Master of Science in Engineering: Game and Software Engineering
June 2024

Evaluation of Hand Collision in Mixed Reality

Adrian Tegelind

Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden

This thesis is submitted to the Faculty of Computing at Blekinge Institute of Technology in partial fulfilment of the requirements for the degree of Master of Science in Engineering: Game and Software Engineering. The thesis is equivalent to 20 weeks of full-time studies.

Contact Information:
Author(s):
Adrian Tegelind
E-mail: adte17@student.bth.se

University advisor:
Associate Professor Veronica Sundstedt
Department of Computer Science

Faculty of Computing
Blekinge Institute of Technology
SE–371 79 Karlskrona, Sweden

Internet : www.bth.se
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57

Abstract

Background. With the growing prospects of extended realities (XR), new use cases and experiences are constantly being developed. Especially with the introduction of mixed reality (MR), allowing for a more seamless blend of the physical and digital space, it provides great opportunities in many fields such as education and training where dangerous procedures can be practiced safely. However, to make these experiences as effective and educational as possible, there is a need to make the experiences realistic.

Objectives. One important aspect of creating realistic experiences is believable collision between the user's physical hand and the digital objects. This study specifically takes aim at this aspect. Trying to find how the performance difference and user experience (UX) is affected by the addition of collision around the user's hands in an MR environment. In order to help guide the way to get the answers to these questions, a set of objectives has been formulated. These objectives are; finding and implementing a hand collision method, designing and performing the user study, and finally finding and utilizing appropriate methods for analyzing the collected data.

Methods. To get a better understanding of the UX and performance of using hand collision, a user study was created where the participants had to complete a series of tasks, with and without collision around their hands. For each task, answering a questionnaire about their experience. Once the data have been collected, it will be analyzed with the help of the SUS scoring system and statistical tests.

Results. The study had 12 participants. With and without hand collision received an average SUS score of 62,5 and 69,2 respectively. The results show that the method using no collision performed better in terms of time to complete the task. However, hand collision performed better with fewer grabs used. No statistically significant difference was detected between having or not having hand collision in terms of intuitiveness and realism. However, participants were observed to intuitively use the hand collision to their advantage.

Conclusions. In conclusion, the participants did not perform better with hand collision, however, did indicate some level of increased intuition and realism. The negative aspects of the hand collision are believed to be attributed to the method used to implement it, and potential in the area exists for further improvements and research.

Keywords: Mixed Reality, HoloLens, Hand Collision, User Study.

Sammanfattning

Bakgrund. Med ett växande potential för extended realities (XR), nya användningsområden och upplevelser utvecklas ständigt. Speciellt med införandet av mixed realities (MR), möjligjorde en mer enad upplevelse av det fysiska och digitala, med stora möjligheter inom utbildning och träning där det farligt sitvationer kan övas på ett säkert sätt. Men, för att göra dessa upplevelser så effektiva och pedagogiska som möjligt behöves mer realistiska upplevelser.
Syfte. En viktig aspekt av att skapa realistiska upplevelser är att skapa trovärdiga kollisioner mellan användarens fysiska hand och the digitala objekten. Detta är ett av målen denna studien tar sikte på. Att försöker hitta hur prestandaskillnaden är och användarupplevelsen (UX) påverkas med tillägget av kollision runt användarens händer i en MR-miljö. För att enklare kunna hitta en väg till svaret för dessa frågor har mål formulerats. Dessa mål är att; hitta och implementera en handkollisions metod, designa och utför en användarstudie, och hitta samt använd lämpliga metoder för att analysera den insamlade datan.
Metod. För att få en bättre förståelse för hur UX och prestanda för användning av handkollsion skapades en användarstudie där deltagarna genomförde en serie uppgifter, med och utan kollision runt deras händer. För varje uppgift besvarades ett frågeformulär om deras upplevelse. När uppgifterna har samlats in kommer de att analyseras med hjälp av SUS poängsystem och statistiska tester.
Resultat. Denna studie hade 12 deltagare. Med och utan handkollision fick en genomsnittlig SUS-poäng av 62,5 respektive 69,2. Resultaten visar att metoden som inte använder någon kollision presterade bättre när det gäller tid för att slutföra uppgiften. Men, kollision fick dock bättre med resultat med ett färre antal grepp som används. Det var ingen statistiskt signifikant skillnad som upptäcktes mellan med och utan handkollision i avsikt på intuitivitet och realism. Dock observerades deltagarna att använda kollisionen på ett mer intuitivt sätt till sin fördel.
Slutsatser. Sammanfattningsvis, deltagarna presterade inte bättre med kollision, men indikerade viss nivå av ökad intuition och realism. De negativa aspekterna av kollisionen tros att vara i grund på den metod som använts för att implementera den, och potential finns inom området för ytterligare förbättringar och forskning.

Nyckelord: Mixed Reality, HoloLens, Handkollision, Användarstudie.

Contents

Chapter 1

Introduction

For this chapter, an introduction of the relevance for the research conducted in this thesis is presented and its potential benefits. This area of research is the use of adding digital hand collision to the user's physical hands via hand-tracking in a mixed reality environment. This is done by using the data collected from the hand-tracking to place and fit digital physics colliders around the user's hands. This should, in theory, allow the user to more realistically interact with digital objects with only there bare hands. The benefits described in this section covers various fields this technology could be used in, as well as sustainability aspects. Furthermore, related areas are covered to provide foundational understanding and terms used for the relevant fields, such as the world of extended realities and simulated physics.

1.1 Background

Augmented reality (AR) generally refers to the overlaying sensory information like visuals layered over the physical world [16]. Virtual reality (VR), on the other hand, is a fully digital experience that completely obscures the real world [16]. Mixed reality (MR) is commonly referred to as a combined experience of the physical and digital world where elements of each world can affect and interact with each other in an immersive way [31]. Extended reality (XR) is an umbrella term used for the various realities like AR, VR, and MR [28]. For more information on the XR and their definitions, see Section 1.6.

XR is currently a large and growing industry, as shown in the white papers published by the XR association and authored by Aaronson et al. [1]. With an estimated $41 billion in worldwide market size for 2023 and is projected to be $100 billion by 2026 [1]. Furthermore, the XR market is calculated to contribute with a $1.5 trillion in global GDP by 2030 [1]. As well as growing economical prospects, AR/MR has also seen significant growth in scientific research as shown by Sünger and Çankaya, based on the number of scientific papers on the subject released on Scopus [32]. They also found that the most prominent research areas in MR are education, advertisement, architecture/decoration, repair/maintenance, and healthcare [32].

MR has potential in many different areas, but one of the more common examples is in education and healthcare where dangerous or risky procedures can be demonstrated and practiced safely in a more realistic environment. Examples of these are surgical procedures, as seen in the work of Hu et al. [11] and military training as seen in the work of Hughes et al. [12]. However, for these experiences to be effective educationally, realistic ways of interfacing with the digital environment are needed to simulate real-life scenarios accurately.

This study aims to test the impact of adding digital collision around the user's physical hands in an MR environment. This is done to study the performance and user experience (UX) given by the addition of hand collision. Collision should, in theory, increase the intuitiveness, realism, and accuracy of the experience. The argument being that having collision around one's hands allows the user to interact with the digital objects in a more similar manner to real life.

1.2 Aim

This study aims to evaluate different direct hand manipulation methods in MR in terms of UX, particularly in relation to intuitiveness and realism. While there are many different ways of interfacing with objects, for example, voice commands, eye-tracking, controllers, and gestures, this study will focus on the addition of hand collision and physics. While haptic feedback could potentially be used to enhance the experience of the simulated hand collision, it will not be tested in this thesis. This is to see if the addition of collision brings any positive aspects in terms of performance and UX. Additionally, this study also performs the experiment in MR, which seems to be a somewhat unexplored area.

1.3 Research Questions

This section covers the project's selected research questions (RQ). UX for RQ1 will be defined according to the system usability scale (SUS) [4] method with some additional attributes added such as realism and intuitiveness. User task performance in RQ2 will be defined as objective data collected during the user study, such as the time it took to complete a task and hand-object interaction metrics like the number of grabs. Number of grabs is a metric that has previously been used to measure accuracy of the manipulation of objects. This can be seen in Navarro and Sundstedt paper for evaluation of manipulation techniques [20]. There are also multiple ways of implementing hand collision but this research question references to the one described in Section 3.1. There are two research questions this work tries to answer:

- RQ1: What is the user experience with and without hand collision in mixed reality?
- RQ2: What is the user task performance with and without hand collision in mixed reality?

1.4 Objectives

To achieve the aim and answer the research questions previously described, a set of objectives (Ob) needs to be accomplished. These objectives will be defined in this section of the thesis. The objectives for this thesis are:

- Ob1: Find and implement an appropriate hand collision method.
- Ob2: Design and perform tests for the user study to collect data for the hand collision experience.
- Ob3: Find and utilize appropriate methods to analyze the collected data.

1.5 Sustainability aspects

In terms of sustainability, AR and MR research has great potential for positive societal impacts, especially in areas such as education, training, [11], [12], [18], [32], and, as shown by the World Economic Forum's report, communication [18]. One previously given example is the simulation of potentially dangerous and risky procedures that allow people to train in a safer and more sustainable way while also being economically viable [11], [12].

More broadly speaking, general training in XR has also been reported to be more cost-saving, scalable, and more efficient than in-person training due to more visual aids and less staff needed to perform the training [18]. Furthermore, this technology has also allowed people to collaborate in ways that make the user feel more connected, allowing remote work while still keeping a socially sustainable work environment [18]. This experience could be further enhanced by having more realistic interactions between the users.

Research in AR and MR has also shown the potential to create more inclusive and accessible work environments by enabling people with disabilities to do work that previously had been too difficult. For example, giving employees with cognitive disabilities more visual instructions through technology like AR instead of language-based instruction like manuals [18]. Overall, researching intuitive and realistic ways to interface with these systems can improve effectiveness and increase these positive aspects.

1.6 Definition of Various Extended Realities

One of the earlier works of defining the various extended realities was done by Milgram and Kishino in 1994 [16]. They categorized different visual displays and defined what is called a virtual continuum as seen in Figure 1.1, a spectrum with an unmodified view of reality on the left and a completely VR view on the right [16].

In between the endpoints of the virtual continuum, there is AR which mainly consists of displaying information from the real world with some added digital elements, and augmented virtuality (AV), which is primarily virtual with some added real-world elements [16]. In this paper, Milgram and Kishino also considered mixed

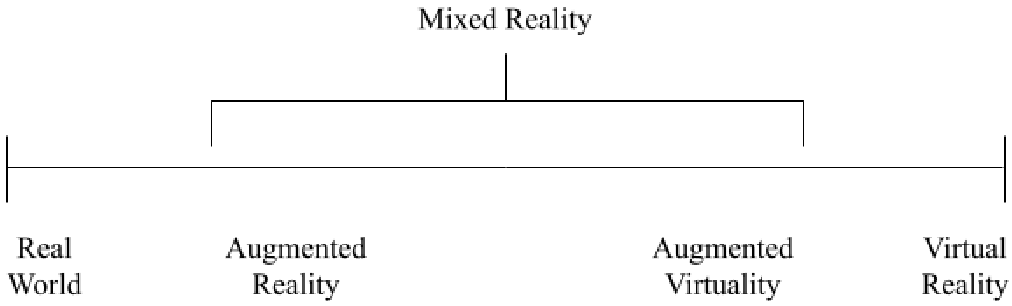


Figure 1.1: A figure showing the virtual continuum.

reality (MR) to be all the versions between (but not including) the endpoints, i.e., all combining of the real and digital world [16].

However, since Milgram and Kishino's definitions were only for displays and in today's world, many AR/AV/VR/MR devices can have a lot of other features attached to them, such as controllers, eye trackers, motion sensors, etc. These additions can change the experience and use of these devices, creating a need for more definitions. MR is the one that has had the most change. One study conducted by Spechier et al. found that the term MR does not have a singular consistent definition [31]. However, one of the more commonly used definitions is that it is a combined experience of the physical and digital world where elements of each world can affect and interact with each other in an immersive way [31]. Finally, there is also the term extended reality (XR). As found by Rauschnabe et al. XR tends to be used as an umbrella term, referring to everything after (but not including) the starting point, the "real world", in the virtual continuum [28]. However, Rauschnabe et al. argue that the name extended reality is a misleading term since it includes the term VR which is its own encapsulated reality rather than an extension to the real world [28]. Instead, they propose the term xReality, where the x implies an unknown variable [28]. Nonetheless, this thesis will refer to XR as extended reality since it is the more widely used and recognized term. In summary, for this thesis, the following definitions will be used:

- Augmented Reality (AR): The layering of digital elements over a mostly physical world.
- Augmented Virtuality (AV): The layering of physical elements over a mostly digital world.
- Virtual Reality (VR): A fully digital and encapsulated environment.
- Mixed Reality (MR): The combining of digital and physical elements where they can interact in an immersive way.
- eXtended Reality (XR): A catch-all term for all of the above.

1.7 Physics In Games and Software

This section will cover basic physics concepts and terms used in games and physics simulations. This is relevant for the study since a lot of physics methods are needed to construct the collision used for the hand collision tested. A lot of the information will come from the documentation of PhysX made by NVIDIA [22], [23], [24], [25] and Box2D's documentation authored by Catto [6], [7] which are two commonly used physics libraries for software and games. As an example of their use, it is stated in Unity Technologies' Unity documentation [33] and Epic Games' Unreal Engine documentation [10] that they both use PhysX. Unity also uses Box2D for 2D physics solutions [33].

In reality, physical objects can be very complex systems where materials warp and deform after physical stress. Deformable bodies in physics are called soft bodies [25]. Soft bodies are of course relevant for things where deformation is relevant like calculations/simulation of things like cloth. However, even though all materials can deform, in many cases it is to such a small degree that it is negligible. In physics, simplifications are often made to remove factors that are mostly negligible. This is done in order to simplify the calculations and to save performance in a simulation. One commonly ignored effect is the deformation of physics objects, for example during a collision. An object that can not be deformed is called a rigid body [24], [7].

1.7.1 Rigid physics bodies

Rigid physics bodies are defined, as previously mentioned, by the fact that they can no deform [24], [7] and come in three variations; static, kinematic, and dynamic. Static physics objects are stationary objects that are unaffected by physical forces and do not move when colliding with other objects [23], [6]. Static physics objects are usually a simplification of objects like walls or the ground. From a mathematical perspective, as also stated in Palmer's book, immovable objects are assumed to have an infinite mass, [26], [24], [6]. Similarly to static physics objects, kinematic physics objects can not be affected by other forces and can exert forces on other objects. However, kinematic objects have the difference that they can move [23], [6]. Kinematic physics objects are often used in video games where you have objects like platforms that move back and forth along a determined path and do not deviate from it. Dynamic physics objects are more in line with real-life objects where they can both inflict physical forces and be affected by them [23], [6].

1.7.2 Joints and constraints

Other common methods in physics programming are joints and constraints. The methods of joints and constraint are used for the control, relation, and movement between physics objects [22], [6]. Joints are used to connect physics objects in a simulation, allowing them to move relative to each other while respecting certain physical constraints [22], [6]. Joints are typically used to model the way physics objects can articulate. There are several types of joints commonly used in physics programming. An example of a very common joint is a revelation joint (sometimes also called a hinge joint) which allows for the rotation between two objects around a single axis, similar to a door hinge [22], [6]. A joint that allows for rotation around all axes is called a ball-and-socket (or spherical) joint [22]. A distance joint keeps the distance between two bodies within a range or at a fixed distance [22], [6]. These distances and rotations enforced by the joints can often be limited to a certain amount and set to be strict or soft [22]. A strict constraint creates a hard stop when the limit of distance or rotation is reached [22]. Soft constraints on the other hand have a slower and more elastic restriction between the bodies connected with a joint [22].

1.8 Outline

Chapter 1 called Introduction will cover the aim and scope of the research, its relevance, and related terms and terminology used throughout the thesis. In Chapter 2, Related Work, covers the prospects with hand-tracking in terms of interacting with digital objects. Furthermore, methods that have been used to enhance the hand-tracking experience will be brought up, amongst which is simulated hand physics, the focus of this thesis. Chapter 3, Method, describes the environment and implementation of the hand collision. It also gives an explanation of the experiment setup and design as well as the ethical aspects of the experiment and data collected. Chapter 3 also covers how the collected data was processed, for example, what statistical calculations were used and the scientific method used for the thesis. Chapter 4, called Results and Analysis, presents the collected data and analyzes it with various calculations. Chapter 5, Discussion, will then proceed to interpret the results from Chapter 4 and present options to what the results could indicate. Finally, Chapter 6, Conclusion, will give a quick summary of the thesis and distill the work down to the main findings.

Chapter 2

Related Work

This chapter will cover various research done in the space of hand-tracking and hand manipulation. Covering some of the research comparing hand-tracking and controller-based manipulation to see the current state of hand-tracking manipulation to the more common controller-based methods. Furthermore, the potential of hand-tracking as a manipulation method is evaluated and research where areas of improvement for performance and experience is brought up. In this case, the use of physics in hand-tracking and manipulation is detected as one of the bigger points of interest to be explored. Later on in the chapter, the gap and improvements will be identified based on the previous research.

2.1 Previous Research

Hand-tracking has for a long time been an attractive prospect for creating realistic and intuitive ways for the user to interact with objects. This section will cover the performance of hand-tracking compared with more traditional methods of interaction like controllers. Furthermore, it will give some motivation to why hand-tracking, as a way of interfacing with objects, is an appealing idea. It will then look at some avenues for improvements to make the experience of having the user use their bare hands to manipulate and interact with digital objects more realistic and effective. The main area that this thesis will focus on is the addition of simulated hand collision physics, allowing the user's hands to collide with digital objects.

2.1.1 Hand-Tracking and Its Potential

A lot of research has been done to compare different types of performance aspects between hand-tracking and controllers. Khundam et al. [13] in 2021 tested the performance between hand-tracking and controller with VR in a medical training program with 30 participants that required the users to perform actions like moving and rotating. They got a mean SUS score of 67,17 for controllers and 60,17 for hand-tracking. However, no statistically significant performance differences between the two methods were found in the end. They do note that hand-tracking is likely still the better option for long-term training [13]. This is because it is more similar to the real world, allowing the user to more intuitively recognize and correct posture, leading to more experience gained from practicing over a longer period of time [13]. Another study by Varela-Aldás et al. in 2023 also compared hand-tracking and controller in VR but in terms of reaction speed, perception of presence, and usability, among other things [34]. They tested this on 20 participants, however, did not find any significant differences between the two options [34].

While few advantages of hand-tracking over regular controllers have been found with current hand-tracking methods, many believe that hand-tracking still has great potential in the future for more effective and immersive interactions [13], [5]. As outlined by Buckingham, hands are essential for human visuomotor behavior, and even obscuring the user's hands can give feelings of disembodiment, hence are very important for immersion and engagement [5]. In a VR environment, hands can be digitally represented when using a controller, having the fingers of the digital hand move when pressing the button on the controller. While this helps with engagement, having the full functional and visual dexterity of one's hands gives a greater sense of engagement and ownership, creating a more immersive experience [5]. This is one of the areas where MR might have some extra potential over VR since the user can see their real hands and generating an accurate digital representation of the user's hands might not be necessary.

2.1.2 Haptic-feedback

However, hand-tracking on its own does not seem to be able to achieve a completely immersive, intuitive, and realistic experience. This has led to other things that have been tested to see what can improve direct interaction with the user's hands. A well-tested method is the addition of haptic feedback. One of these examples is Kim et al., who tested haptic feedback wristbands in VR on 21 participants and found in their study the participants experienced a higher level of immersion while using it [15]. Vermeulen et al. also tested haptic feedback gloves on 22 participants and found that they both improved performance and user enjoyment [35]. However, it is noted that the small improvement the haptic feedback gloves bring might not be worth the cost from a commercial perspective [35].

2.1.3 Hand Collision

Another avenue that has seen some exploration is the addition of collision physics that follows the user's hands. In terms of improving the UX for hand-tracking based manipulation, hand physics is a way to improve the UX while not necessarily increasing the cost of the XR products themselves. This is because it is a digital solution compared to haptic feedback gloves which require additional hardware. Boonbrahm and Kaewrat have written about their experience with implementing physics for the fingertips in an MR environment [3]. They tested different shaped colliders' effects on how effective they were at grabbing objects. Multiple different collider shapes such as cylinders, rectangles, and hexagonal shapes were tested and the best one was the cylindrical shaped collider [3]. Furthermore, they also compared their collision based grabbing to some usual gesture-based methods. While they felt that the collision felt a bit more natural than the gesture detection, it was only tested by the researchers themselves [3]. They also implemented this through the use of a webcam to perform hand-tracking, which do not provide as accurate measurement as dedicated headsets that tend to use more advanced sensors [3].

Nasim and Kim also tried a physics-based grasping method with assisted grabbing in VR [19]. They state that very little research have been done in the area of hand-tracking with collision physics and that more techniques for hand collision need to be explored and tested [19]. They mainly list to directions to take in terms of implementing hand collision. One way is a kinematic version, where kinematic colliders strictly follow the hand. This version is cheaper to implement but they note that it is generally less reliable with the colliders sometimes penetrating the objects which can cause odd physics behavior [19]. The other method is to use dynamic hand colliders that follow the less strictly which allows other objects to push away the hand colliders, avoiding penetration between the physics objects [19]. In the end they chose to go forward with the dynamic collider version since they prioritize accuracy over performance [19]. They tested their physics-based grabbing method with dynamic colliders against a more common grab detection algorithm with a group of 12 participants [19]. Overall, on average, the more physics-based grabbing seemed to have performed better [19]. However, no statistical test was performed making it hard to draw any strict conclusions from their results [19].

2.2 Information Gaps and Improvements

While there are some areas that have been explored in effort to improve the UX of hand-tracking in XR environments, hand physics remains somewhat unexplored [19]. The research that has been done seems to have had few to no participants and different hand physics methods remains untested. Furthermore, the user studies done regarding hand collision physics have also been lacking in statistical tests and analysis, giving uncertain results. Additionally, most of the research listed in Section 2.1 has been done in VR and very little in actual proper MR headsets. This also prompts more research since the potential in MR environments where the visual connection between the user and their hands has the potential to create more realistic and intuitive experiences. MR headsets often have worse performance than VR headsets because they more often run the software locally on the device while VR runs software more often on a PC and streams the data to the headset. While an MR headset also can stream data from a PC and some VR headset can run the software locally, in theory, the MR headset needs to be more mobile and therefore needs to run software locally. The ambition of an MR headset is to move around the physical world (unfettered to a PC) and display holograms in the real world [18] [32], while with a VR headset the user is often stationary in real space but moving around in virtual space. Nasim's and Kim's study also focuses on a dynamic collider implementation while the kinematic collider implementation still could be worth testing [19].

Chapter 3

Method

This part of the thesis will cover the implementation and development of the hand collision methods and an overview of the scientific method used. This will involve covering the methods used to analyze the collected data, such as the system usability scale method and the statistical tests used. It will also describe the experiment that was performed and the demographic, as well as the recruitment process of the participants. Finally, an ethical analysis is performed to assess and minimize the potential risk of the participants to ensure that the study is performed as ethically as possible.

3.1 Implementation and Environment

In the final hand collision implementation cylindrical colliders were used due to them generally being easier to grab objects with [3]. The hand colliders used were also kinematic physics colliders due to it being a more performance effective implementation [19]. This is relevant due to MR headset often having quite limited hardware that runs the software locally on the headset.

The program was developed for a HoloLens 2 MR headset and the software was made with the Unity engine (version 2020.3.28f1). For the development of the HoloLens 2 application, the Mixed Reality Toolkit 2 (or also called MRTK2) plugin was used to provide a basic interface for HoloLens 2 development. The MRTK2 was used to track the positioning and joints of the user's hands. This data was then used to orientate and place kinematic physics colliders. Some changes were also made in the Unity physics setting to get more accurate physics simulations. These settings can be seen in Table 3.1. The motivation for using Unity for the implementation, compared to Unreal Engine or programming it from scratch is that it is easier and faster to work with. Programming from scratch is too time consuming. In terms of using Unity compared to Unreal Engine is mainly motivated by personal preference and experience.

Table 3.1: Physics settings used in Unity (Edit->Project Settings->Physics)

Default Solver Iterations	25
Default Solver Velocity Iterations	15
Enable Adaptive Force	True
Friction Type	One Directional Friction Force
Solver Type	Temporal Gauss Seidel

3.2 Scientific Method

For this thesis, a scientific method will be used to formulate questions, collect and process data to produce results that conclusions can be derived from. The initial step of the process was background research to learn the area and locate relevant gaps in the field that needed more research. Once gaps are found, research questions are formulated and defined as seen in Section 1.3 as well as defining objectives as seen in Section 1.4. These objectives will help define the path needed to take in order to answer the research questions. A user study will then be performed to collect the relevant data related to the research questions. The relevant metrics are user performance and UX. User performance is objective data collected as the participant completes tasks. As for UX data is subjective and will be collected with the help of questionnaires. Once the data have been collected it will be analyzed. The objective data will be analyzed with statistical methods as motivated in Section 3.2.2. As for the subjective data, the questions will mostly be collected and measured according to the SUS method as seen in Section 3.2.1. Once the data have been analyzed, discussion surrounding the results will be performed and conclusions about the research questions can be drawn.

3.2.1 System Usability Scale Method

The System Usability Scale (SUS) is created by Brook and is a method for measuring the usability of a system [4]. This method is commonly used because it gives a standardized metric for usability comparisons between different systems. This is done by allowing the user to test a system and then answer a questionnaire consisting of 10 questions. Each question is answered on a scale from zero to four, giving a total of five options, which corresponds to strongly disagree and strongly agree. The questions used in SUS goes as follows:

- I think that I would like to use this system frequently.
- I found the system unnecessarily complex.
- I thought the system was easy to use.
- I think that I would need the support of a technical person to be able to use this system.
- I found the various functions in this system were well integrated.
- I thought there was too much inconsistency in this system.
- I would imagine that most people would learn to use this system very quickly.
- I found the system very cumbersome to use.
- I felt very confident using the system.
- I needed to learn a lot of things before I could get going with this system.

Once a person has filled out the questionnaire, some of the scores have to be inverted since some of the questions are asking about negative aspects of the system. Some of the score has to be inverted due to the fact that the numbers need to be summed up, and a higher value should relate to a better score in the final result. Once appropriate questions have been inverted, all the questions are added together and multiplied by 2,5 giving a final score between 0-100. An average is then taken of all the answers from the participants, and a final score is acquired. As found by Bangor et al., an ok score is above 50, a good score would be around 70, and an excellent score is around 85 [2].

3.2.2 Statistical Tests Motivation

In this study, multiple sets of samples need to be analyzed, which is done to a large degree through statistical tests. This section covers the various statistical methods used on the data as well as how they were used and the motivation behind their use.

A common first step before determining what statistical analysis to use, an initial calculation is done to determine what type of distribution the data have. This is often done to see if the data follows a normal distribution or not. For this study, Shapiro's and Wilk's test for normality called the Shapiro-Wilks test [30] is used to check if the data follows a normal distribution. In this case, the Shapiro-Wilks test is used due to its higher accuracy with smaller sample sizes, as shown by Mishra et al. [17]. Specifically, the Shapiro-Wilks test works better for 50 or fewer samples compared to the Kolmogorov-Smirnov test which works better for sample sizes larger than 50 [17]. In terms of how the method was applied, the calculations were done with Python, specifically a module called stats from the scipy library was used where the function `shapiro` was used to perform the actual Shapiro-Wilks test.

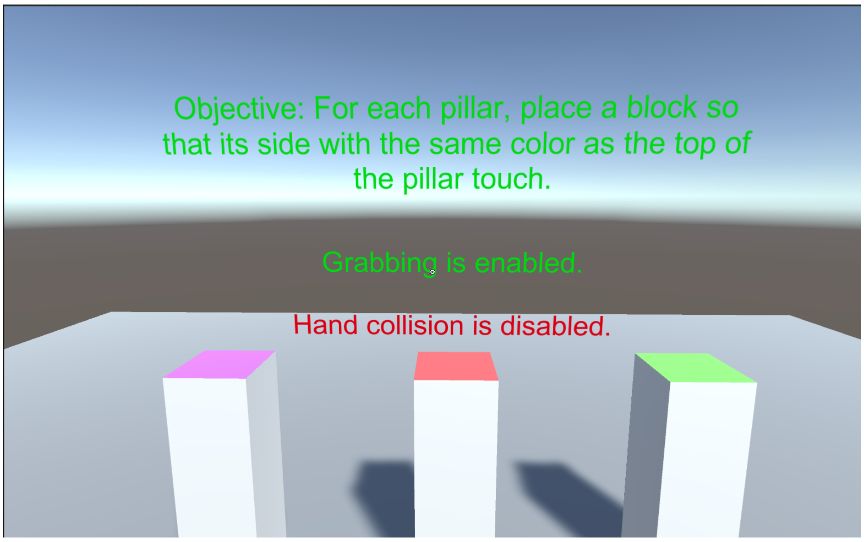
For the statistical analyses a paired t-test, Wilcoxon-test, and Mann-Whitney U-test was used. As described by Cunningham and Wallraven as well as Witte and Witte, the paired t-test is appropriate for comparing two samples of parametric data where they are collected from the same individual [36], [8]. This makes it a good option for comparisons between measurements of parametric data where the participants had hand collision and no hand collision. The paired t-test calculations were performed with Python, also with the stats module from scipy library using the function `ttest_rel`. The Wilcoxon-test is another alternative for similar situations as the paired t-test except when the data is non-parametric [37]. This test was also calculated with the same module but using the `wilcoxon` function. To analyze the five-point Likert data collected through the questionnaires that is not analyzed with the SUS method, the Mann-Whitney U-test was used. As shown by de Winter and Dodou, both t-test and Mann-Whitney U-test are good options for analyzing Likert data [9]. However, the t-test is generally a bit more accurate for scenarios where the data very closely follows a normal distribution compared to the Mann-Whitney U-test [9]. This is why the Mann-Whitney U-test was used for this study since the compared Likert data do not quite follow a normal distribution. Similarly to the other calculations, it was done with Python through the same module using the `mannwhitneyu` function.

For the statistical tests the significance level 0,05 was chosen. As described by Kim and Choi, to optimize the chosen significance level, the sample size, Type I and

II error rates, and trends about the subject area is needed [14]. However, beyond the sample size, these things are not known for this thesis and a common convention is to otherwise pick 0,05 as the significance level [14].

3.3 Experiment Description

In this study, the participants will complete a set of tasks with and without collision around their hands. This is done in an MR environment, specifically with the HoloLens 2 headset. The grabbing system is the default HoloLens 2 grabbing with long-distance grabbing disabled. The hand collision made for this study and implementation details can be seen in Section 3.1. While completing the tasks, both subjective and objective data are collected. The subjective data is about the participant's experience which is collected with the help of multiple questionnaires taken throughout the test. The objective data is collected based on the user's performance. This data is the time it took to complete a task and the number of grabbing actions the participant performed. Throughout the experiment, the participant receives information and instructions on what to do through text displayed in front of them, as seen in Figure 3.1.



Objective: For each pillar, place a block so that its side with the same color as the top of the pillar touch.

Grabbing is enabled.

Hand collision is disabled.

Figure 3.1: An image showing how information is displayed for the participant. Color coding with green and red is used for the text to highlight when hand collision/grabbing is disabled or not.

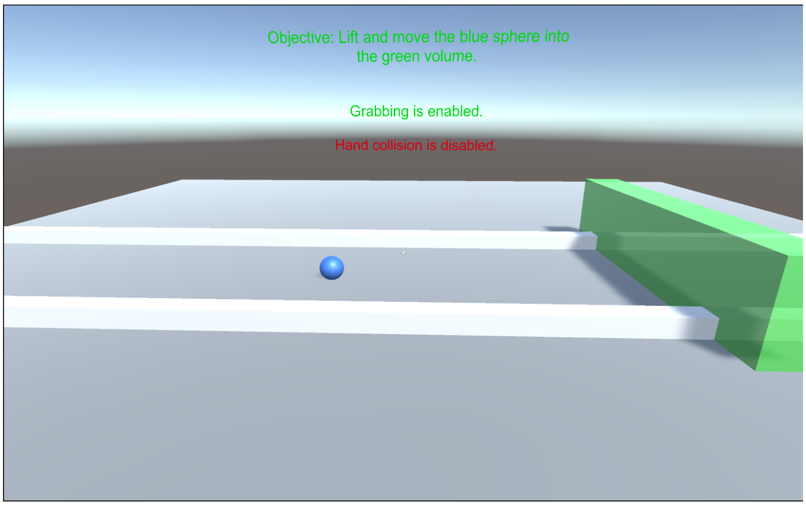
There are primarily two types of tasks the participants have to complete. In this study, these tasks are referred to as the "sphere task" and the "cube task". However, before the task can start the participants are asked to read and sign a consent form and perform a demographical questionnaire. In this questionnaire, the participant provides information such as age, gender, and experience with VR/MR.

3.3.1 Sphere Task

The first task performed is the sphere task. This task requires the participant to move a sphere from its starting position into a goal visualized as a green volume. The setup for this task can be seen in Figure 3.2. The task has two variations, one where the player moves the sphere by grabbing it with no collision around the hands, and moves it into the goal. In the second part, the grabbing is turned off and the participant has to use the collision to push the sphere into the goal. After the two parts are done the participant will have to answer a questionnaire with the following questions:

- Did you prefer to lift or push the sphere? (three options: lifting, pushing, indifferent).
- How realistic did the lifting feel? (five options: strongly disagree, disagree, indifferent, agree, strongly agree).
- How realistic did the pushing feel? (five options: strongly disagree, disagree, indifferent, agree, strongly agree).

The order of having to grab and lift the sphere or push the sphere into the goal does not always occur in that order. Counterbalancing is utilized to prevent any sort of primacy or recency bias to affect the results as suggested by Price [27]. Overall, the point of this task is to introduce the participants to the basic features of the system and to get information about the UX of the hand collision in isolation.



Objective: Lift and move the blue sphere into the green volume.

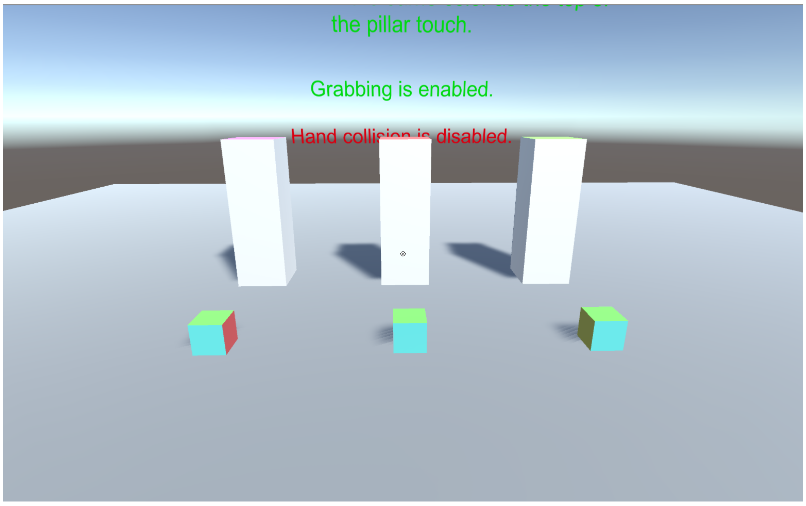
Grabbing is enabled.

Hand collision is disabled.

Figure 3.2: An image showing the setup for the sphere task.

3.3.2 Cube Task

The second task the participant needs to perform is the cube task. For this task grabbing is always enabled but hand collision is toggled on and off. This task presents three cubes and three pillars in front of the participant who needs to place one cube on each pillar. See Figure 3.3 for the cube task setup. However, each pillar has a specific color on the top, and the cubes have one color on each side. The cubes have to be placed on the same colored side as the top of the pillar facing each other. If a cube is correctly placed on a pillar it will light up in the color green. See Figure 3.4 for what an activated pillar looks like, showing a correctly placed cube. Once all pillars are activated, the participant has to then remove the blocks from the pillars. The process of placing the cubes on the pillars and knocking them down is done with grabbing and then again with grabbing and collision. This is done twice, once with medium-sized pillars/cubes and another time with a smaller set. See Figure 3.5 for size comparison. In total, the task is done four times, medium and small-sized pillars/cubes without hand collision and then again with hand collision. The varying size is used to test the participant's accuracy, specifically if they can easily move and manipulate small objects. The ordering of the small and medium sets of cubes/pillars as well as with/without hand collision also utilizes counterbalancing and might be executed in different orders.



the pillar touch.

Grabbing is enabled.

Hand collision is disabled.

Figure 3.3: An image showing the setup for the cube task. The cubes in front of the pillars needs to be placed on the top of the pillars with the correct side facing down. The cube's bottom facing side needs to match the color of the pillar's top facing side.

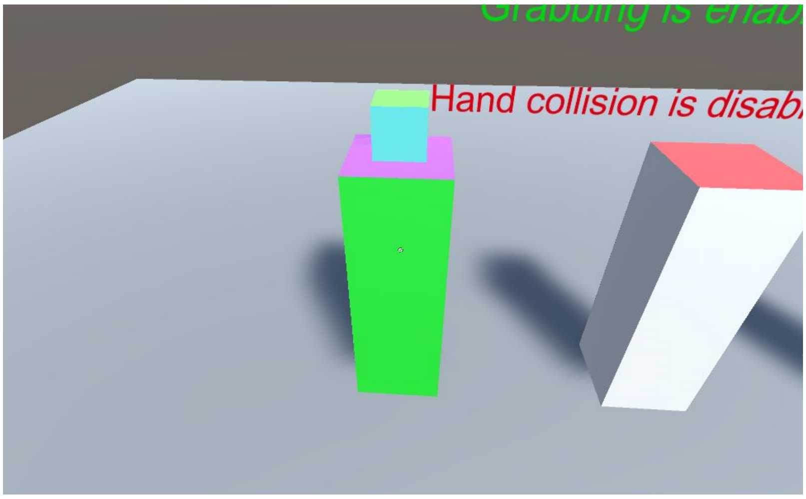


Figure 3.4: An image showing an activated pillar with a cube successfully placed with its purple side against the pillar's purple top side. Once all pillars have been activated they also need to be deactivated again to complete the cube task.

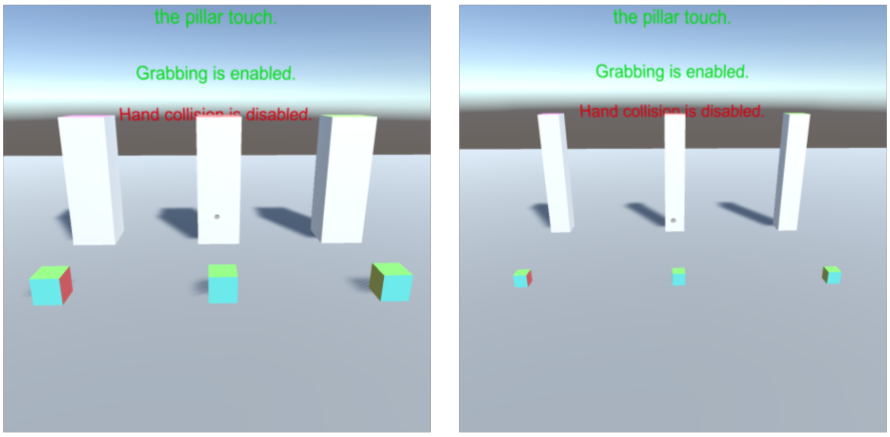


Figure 3.5: An image showing the size difference between the medium and small-sized cubes.

Once the participants had completed both a medium and small-sized set of pillars/cubes, they also filled out a questionnaire. This questionnaire consists mainly of the SUS questions (see section 3.2.1 for the full list of the SUS questions) with two additional questions. The additional questions added are (answered on a five point scale with the options strongly disagree, disagree, indifferent, agree, strongly agree):

- I think the system worked as I expected.
- I think the system felt realistic to use.

The additional questions are added since the SUS questions do not cover topics like intuition and realism which these additional questions are meant to cover.

3.4 Motivations and Alternate Experiment Designs

This experiment setup described in Section 3.3 was not the only one considered. Initially, the idea was to provide the participants with digital building blocks that could attach to each other, similarly concept to LEGO building bricks. The participants would receive instructions on how to assemble the block into a specific figure. This would require the participants to interact and manipulate the objects both through the use of translation and rotation in order to assemble them into the final figure. This idea did, however, ended up being rejected mainly due to two reasons. One reason was that there was no clear simple way of communicating how to assemble the blocks. Providing the participants with clear and easy instructions is very important to keep the experiment open for as a large pool of people as possible. This alternate experiment implementation was also more technically challenging to implement but there was no clear benefit with it over the current one being used. Both methods create scenarios for the participants to interact with digital objects in ways that require the participants to move and rotate the objects.

3.5 Participants and Demographic

As this thesis uses a user study, participants are required. The target group for participants is broad for this study the focus is on realism and intuitiveness. This means that the tasks should be accessible for most individuals regardless of age or experience. The exception is people under the age of 18 can not participate due to ethical reasons regarding user studies with minors. Other groups that can not participate in the study is people who do not have full functionality with their hands since the tasks require the utilization of the user's hands. There is also some color coding used in the task making it unsuitable for people with color blindness. Furthermore, the participants should not be sensitive to motion sickness or have epilepsy. The final number of participants were 12, 8 male and 4 female, between the age of 21-26.

3.5.1 Recruitment

The recruitment of participants for the user study was mainly done through the help of posters. These posters were placed on several billboards around Blekinge Institute of Technology (BTH). Digital resources like social media were also used for the recruitment. In this case, BTH's Discord server for Master of Science in Game and Software Engineering was used to post messages about the opportunity to partake in the user study.

3.5.2 Ethical Analysis

This thesis experiment requires testing with people to evaluate the UX with different direct hand manipulation methods. Due to the nature of conducting tests with people, an analysis has been made to determine potential ethical issues regarding health and user data. This is done to ensure that no unnecessary risks are taken and the risks that are involved are proportionally appropriate for the potential benefits.

One known health risk in VR is motion sickness which is generally perceived to be due to the brain receiving conflicting information from visuals and the inner ear [29]. However, because the user study takes place in MR the visuals should align better with the motion and orientation of the user, reducing the risk of motion sickness. Furthermore, MR does also not obstruct the visuals as much as VR which means the risk of users tripping or bumping into their surroundings should be reduced. Overall the risk of the user study is perceived to be less dangerous than similar tests conducted in VR. In terms of user data, all data collected will not contain identifiable traits and will not be traceable to any of the participants. The duration of the test will also be between 20-30 minutes to not be too long and cause fatigue. All participants will be 18 years or older and will read and sign a consent form to participate. Lastly, approval from an ethics board is not needed since the user study is not aimed to physically or psychologically affect the participants, as well as no sensitive data will be collected. However, advisory feedback has still been sought from the Ethical Advisory Board in South East Sweden. The received feedback did not indicate any major ethical issues with the user study.

3.6 Experiment Overview and Summary

In summary; the participant will read the information letter as well as read/sign a consent form and provide relevant demographic information about themselves. Once that is done, the participants will have to complete the sphere task by only grabbing and then again with only hand collision. The order for the method of moving the sphere is counterbalanced. When the sphere task is done, the participant will fill out a questionnaire about the experience. Next up is the cube task. This task is done once without hand collision on a medium-sized set of pillars/cubes and then again on a small-sized set. After that, the participant will fill out another questionnaire about the experience. Finally, the cube task is done again with a medium and small-sized set but this time with collision, and a final questionnaire is filled out. For the cube tasks, the size of the sets is counterbalanced as well as the hand collision or no hand collision being used. For a graphical overview, see Figure 3.6.

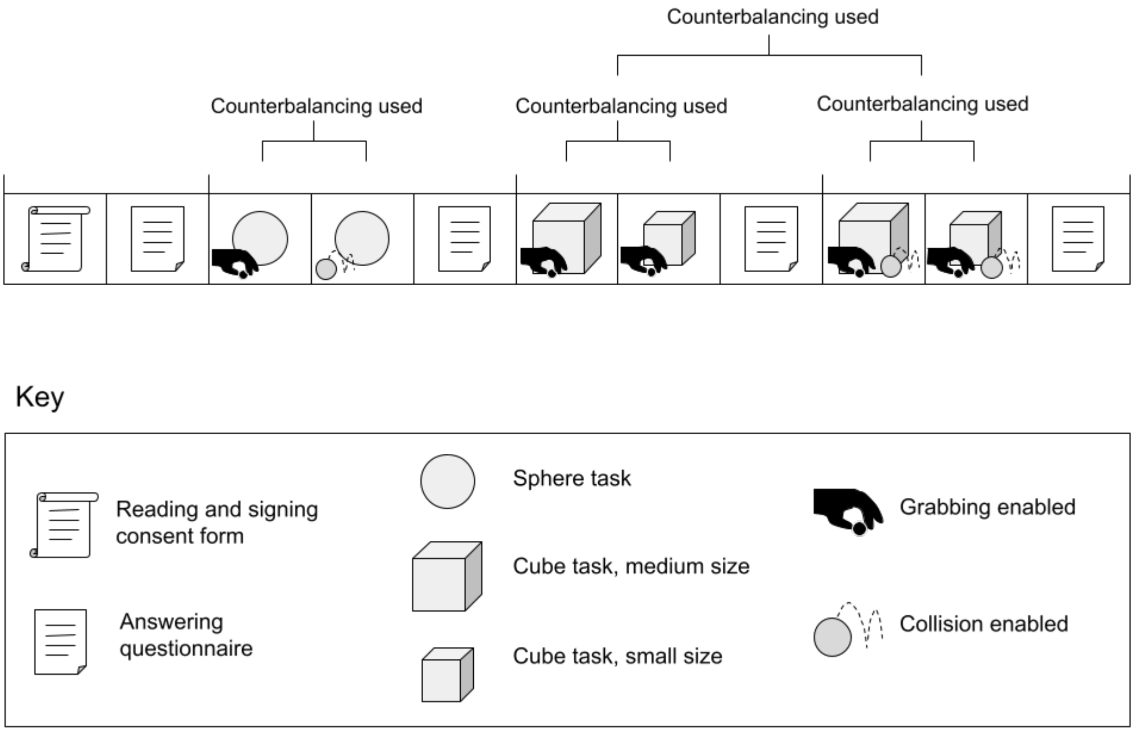


Figure 3.6: An image showing an overview of the experiment.

Chapter 4

Results and Analysis

This chapter will cover the results and analysis of the data collected throughout the study. This is both the objective performance data collected from the participants while completing their tasks and the subjective data based on the participants' experience collected through questionnaires. The section Results 4.1 will present the collected data through the use of various graphs. It is also divided into two subsections, one covering the objective performance data and the other the subjective questionnaire data. The section Analysis 4.2 covers the summarizing of the data, such as statistical tests and alike. Similarly, the analysis section is also divided up into two subsections. One covers the objective performance-based data and the other subsection focuses on the subjective questionnaire data.

4.1 Results

The data collected in this study is presented in this section and is shown in the graphs below. The times presented in the tables are measured from the first contact with the objects. For example, in the cube task, the time does not start to count until the user has touched one of the cubes. This is done to give the participants time to read and understand the task before attempting it. The tasks also have a time cap to prevent them from spending too long. This time cap is 180 seconds and if the task is not completed within this time the experiment will simply move on to the next step. Some tasks were also skipped due to bugs occurring and invalidating the data or the participant needing to leave. This is why medium and small tasks have different counts of participants.

4.1.1 Task Performance Results

This section presents the objective performance measurements collected throughout the tasks, such as time and number of times the participant grabbed an object. This is presented with multiple different graphs showing the results of the performance for both the cube and sphere task. The graphs will also show if the data holds normality. This means that the p-value calculated from the Shapiro-Wilks test was greater than the significance level of 0,05 which means that it holds normality. And if the p-value is lower than the significance level, it does not hold normality. The motivation for using the Shapiro-Wilks test can be seen in Section 3.2.2. The graphs will also show what the mean value is rounded off to two decimal points. The mean refers specifically to the arithmetic mean.

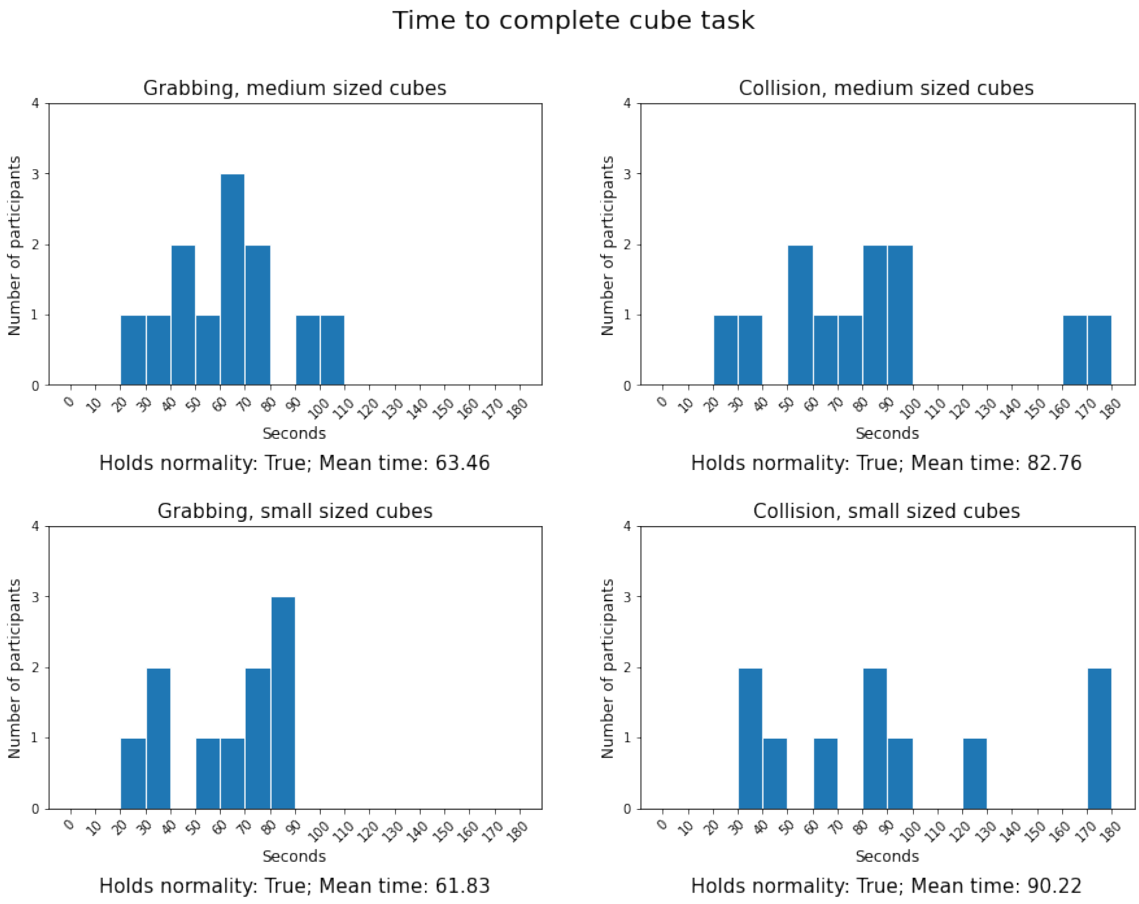


Figure 4.1: This figure shows four histograms presenting the data for how long it took the participants to complete the cube task. The left column of graphs presents the data for only having grabbing, while the right graphs have hand collision enabled.

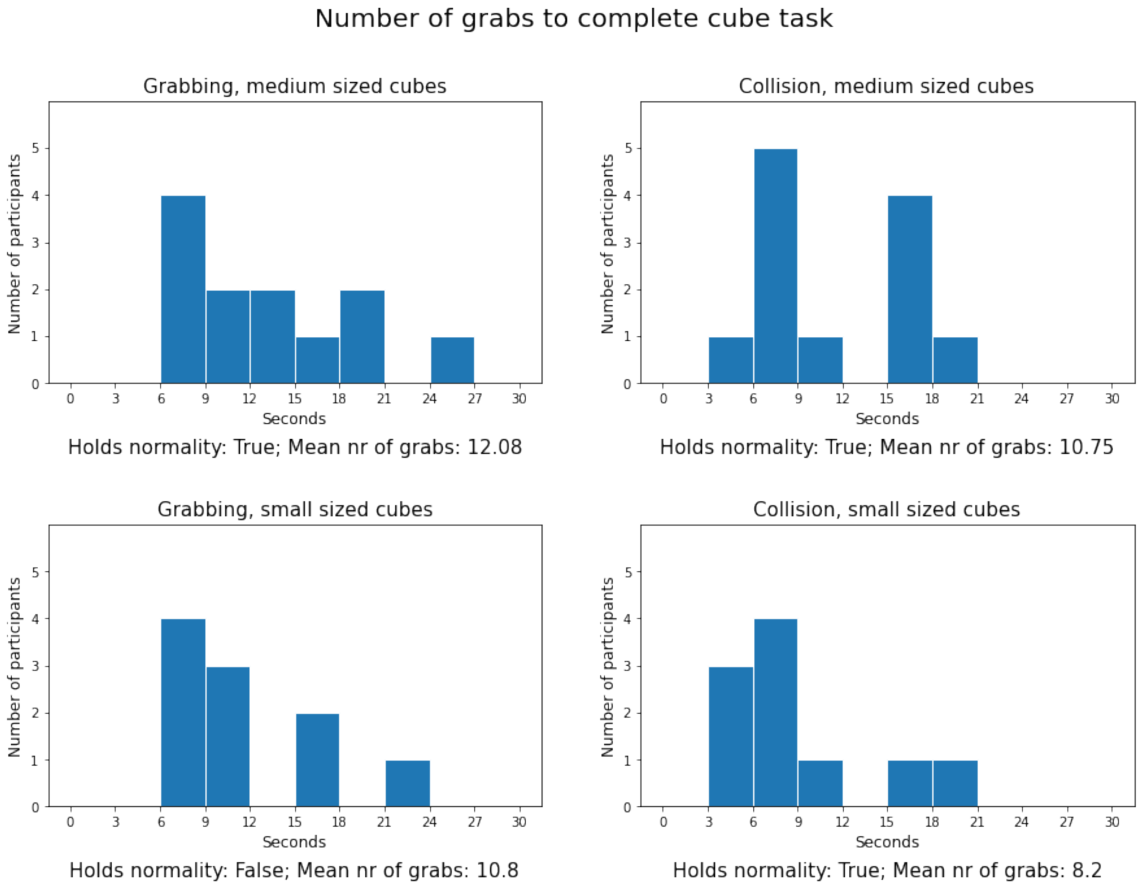


Figure 4.2: This figure shows four histograms presenting the data for how many grabs it took for the participants to complete the cube task. The left column of graphs presents the data for only having grabbing, while the right graphs have hand collision enabled.

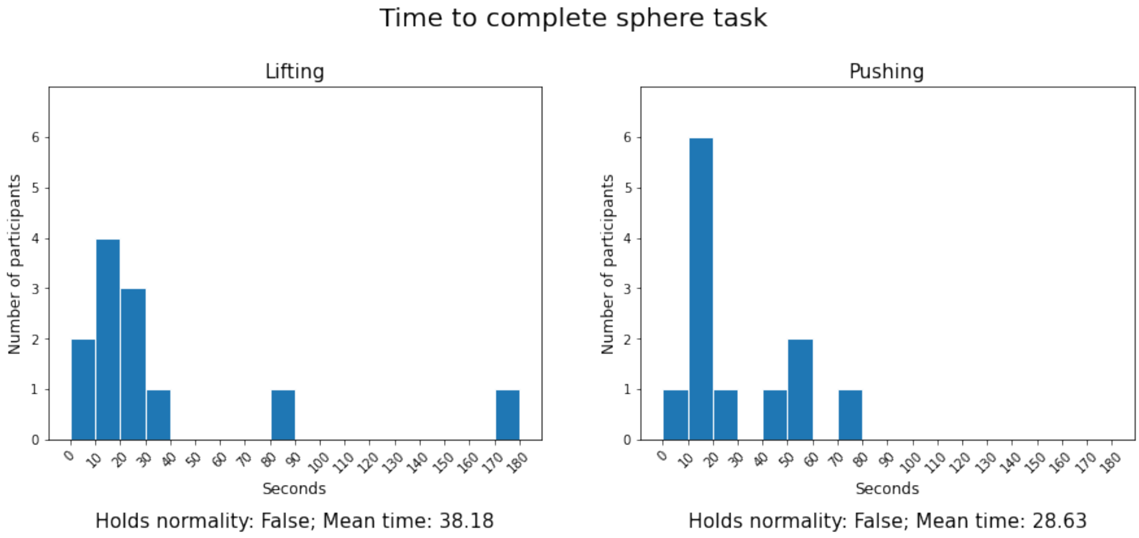


Figure 4.3: This figure shows two histograms presenting the data for how long it took the participants to complete the sphere task.

4.1.2 Questionnaire Results

This section presents the subjective data collected through the questionnaires, such as the preferences and SUS questions as seen in Section 3.2.1. The results from the questionnaires the participants answered in the study are presented in the form of different graphs. The results for the sphere task questionnaire can be seen in Figures 4.4 and 4.5. The results for the cube task can be seen in Figures 4.6 through 4.17 where the questions were asked twice, once for hand collision and once without, and the results are presented next to each other in each figure.

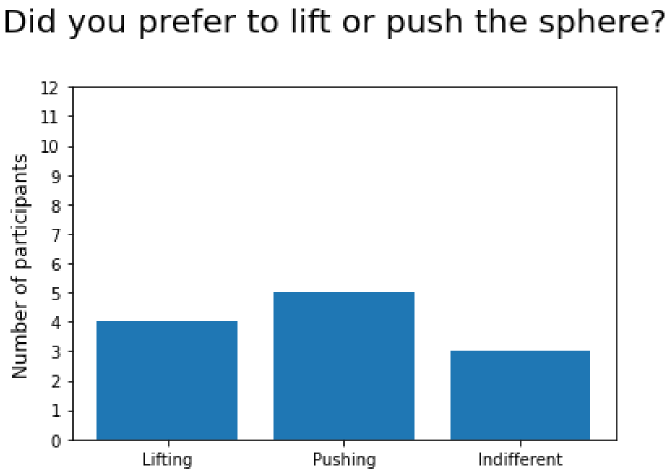


Figure 4.4: A figure showing a bar graph that presents the result from one of the questions from the sphere task questionnaire. The question is regarding which method of moving the sphere the participant preferred.

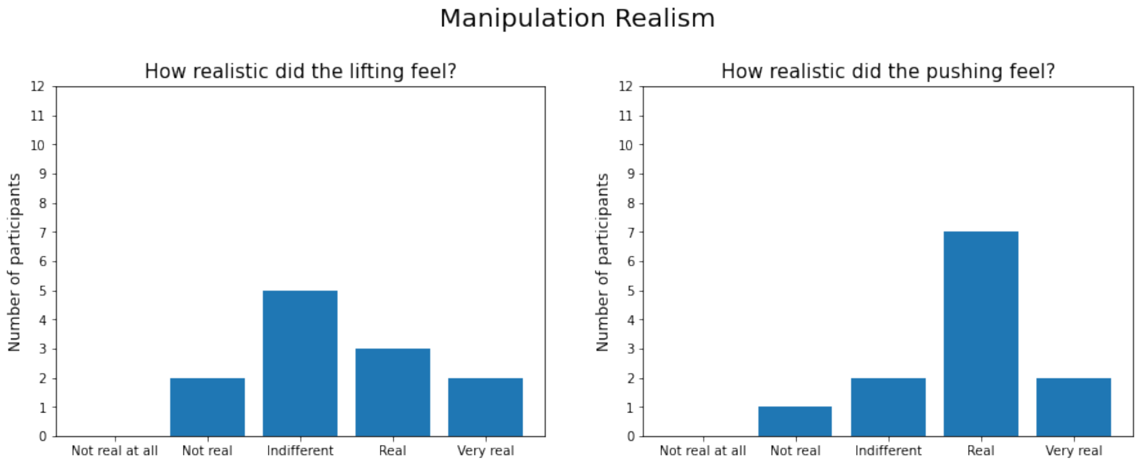


Figure 4.5: A figure showing two bar graphs that present the result from two of the questions from the sphere task questionnaire. These two questions are regarding how realistic the different manipulation methods felt.

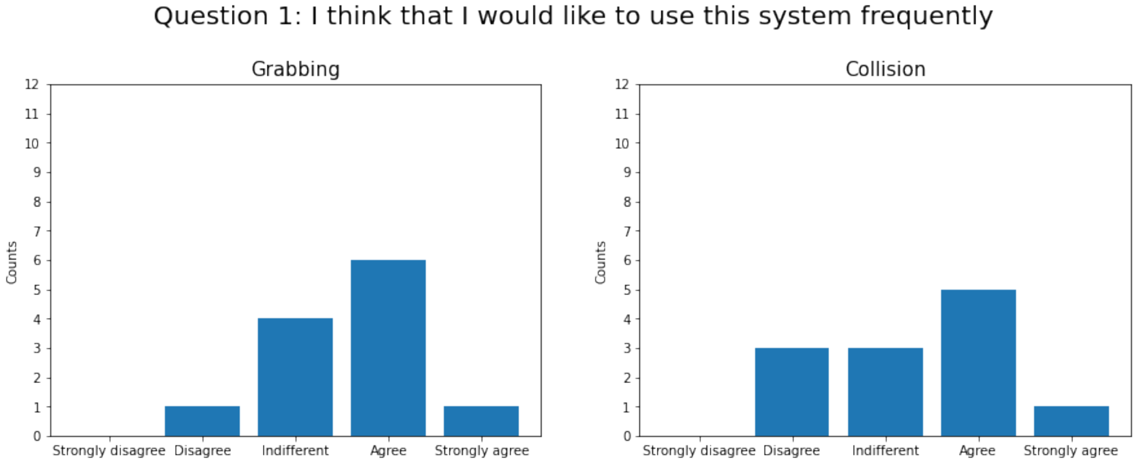


Figure 4.6: Two bar graphs showing the result for the 1st SUS question.

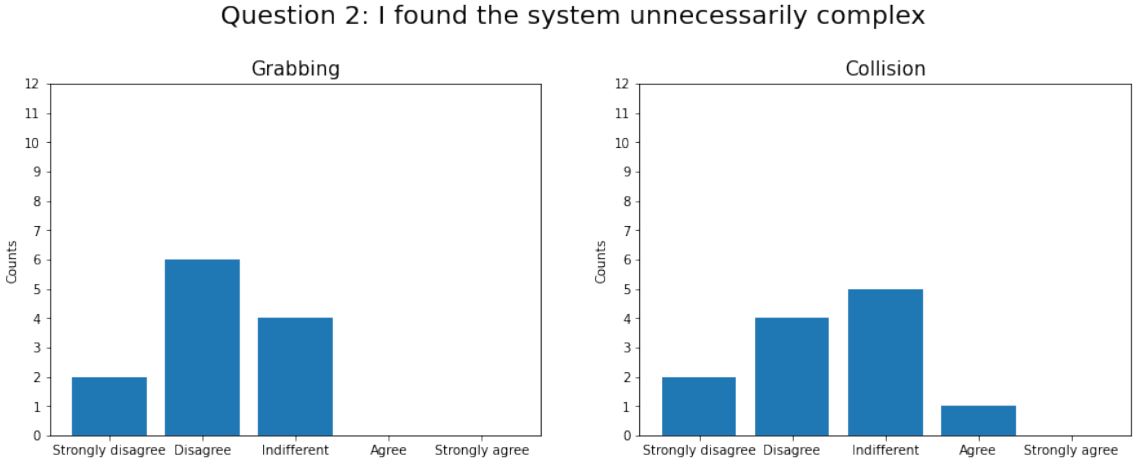


Figure 4.7: Two bar graphs showing the result for the 2nd SUS question.

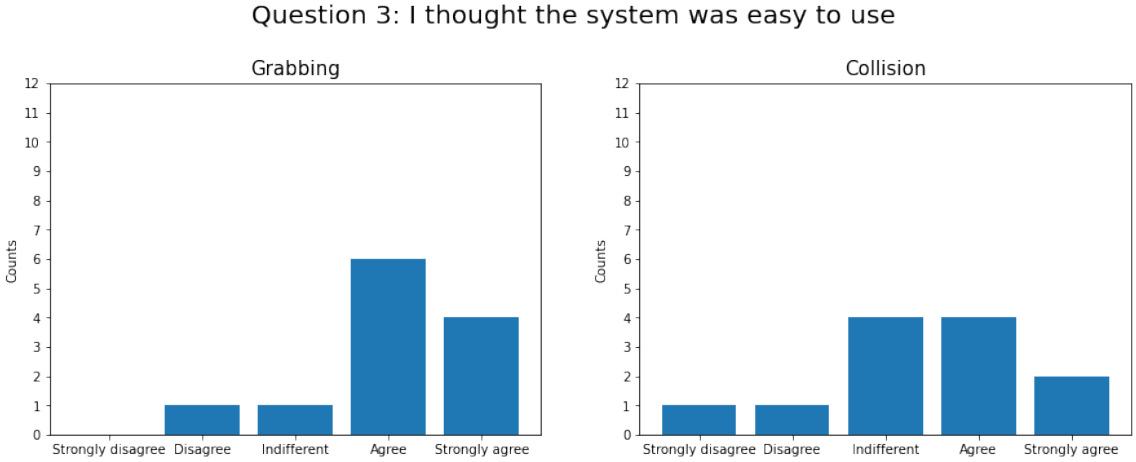


Figure 4.8: Two bar graphs showing the result for the 3rd SUS question.

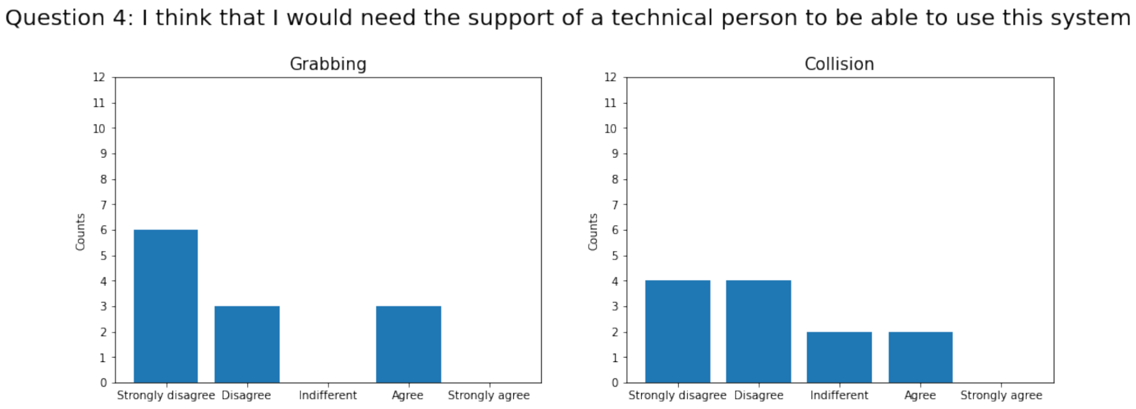


Figure 4.9: Two bar graphs showing the result for the 4th SUS question.

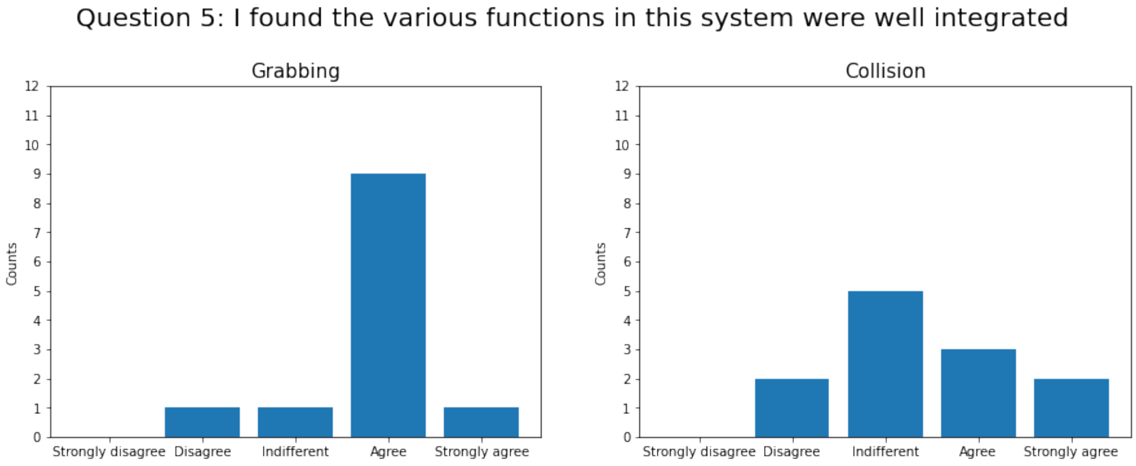


Figure 4.10: Two bar graphs showing the result for the 5th SUS question.

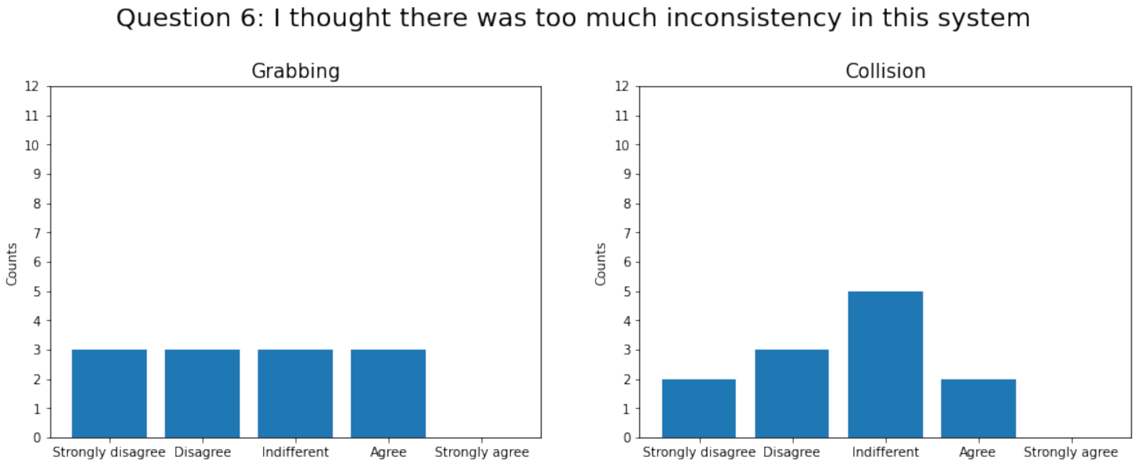


Figure 4.11: Two bar graphs showing the result for the 6th SUS question.

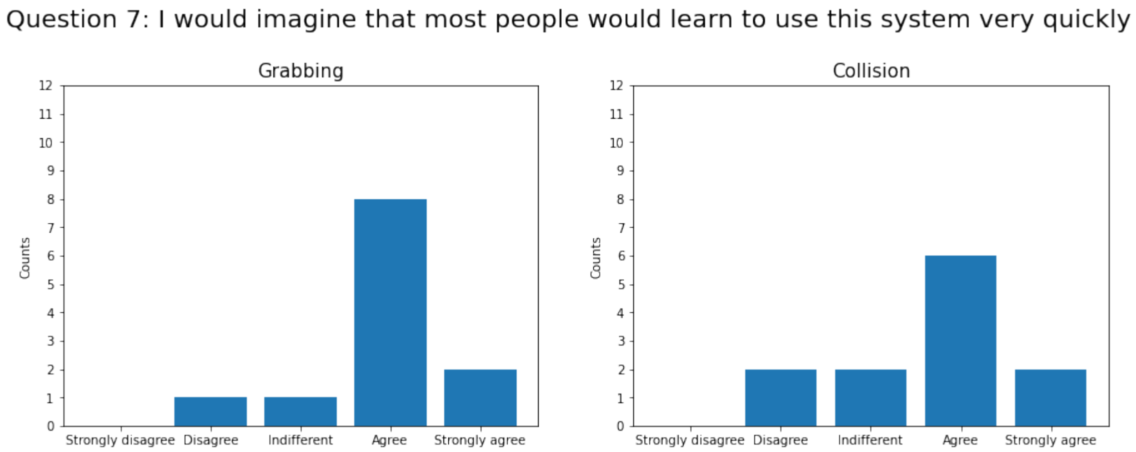


Figure 4.12: Two bar graphs showing the result for the 7th SUS question.

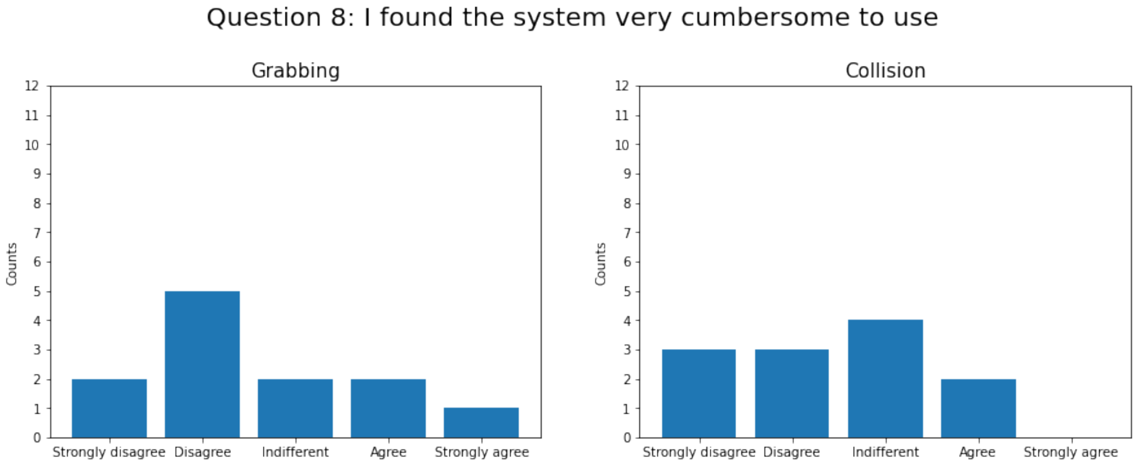


Figure 4.13: Two bar graphs showing the result for the 8th SUS question.

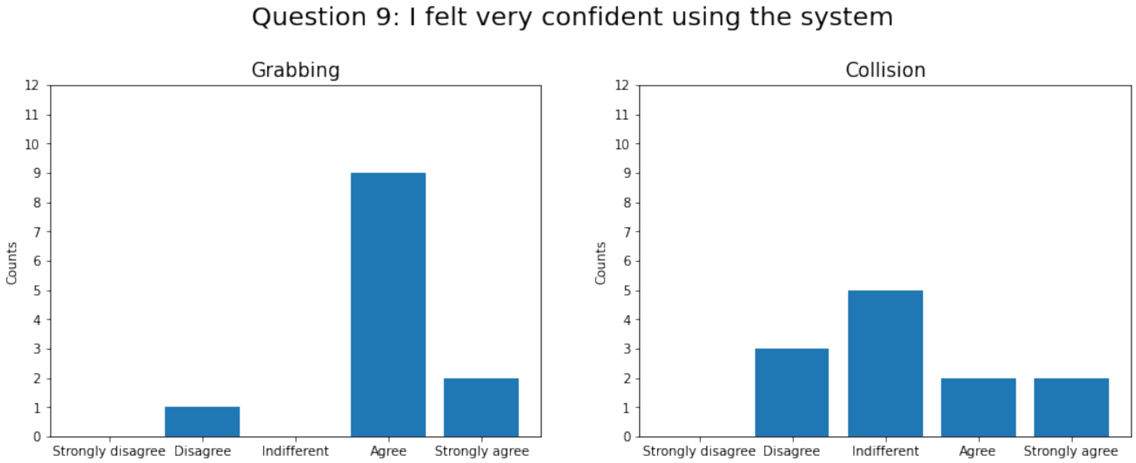


Figure 4.14: Two bar graphs showing the result for the 9th SUS question.

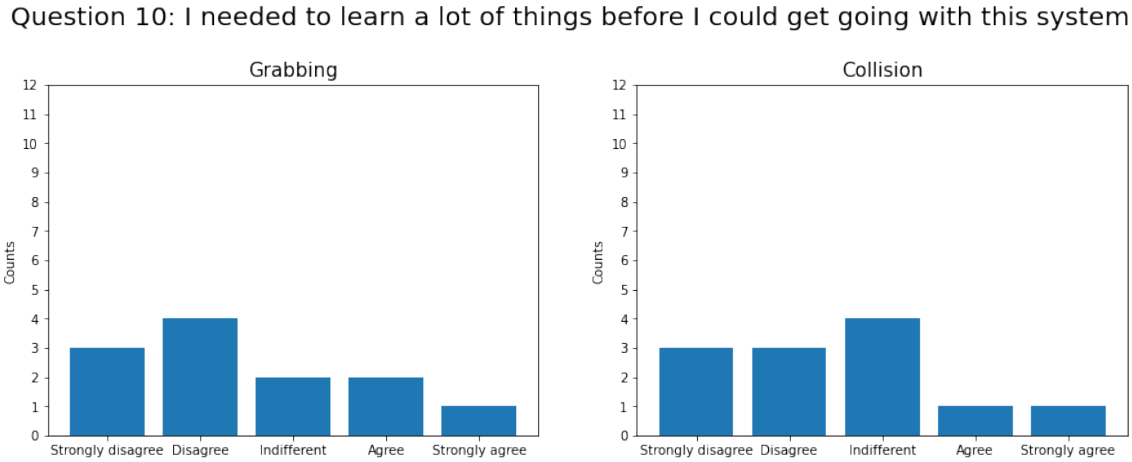


Figure 4.15: Two bar graphs showing the result for the 10th SUS question.

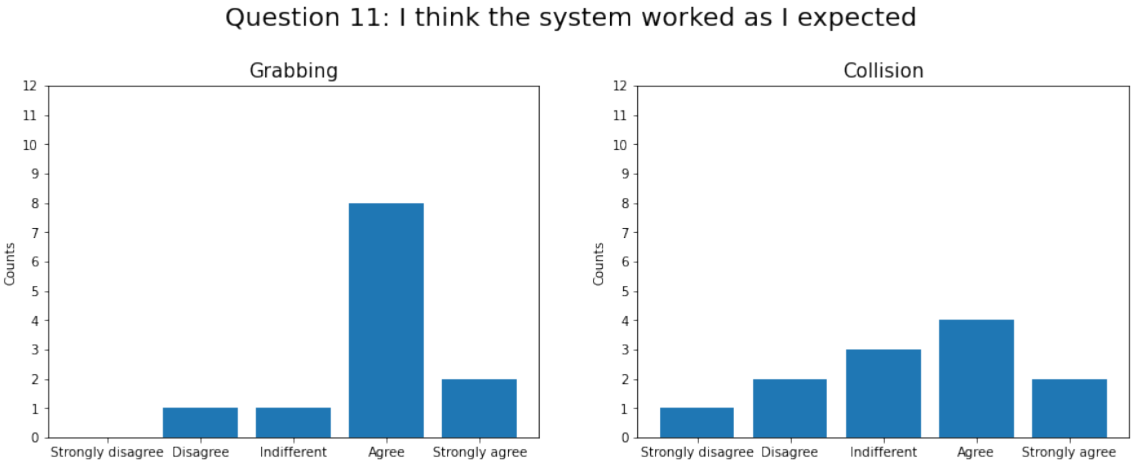


Figure 4.16: Two bar graphs showing the result for one of the additional questions added beyond the SUS questions.

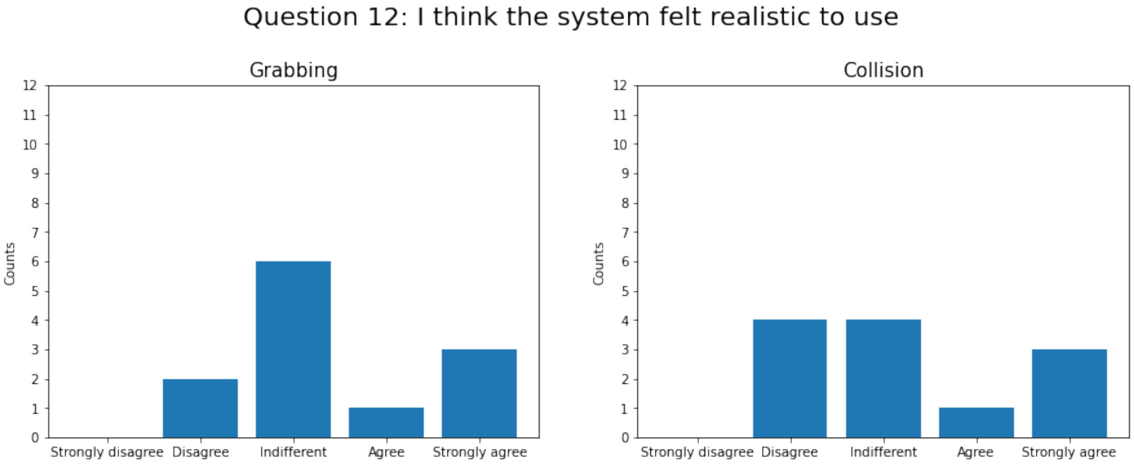


Figure 4.17: Two bar graphs showing the result for one of the additional questions added beyond the SUS questions.

4.2 Analysis

This part will analyze and process the data presented in Section 4.1. The performance-based data will mainly be analyzed and processed through the use of various statistical tests as motivated in Section 3.2.2. The subjective data is mainly analyzed with the SUS method as presented in Section 3.2.1 with a few exceptions to the additional questions that were added that will be analyzed with statistical methods. Some of the data will also be analyzed with the use of a graph, since there are too many buckets and few data points to get an accurate estimate. All the statistical tests performed in this section used a significance level of 0,05 as motivated in Section 3.2.2.

4.2.1 Task Performance Analysis

This section will cover the analysis of the objective performance related data collected during the tasks. The analysis is done through statistical tests. In Table 4.1 the statistical test for detecting a statistically significant performance difference between using hand collision and no hand collision can be seen. The performance metric in this case is time to complete the cube task. This was done with a paired t-test for both the small- and medium-sized cubes. The results shows that the null hypotheses, stating that there is no difference, is rejected, concluding that there is a noticeable performance difference in terms of time.

Table 4.1: A table showing two statistical tests testing if there is a statistically significant performance difference in time performance for medium and small-sized cubes in the cube task.

Cube size:	Medium	Small
Statistical test:	Paired t-test	Paired t-test
h_0:	There is no difference with and without hand collision in terms of time.	There is no difference with and without hand collision in terms of time.
P-value:	0.043	0.032
Result:	h_0 Rejected	h_0 Rejected

Regarding the number of grabs needed to complete the cube task, the statistical test can be seen in Table 4.2. For this test, the paired t-test and Wilcoxon test were used for medium- and small-sized cubes, respectively. Wilcoxon test was used for the small-sized cubes since they did not follow a normal distribution. The test found that the null hypotheses for medium-sized cubes holds true but is rejected for the small cubes.

Table 4.2: A table showing two statistical tests testing if there is a statistically significant performance difference in the number of grabs for medium and small sized cubes in the cube task.

Cube size:	Medium	Small
Statistical test:	Paired t-test	Wilcoxon test
h_0:	There is no difference with and without hand collision in terms of the number of grabs.	There is no difference with and without hand collision in terms of the number of grabs.
P-value:	0.079	0.049
Result:	h_0 Holds true	h_0 Rejected

As for the sphere task, it can be seen in Table 4.3 where the Wilcoxon test was used. This was done to calculate if there was a statistically significant difference in the time it took to complete the sphere task when only lifting and only pushing the sphere. Wilcoxon was used since the data did not follow a normal distribution. The test found that the null hypothesis holds true and that there does not seem to be a significant performance difference in time for pushing or lifting the sphere.

Table 4.3: A table showing a statistical test testing if there is a statistically significant performance difference between lifting or pushing the sphere in the sphere task.

Statistical test:	Wilcoxon test
h_0:	There is no difference between lifting and pushing in terms of time.
P-value:	0.470
Result:	h_0 Holds true

4.2.2 Questionnaire Analysis

This section will cover the analysis of the subjective data collected through the questionnaires regarding the participants' experience using the system. The data will be analyzed with the SUS method as well as statistical tests and with the help of a graph.

The SUS score was calculated as described in Section 3.2.1. The result was a score of 69,2 for no hand collision and 62,5 for hand collision. As for the additional questions added about realism and intuitiveness, a statistical test was performed as seen in Table 4.4. The Mann-Whitney U-test was used for both questions since it performs well with comparing Likert data [9]. The result shows no statistically significant difference in intuitiveness or experience for hand collision and no hand collision, and the null hypotheses for both questions holds true.

As for the questions asked in the questionnaire for the sphere task, there were two questions asked about how realistic the lifting and pushing felt. These two questions were compared with the Mann-Whitney U-test since they are both Likert data as seen in Table 4.5. The result shows that there was no statistically important difference in terms of realism between lifting and pushing the sphere. The null hypothesis therefore holds true.

Table 4.4: A table showing two statistical tests testing if there is a statistically significant experienced difference in intuitiveness and experience between hand collision and no hand collision in the cube task.

Question:	I think the system worked as I expected.	I think the system felt realistic to use.
Statistical test:	Mann-Whitney U-test	Mann-Whitney U-test
h_0:	There is no experienced difference in intuitiveness when using hand collision and no hand collision.	There is no experienced difference in realism when using hand collision and no hand collision
P-value:	0.215	0.648
Result:	h_0 Holds true	h_0 Holds true

Table 4.5: A table showing a statistical test testing if there is a statistically significant experienced difference in realism between lifting or pushing the sphere in the sphere task.

Statistical test:	Mann-Whitney U-test
h_0:	There is no difference in preference between lifting or pushing in terms of realism.
P-value:	0.247
Result:	h_0 Holds true

Finally, there is the result for the amount of previous experience the participant has with MR/VR as asked through the questionnaire. Since the data have many categories and relatively few data points, it might be more effective to look for correlations through a more visual method. This was done with graphs as seen in Figure 4.18. The graphs in this figure shows participants experience with MR/VR as color-coded points, with their performance plotted in the y-axis.

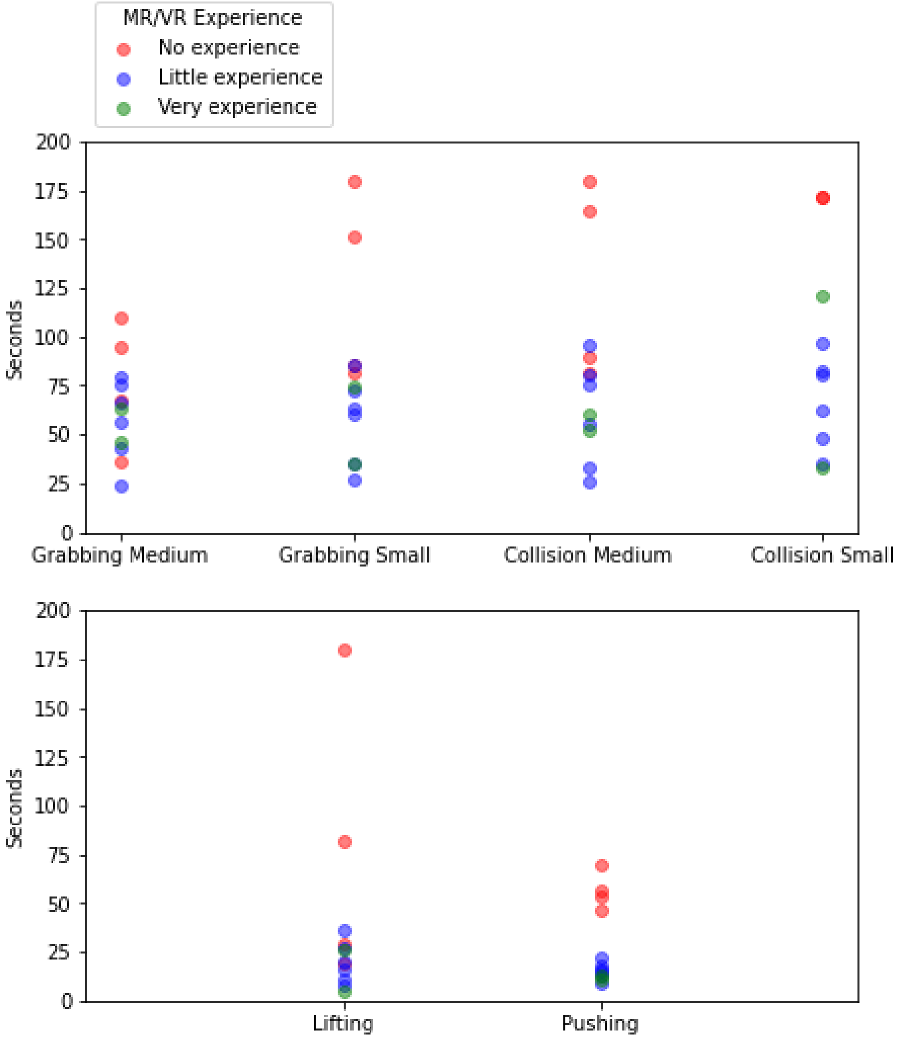


Figure 4.18: Two graphs showing the performance data (time spent until the task was completed) relative to experience in MR/VR. The upper graph shows performance in the cube task and the lower graph shows performance in the sphere task.

Chapter 5

Discussion

This chapter will cover the discussion surrounding the data presented in the Chapter 4 Results and Analysis. This means trying to understand and theorize why the results are the way they are. The different findings are then summarized and compared to previous research. Finally, any issues will be identified and potential solutions proposed.

5.1 Results and Analysis Discussion

This section will cover the discussion of the data presented and analyzed in the previous chapter, Chapter 4. This section will follow a similar structure of Chapter 4, being divided into two subsections. One covering the objective performance based data and the other covers the subjective questionnaire data.

5.1.1 Task Performance Discussion

This section will cover the discussion surrounding the results presented in Section 4.1.1 and the analysis performed in Section 4.2.1. In other words, this section covers the discussion about the objective performance based data collected and analyzed.

For the time required to complete the cube task, the result of the analyses can be seen in Table 4.1. The statistical test performed found that there is a statistically significant performance difference for both cube sizes. Looking at the graphs from Figure 4.1, it can be seen that the average time to complete the task is greater for hand collision, for both cube sizes. With these two pieces of information, it is reasonable to conclude that the hand manipulation without collision performed better in terms of time performance.

Another performance metric measured in the cube task was the number of grabs used to complete the task. The result of the grab count analysis can be seen in Table 4.2. The results of the statistical test presented in the table shows no statistically significant difference for the medium-sized cubes, but does however find a significant difference for the small-sized cubes. Looking at the graphs in Figure 4.2, it can be seen that with hand collision, the average grab count is lower. With these two pieces of information, it can at least be reasonably concluded that hand collision performed better for the small cubes with a lower grab count. Although the medium-sized cubes did not quite reach a low enough p-value for it to go below the significance level, it did get close. This indicates that there is a good chance that hand collision also performs better for the medium-sized cubes in terms of grab count. The reasoning behind why the performance difference was greater for the small-sized cubes might be because they are harder to interact with. Being smaller makes the cubes more difficult to move and manipulate, which might accentuate the more effective option. The reason why collision performed better in this regard is probably because of the second stage of the cube task where the participants have to remove the cubes from the pillars. While the hand method can only utilize grabbing and had to grab to remove the cubes, the hand collision has another alternative to knock away the cubes, reducing the number of grabs needed. This was a method observed to be performed by the participants. Utilizing the collision to knock down the cubes from the pillars was also intuitively done by the participants and were never told to do this. This potently shows some degree of added intuitiveness and realism with this method. Since the participants were able to recognize real-life applications of their hands and relate that use-case to this scenario.

Finally, there is the collected data from the sphere task. The metric collected here is time to complete the sphere task. The statistical test for this data can be seen in Table 4.3. The result shows no statistically significant difference in time between only pushing the sphere or only grabbing the sphere to complete the task. As seen in the graphs in Figure 4.3, the average time was slightly faster for pushing the sphere, however, the average time difference seems quite minimal. Overall, there does not seem to be much of a difference between the alternatives of moving the sphere. It is also worth noting that it might not be fair to compare these two ways of moving the sphere with another. This is because they are not quite the same thing and may not be directly comparable. The sphere task was mostly there to give the participants some initial intuition of the system's features and collect the participants' subjective opinions on the features.

5.1.2 Questionnaire Discussion

This section covers the discussion for collected subjective data gathered through the questionnaire. This means both the results from Section 4.1.2 and the analysis from Section 4.2.2.

As presented in the beginning of Section 4.2.2, the calculated SUS scoring was 69,2 for no hand collision and 62,5 for hand collision. This means that the method with no hand collision received a good score, while hand collision received a moderately good score [2]. Looking through the results in Figures 4.6 to 4.15, the scores do seem quite similar for both hand collision and no hand collision. Out of the ten standard SUS questions, notable questions that seem to deviate the most from each other are Figures 4.10 and 4.14. Question 5 was about how well the features were integrated in the system. Here the hand collision seems to have performed slightly worse. This is most likely due to the kinematic bodies used for the physics collision inherently not reacting to the collisions themself, sometimes leading to unrealistic scenarios. Question 9 was regarding how confident the user felt while using the system. For this question, hand collision also performed worse. The reason for this is likely similar to the other one. That is, the kinematic bodies not reacting to physics leading to, for example, squishing a dynamic body between the fingers which generates a high force on the dynamic body flinging it away. Unpredictable behavior like this might confuse the user and lower their confidence using the system. While these could be the reasons for the resulting difference in questions 5 and 9, the difference might not be high enough and could also be the results of random chance.

For the additional questions added regarding intuitiveness and realism, the analysis can be seen in Table 4.4. The statistical tests did not find any statistically significant differences between having hand collision and no hand collision, both with regard to intuitiveness or realism. Looking at Figures 4.16 and 4.17 it seems that both methods have similar results but with the biggest discrepancy for question 11. This could be for the same reason before of kinematic bodies sometimes creating unnatural behavior, or simply due to random chance.

For the sphere task, the participants were asked how realistic they felt pushing and lifting the sphere. A statistical analysis can be seen in Table 4.5 to test if there is an experienced difference in realism between the two. The test did not find a statistically significant difference between the two methods in terms of realism. Looking at Figure 4.5 it does look like pushing the sphere was experienced to be slightly more realistic. This could be because pushing an object has less chance of pinching or crushing the dynamic object, avoiding the unrealistic physics behavior that otherwise can happen. Or it is simply the result of random chance. There was also one question regarding the participants' preferred way of moving the sphere. The result of this question can be seen in Figure 4.4. No statistical test was performed on this data, since it is quite clearly evenly distributed and no clear difference in preference is seen.

Another potentially interesting variable to look at is experience level in MR/VR environment the participants provided in the questionnaire in relation to performance. This can be seen visualized in Figure 4.18. Based on the visualization, it does seem to be a correlation of performance and having no previous experience in MR/VR. However, there is no clear indicator that it had any effect on performance between hand collision and no hand collision.

5.2 Findings and Previous Research

The hand collision did not improve time performance or UX according to the SUS scoring. This is probably related to the method used for creating the hand collision, which is the use of kinematic bodies that strictly follows the user's hands. This means that sometimes unrealistic scenarios occurred where a dynamic object sometimes got squished by the kinematic colliders, flinging the dynamic object away. While there was no statistically significant difference between with/without hand collision for intuitiveness and realism, this finding is consistent with what Nasim and Kim stated in their paper. That using kinetic bodies that strictly follows the user's hand could have issues with penetrating other object, causing instabilities in the simulation [19].

There is still some argument to be made about the intuitiveness and realism that the hand collision might add. This is because participants were observed to intuitively utilize the hand collision to complete the task without being instructed to do so. Showing a level of intuition and realism where real-life logic and experience it effectively utilized in this setting. This is also reflected in the grab count, with the number of grabs being lower for hand collision. While not much previous research seems to have put much focus on the UX relating to intuitiveness and realism, there is Boonbrahm and Kaewrat talk about their own experience in their paper [3]. While they only tested their hand collision method on themselves, they did find that the collision added a bit more of a natural feel [3]. With these pieces of information, there is an indication of more potential research to focus on what hand collision can add in terms of UX.

5.3 Identified Issues

The reason why hand collision generally performed worse is probably due to the collision being kinematic. This gives very little feedback to the user since kinematic bodies only exert force in one direction, only affecting other physics bodies while not being affected itself. This leads to some unrealistic and unintuitive scenarios, hence, most likely why the hand collision performed worse with the SUS score. One common issue was the objects easily got squished between different parts of the collision around the hand. This issue was most likely also exacerbated by their only being objects composed of rigid physics bodies that do not deform. For example, picking up an object with the index finger and thumb would require the user to maintain an exact and constant distance between the fingers. Since the fingers can not get pushed back and the object being picked up can not be deformed, a very high force would be generated. This force would often fling or glitch the object away/through the hands, creating confusing and unrealistic situations.

More generally, there is a problem of affects only acting in one direction in MR. While the user can easily interact with the digital world and affect it, the digital world can not easily affect the user back. This might ruin the illusion of the physical and digital world being one, breaking the immersion. This, however, is not to say that the digital world can not affect the physical at all. It can since it creates visuals that are displayed to the user, generate sounds, or create the sensation of touch with for example haptic feedback. It is however a challenge to connect the worlds in a seamless way, especially with physics, since it relies heavily on being able to push the user back when the user exerts a force on an object. As stated by Isaac Newton's third law of motion, for every action there is an equal and opposite reaction [21].

5.4 Improvements

There are a couple of potential ways to fix these issues with unrealistic collisions with the hands, as mentioned in Section 5.3. One way would be not using completely rigid physics bodies for the hand colliders and the generic objects. This would make the physics bodies slightly more realistic, allowing for some degree of deformation. This would overall require less precision from the user when lifting and picking up objects. The quite significant downside of this is that soft body simulations are significantly more expensive than rigid body simulations. This can especially be a problem if the MR/VR headset runs the software locally which already generally has limited hardware performance.

Another potential way of solving this issue could be to not have the hand collision follow the user's hand exactly. For example, the hand collision could be some form of dynamic body that allows for some reaction to other physics bodies. Meaning that the digital version of the hand can be pushed back and more naturally can interact with objects while still trying to follow the real hand as closely as possible. An example of this idea can be seen in Figure 5.1. A way to create this could be to make the collision bodies dynamic physics bodies instead of kinematic bodies. Furthermore, joints and constraints should also be used between the physics bodies, allowing them to stay connected, rotate relative to each other like real joints, and not separate during a collision. This allows a collider to affect connected colliders, creating a more realistic physical response to collisions. This is most likely a more expensive method, however, no significant performance issues were noticed for the kinematic implementation. This means that there might the dynamic implementation might be a viable option.

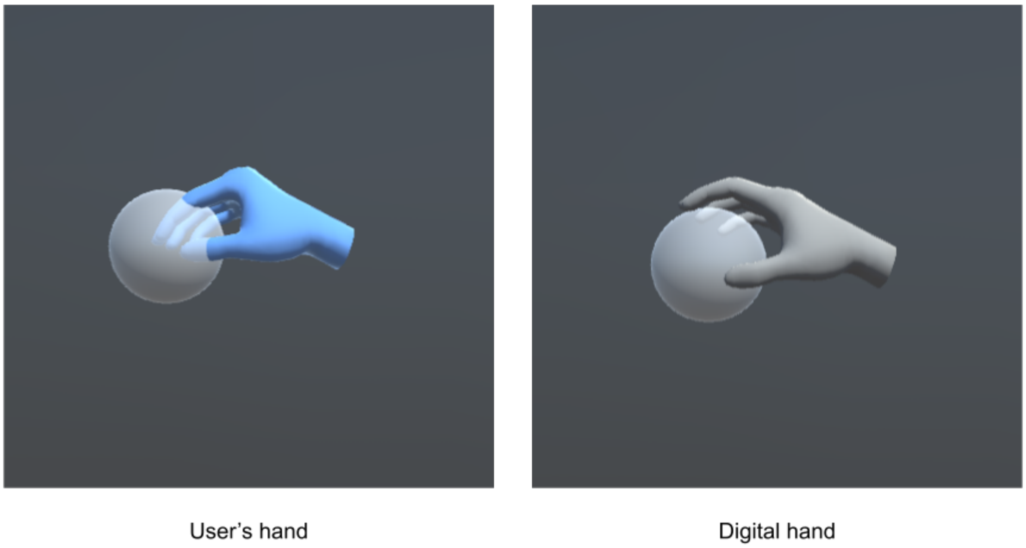


Figure 5.1: The blue hand on the left represents the user's hand trying to grab the digital sphere. The gray hand on the right shows the digital hand with the collision that tries to follow the real hand but collides with the sphere.

Chapter 6

Conclusions and Future Work

This chapter will cover a summary of the thesis. Providing a quick summary of the purpose, execution, and results. Furthermore, it will cover a conclusion that can be drawn based on the results, analysis, and discussion previously presented in the thesis. It will also list the answers to the previously stated research questions and the results for the objectives. In this case, it seems that the addition of collision did not have any major impact on performance or UX, but has potential for future research. Finally, a look into the future of what the next step could be to improve upon this research as well as the collision method used.

6.1 Thesis Summary

This work has performed a user study on the addition of collision around the user's hands in an MR environment. A relevant area for many fields but especially for education, training, [11], [12], [18], [32], and medicine [11]. This study hopes to gain insight into whether collision can provide better and more realistic MR experiences. The MR environment was created with the use of the MR headset HoloLens 2. The headset was used to track the user's hands and joint positions, which was used to fit kinematic colliders to the user's hands. This method was then tested with a group of 12 participants. They were tasked with completing a series of tasks with and without the hand collision, as well as answering a set of questionnaires about their user experiences. The collected data that is analyzed for this study is the objective performance-related data collected from the user when completing the task. This data is the time and number of grabs the participant used to complete the task. The other data was the subjective experience-based data that was collected through the answers the participants provided in the questionnaire. The data was analyzed through the use of various statistical tests and the SUS method. The results from these calculations show that the participants did not perform particularly better or had a strong preference for using the hand collision. However, the utilization by the participants of the collision showed some degree of potentially added intuition and realism.

6.2 Conclusion

Based on the statistical test in Table 4.1 and the data from Figure 4.1 show that using hand collision performed worse in terms of the time it took to complete the tasks. However, in terms of grab count, the hand collision performed better based on the statistical test seen in Table 4.2 and the data presented in Figure 4.2. Overall, based on the objectively collected data, the results are a bit mixed, with hand collision potentially performing a bit worse. This might be due to the lack of feedback creating inaccuracy, potentially due to the lack of tactile sensation, but also in terms of the lack of bi-directional physical interaction, e.g. the hands not being pushed back by the digital objects. This follows a similar line of thinking supported by other researchers [19].

As for the subjective data, the SUS scores, as seen at the beginning of Section 4.2.2, show that no hand collision received a good score while hand collision performed well but not quite as well. For the additional questions about realism and intuition, the statistical test can be seen in Table 4.4, which found no statistically significant difference. It is worth mentioning that participants were observed to utilize the collision to their advantage. The hand collision was used to quickly knock down the cubes from the pillars during the cube task, which was a strategy intuitively utilized and discovered. While the overall scores were quite similar for both methods, the non-hand collision seems to have performed better. This could also be due to previously mentioned issues with feedback, both in terms of physics and tactile sensation.

In conclusion, the collision method seems to have generally performed worse than the one without collision. However, the differences are quite small, and many of them could be due to randomness. Furthermore, while the collision performed worse in many areas, it did also show promise in some areas that warrant further research. Indicating that this method of creating collision around the user's hands might not be ideal. However, there are other ways that could work better [19]. With further improvements, such as increased accuracy and feedback, the collision might have potential.

6.2.1 Objectives Results

The thesis had a set of objectives it formulated that would help find a path to acquire the answers to the research questions. There where a total of three objectives listed as Ob1 to Ob3 in Section 1.4 called Objectives. The following listing is the results of the objectives:

- Ob1: An appropriate hand collision method was found and implemented. It was a method of using kinematic colliders that shaped and followed the user's hands which was a fitting method. This is because it is cheap in terms of performance which makes it ideal for mixed reality environment.

- Ob2: An appropriate experiment design and setup was created and used to collect user data. It puts the user in scenarios that had the user interact with digital objects through the use of their hands, both rotating and translating the object through space.

- Ob3: Appropriate methods for analyzing the data were found and utilized. This includes the SUS scoring method, paired t-test, Wilcoxon-test, and Mann-Whitney U-test were all used depending on what data was analyzed and what features the data had, like distribution, being related data, etc.

6.2.2 Research Question Answers

In Section 1.3 the research questions for this thesis were presented. The research questions presented were (RQ1) "What is the user experience with and without hand collision in mixed reality?" and (RQ2) "What is the user task performance with and without hand collision in mixed reality?". In this section, the research questions will be answered.

RQ1: The UX for the kinematic hand collision was moderately good according to the calculated SUS scoring. It performed, however, a bit worse the not using any hand collision in terms of UX. In terms of intuitiveness and realism, there were no statistically significant differences found for UX between hand collision and no hand collision. Overall, the UX seems to be acceptable for kinematic hand collision but not reliable enough to enhance the experience in a significant way in mixed reality.

RQ2: The kinematic hand collision performed worse in terms of time to complete the tasks relative to not having hand collision. It did however, in some instances utilize less number of grabs which indicates that it is, in some scenarios more accurate or more versatile. Overall, it does seem like performance is worse with kinematic hand collision in mixed reality due to the unstable physics interactions that can sometimes happen when the kinematic colliders penetrate other objects.

6.3 Future Work

For the future, there are many potential avenues to explore. The first is to increase the size of the study and have a larger number of participants. This study had 12 participants, which is a relatively small group. While some of the statistical tests still were able to draw some statistically significant results, more participants will only increase the accuracy.

Another area of interest is to test different collision methods for the hands. This study used kinematic bodies that strictly follow the user's hands without any sort of deviation from the hands. As suggested in Section 5.4, potentially better ways to do this would be to allow the rigid bodies to have some degree of deformation, essentially using soft bodies instead. This could help mitigate the issues of the kinematic hand colliders intersecting the dynamic objects. However, this could potentially be too performance-taxing and a better method might be to use colliders that do not strictly follow the user's hands. For example, one could use dynamic colliders that try to follow the user's hand as well as having the colliders connected with joints, allowing the colliders to push each other more naturally. This method seems like it has potential since Nasim and Kim used dynamic colliders for their experiment, which seems to have received positive results [19].

Something else that has seen some promising results is the addition of tactile sensation to MR, specifically haptic feedback [15], [35]. However, this could, in the future, be tested with hand collision to see if they have any additive effects. For example, haptic feedback could be used as a feedback system for how much pressure the user is applying to an object through the collision. Potentially giving the user a greater scene of control and immersion.

References

[1] S. A. Aaronson, A. Zable, J. V. O'Hara, and M. Lutz, "Reality check: Why the U.S. government should nurture XR development," XR Association, Tech. Rep., 11 2023, accessed: May. 1, 2024. [Online]. Available: https://xra.org/wp-content/uploads/2023/11/FINAL-XRA-REALITY-CHECK-White-Paper.pdf

[2] A. Bangor, P. Kortum, and J. Miller, "Determining what individual sus scores mean: Adding an adjective rating scale," *JUX: Journal of User Experience*, vol. 4, pp. 114–123, 2009. [Online]. Available: https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/

[3] P. Boonbrahm and C. Kaewrat, "Assembly of the virtual model with real hands using augmented reality technology," in *Virtual, Augmented and Mixed Reality. Designing and Developing Virtual and Augmented Environments*, R. Shumaker and S. Lackey, Eds. Cham: Springer International Publishing, 2014.

[4] J. Brooke, "SUS-a quick and dirty usability scale," in *Usability evaluation in industry*, 1st ed., P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, Eds. Gunpowder Square, London: Taylor & Francis, 1996, pp. 189–194.

[5] G. Buckingham, "Hand tracking for immersive virtual reality: Opportunities and challenges," *Frontiers in Virtual Reality*, vol. 2, Oct 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2021.728461

[6] E. Catto. Dynamics module. Accessed: 2023-10-23. [Online]. Available: https://box2d.org/documentation/md__d_1__git_hub_box2d_docs_dynamics.html

[7] ——. Overview. Accessed: 2023-10-23. [Online]. Available: https://box2d.org/documentation/index.html

[8] D. W. Cunningham and C. Wallraven, *Experimental Design: From User Studies to Psychophysics*, 1st ed. Massachusetts: Taylor & Francis, 2011, ch. 12.

[9] J. F. de Winter and D. Dodou, "Five-point likert items: t test versus mann-whitney-wilcoxon (addendum added october 2012)," *Practical Assessment, Research, and Evaluation*, vol. 15, 2010. [Online]. Available: https://scholarworks.umass.edu/pare/vol15/iss1/11

[10] Epic Games. Unreal engine documentation: Physics. Accessed: 2023-10-22. [Online]. Available: https://docs.unrealengine.com/4.27/en-US/InteractiveExperiences/Physics/

[11] H.-z. Hu, X.-b. Feng, Z.-w. Shao, M. Xie, S. Xu, X.-h. Wu, and Z.-w. Ye, "Application and prospect of mixed reality technology in medical field," *Current Medical Science*, vol. 39, no. 1, pp. 1–6, Feb 2019. [Online]. Available: https://doi.org/10.1007/s11596-019-1992-8

[12] C. Hughes, C. Stapleton, D. Hughes, and E. Smith, "Mixed reality in education, entertainment, and training," *IEEE Computer Graphics and Applications*, vol. 25, no. 6, pp. 24–30, 2005.

[13] C. Khundam, V. Vorachart, P. Preeyawongsakul, W. Hosap, and F. Noël, "A comparative study of interaction time and usability of using controllers and hand tracking in virtual reality training," *Informatics*, vol. 8, no. 3, p. 60, Sep 2021. [Online]. Available: http://dx.doi.org/10.3390/informatics8030060

[14] J.-H. Kim and I. Choi, "Choosing the level of significance: A decision-theoretic approach," *Abacus*, vol. 57, 11 2019.

[15] M. Kim, C. Jeon, and J. Kim, "A study on immersion and presence of a portable hand haptic system for immersive virtual reality," *Sensors*, vol. 17, no. 5, 2017. [Online]. Available: https://www.mdpi.com/1424-8220/17/5/1141

[16] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Trans. Information Systems*, vol. E77-D, no. 12, pp. 1321–1329, 12 1994.

[17] P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive statistics and normality tests for statistical data," *Annals of Cardiac Anaesthesia*, vol. 22, pp. 67–72, 2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6350423/

[18] M. Moencks and E. Roth, "Augmented workforce: Empowering people, transforming manufacturing," World Economic Forum, Tech. Rep., 1 2022, accessed: Feb. 11, 2023. [Online]. Available: https://www.weforum.org/whitepapers/augmented-workforce-empowering-people-transforming-manufacturing/

[19] K. Nasim and Y. J. Kim, "Physics-based assistive grasping for robust object manipulation in virtual reality," *Computer Animation and Virtual Worlds*, vol. 29, no. 3-4, p. e1820, 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cav.1820

[20] D. Navarro and V. Sundstedt, "Evaluating player performance and experience in virtual reality game interactions using the htc vive controller and leap motion sensor," in *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - HUCAPP,*, INSTICC. SciTePress, 2019, pp. 103–110.

[21] I. Newton, *Principia Mathematica Philosophiae Naturalis*, 1st ed. Londini: Jussu Societas Regiæ ac typis Josephi Streater, 1687.

[22] Nvidia. Joints. Accessed: 2023-10-24. [Online]. Available: https://nvidia-omniverse.github.io/PhysX/physx/5.3.0/docs/Joints.html

[23] ——. Rigid body dynamics. Accessed: 2023-10-24. [Online]. Available: https://nvidia-omniverse.github.io/PhysX/physx/5.3.0/docs/RigidBodyDynamics.html

[24] ——. Rigid body overview. Accessed: 2023-10-24. [Online]. Available: https://nvidia-omniverse.github.io/PhysX/physx/5.3.0/docs/RigidBodyOverview.html

[25] ——. Soft bodies. Accessed: 2023-10-24. [Online]. Available: https://nvidia-omniverse.github.io/PhysX/physx/5.3.0/docs/SoftBodies.html

[26] G. Palmer, *Physics for Game Programmers*, 1st ed. New York: Apress, 2007.

[27] P. C. Price, *Research Methods in Psychology*, 2nd ed. California: The Saylor Foundation, 2013, ch. 6.

[28] P. A. Rauschnabel, R. Felix, C. Hinsch, H. Shahab, and F. Alt, "What is XR? towards a framework for augmented and virtual reality," *Computers in Human Behavior*, vol. 133, 2022.

[29] J. T. Reason, "Motion sickness adaptation: A neural mismatch model," *Journal of the Royal Society of Medicine*, vol. 71, no. 11, pp. 819–829, 1978. [Online]. Available: https://doi.org/10.1177/014107687807101109

[30] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965. [Online]. Available: http://www.jstor.org/stable/2333709

[31] M. Speicher, B. D. Hall, and M. Nebeling, "What is mixed reality?" in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–15, accessed: Feb. 11, 2023. [Online]. Available: https://doi.org/10.1145/3290605.3300767

[32] I. Sünger and S. Çankaya, "Augmented reality: Historical development and area of usage," *Journal of Educational Technology and Online Learning*, vol. 2, no. 3, pp. 118–133, 2019.

[33] Unity Technologies. Unity user manual: Physics. Accessed: 2023-10-22. [Online]. Available: https://docs.unity3d.com/Manual/PhysicsSection.html

[34] J. Varela-Aldás, J. Buele, I. López, and G. Palacios-Navarro, "Influence of hand tracking in immersive virtual reality for memory assessment," *International Journal of Environmental Research and Public Health*, vol. 20, no. 5, 2023. [Online]. Available: https://www.mdpi.com/1660-4601/20/5/4609

[35] Y. Vermeulen, S. V. Damme, G. V. Wallendael, F. D. Turck, and M. T. Vega. (2022, 12) Haptic interactions for extended reality. Accessed Jan. 26, 2023. [Online]. Available: https://arxiv.org/abs/2212.04366

[36] R. Witte and J. Witte, *Statistics*, 11th ed. New Jersey: Wiley, 2017, ch. 15.

[37] ——, *Statistics*, 11th ed. New Jersey: Wiley, 2017, ch. 20.

Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden