being there”.
9. Somehow, I felt that the virtual world surrounded me.
10. I felt present in the virtual space.
11. I still paid attention to the real environment.
12. The virtual world seemed more realistic than the real world.
13. I felt like I was just perceiving pictures.
14. I was completely captivated by the virtual world.

A.3.2 Result from the IPQ

	NPC														
	User/Task	1 - NPC	2 - flippe	3	4	5	6	7	8	9	10	11 - flipped	12	13 - flippe	14
	1	−1	−1	3	−2	1	2	0	2	3	2	−2	−3	3	3
	2	1	−1	2	2	1	1	2	2	3	2	2	1	2	2
	3	−1	1	−1	1	1	1	−1	1	1	0	−3	−1	2	−2
	4	1	−2	0	2	−2	0	0	0	1	0	1	−3	−1	−1
	5	2	2	−2	2	−1	2	1	1	2	1	−2	−1	3	1
	6	2	0	3	−1	1	2	2	1	3	2	2	−3	3	3
	7	2	1	3	2	0	2	2	2	3	3	2	−3	3	2
	8	−1	−1	3	1	1	2	−3	2	2	2	−1	−3	3	2
	9	−3	2	3	2	2	3	2	3	3	3	0	−1	2	3
	10	−2	1	1	1	1	1	−2	1	1	1	−3	−3	0	0
	11	2	2	1	1	−2	1	−2	1	1	2	−1	−2	2	1
	12	−2	1	3	2	1	1	3	2	2	3	3	−3	3	2
	13	2	2	3	2	2	1	2	2	2	2	−1	−2	3	2
	14	0	−2	2	1	−1	0	−3	2	1	2	0	−3	1	−2
	15	−2	2	1	1	1	1	0	1	1	2	−1	−3	2	0
	16	2	−3	2	1	−3	0	1	1	1	1	1	−3	2	0
	17	1	−2	0	2	−1	3	−3	1	1	1	−2	−3	3	−2
Mean Value (w/ out User 1):		0,25	0,2	1,5	1,4	0,1	1,3	0,1	1,4	1,8	1,7	−0,2	−2,3	2,1	0,7

Figure A.9 Mean result for each item in IPQ for NPC environment

Robot

	Task	1 - robot	2 - flippe	3	4	5	6	7	8	9	10	11 - flipped	12	13 - flippe	14
	1	−2	−1	3	1	1	0	−2	0	3	0	−3	−2	3	1
	2	0	0	2	2	0	1	2	1	3	2	2	1	2	2
	3	1	−1	−1	0	−1	0	0	0	−1	−1	−3	−2	1	−3
	4	−1	−2	−1	1	−2	0	−1	−1	1	0	1	−3	−1	−1
	5	0	1	−2	1	1	1	1	0	2	1	−2	−2	3	0
	6	2	0	3	−1	−2	1	2	0	3	2	2	−3	3	3
	7	2	0	1	−2	−1	0	0	0	1	1	1	−3	3	0
	8	−1	−2	3	−1	0	2	−3	2	2	2	−1	−3	3	2
	9	0	−2	2	2	0	2	1	2	2	2	0	−2	1	2
	10	1	−2	−1	−2	−1	−2	−2	−3	−2	0	−3	−3	0	0
	11	2	0	1	1	−1	1	−2	1	1	2	0	−2	2	1
	12	1	0	3	2	0	0	3	2	2	2	3	−3	3	2
	13	0	−2	1	−2	−2	−1	−2	1	2	1	−1	−3	3	1
	14	−1	−3	2	−1	−3	−1	−3	0	1	2	0	−3	1	−2
	15	−1	0	0	−1	−1	0	−1	0	0	1	−2	−3	1	−1
	16	2	−1	2	1	−1	0	1	1	1	1	1	−3	2	0
	17	1	−2	0	2	−1	−3	−3	1	1	1	−2	−3	3	−2
Mean Value (w/ out User 1):		0,5	−1	0,9	0,1	−0,9	0,1	−0,4	0,4	1,2	1,2	−0,3	−2,5	1,9	0,3

Figure A.10 Mean result for each item in IPQ for Robot environment