Informatics

In-Situ-Questionnaires for Haptic Experience in VR

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Media and Human-Centered Computing

eingereicht von

Bc. Jana Varečková

Matrikelnummer 01528537

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ. Prof. Dr. Mag. Hannes Kaufmann
Mitwirkung: Dipl.-Ing. Mag Emanuel Vonach, Bakk.

Wien, 12. Oktober 2020

_____   _____
      Jana Varečková   Hannes Kaufmann

TU WIEN Informatics

In-Situ-Questionnaires for Haptic Experience in VR

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Media and Human-Centered Computing

by

Bc. Jana Varečková

Registration Number 01528537

to the Faculty of Informatics

at the TU Wien

Advisor: Univ. Prof. Dr. Mag. Hannes Kaufmann
Assistance: Dipl.-Ing. Mag Emanuel Vonach, Bakk.

Vienna, 12th October, 2020

_____   _____
Jana Varečková   Hannes Kaufmann

Erklärung zur Verfassung der Arbeit

Bc. Jana Varečková

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 12. Oktober 2020

                                 Jana Varečková

Acknowledgements

I would like to thank Emanuel Vonach for his support and guidance through-out the different phases of this master thesis. A special thanks is addressed to Hannes Kaufmann for supervising.

Further I would like to express gratitude to my parents Beata and Pavol, my sister Katka and my boyfriend Miro for their support. I also thank Misko for proof-reading.

And a big thank you to all participants in the user study.

Abstract

Real (physical) objects are being used in virtual reality (VR), where users see its 3D model and interact with the real object. However, creating or finding an exact replica from a 3D model can be costly. As users can adapt to small proprioceptive mismatches, the objects used do not have to be identical. There is no questionnaire established in virtual reality, which would focus on object features and would be able to compare objects and their suitability to be used as substitutes.

Therefore, our aim was to design a questionnaire, which would distinguish, whether the real object corresponds to a 3D model the user sees in VR based on the evaluation of object properties. To avoid incorrect completion of the questionnaire due to faulty memory, this questionnaire is asked while the users are still in VR and can interact with the object. Therefore, we named it the In-Situ-Questionnaire. To research different aspects of developing a suitable questionnaire, we designed three slightly different questionnaires, where the main change is a scale. To verify validity of the In-Situ-Questionnaires, a simple VR environment was created to test it during user studies. There the users could see their hands and interact with three different chairs. Each chair was evaluated separately. For comparison with the In-Situ-Questionnaires, we also created a Post-Questionnaire, where we used commonly used questionnaires.

Analyzing of the responses from the user studies, we concluded that the designed In-Situ-Questionnaires are able to compare different objects and indicate, which object is the most suitable as a substitute.

Kurzfassung

Reale (physische) Objekte werden in der Virtuellen Realität (VR) verwendet, in der die Benutzer das 3D-Modell sehen und mit dem realen Objekt interagieren. Das Erstellen oder Finden einer exakten Replik aus einem 3D-Modell kann jedoch teuer sein. Da sich Benutzer an kleine propriozeptive Diskrepanzen anpassen können, müssen die verwendeten Objekte nicht identisch sein. In der virtuellen Realität gibt es keinen etablierten Fragebogen, der sich auf Objektmerkmale konzentriert und Objekte und ihre Geeignetheit als Ersatz vergleichen könnte.

Daher war es unser Ziel, einen Fragebogen zu entwerfen, der anhand der Bewertung der Objekteigenschaften unterscheidet, ob das reale Objekt einem 3D-Modell entspricht, das der Benutzer in VR sieht. Um ein falsches Ausfüllen des Fragebogens aufgrund eines fehlerhaften Gedächtnis zu vermeiden, wird dieser Fragebogen gestellt, während sich die Benutzer noch in VR befinden und mit dem Objekt interagieren können. Deshalb haben wir es den In-Situ-Questionnaire genannt. Um verschiedene Aspekte der Entwicklung eines geeigneten Fragebogens zu untersuchen, haben wir drei leicht unterschiedliche Fragebögen entworfen, wobei die Hauptänderung eine Skala ist. Um die Gültigkeit der In-Situ-Questionnaires zu überprüfen, wurde eine einfache VR-Umgebung erstellt, um sie während Benutzerstudien zu testen. Dort konnten die Benutzer ihre Hände sehen und mit drei verschiedenen Stühlen interagieren. Jeder Stuhl wurde separat bewertet. Zum Vergleich mit den In-Situ-Questionnaires haben wir auch einen Post-Questionnaire erstellt, in dem wir häufig verwendete Fragebögen benutzt haben.

Die Analyse der Antworten aus den Benutzerstudien ergab, dass der entworfene In-Situ-Questionnaires verschiedene Objekte vergleichen und angeben kann, welches Objekt als Ersatz am besten geeignet ist.

Contents

Introduction and Motivation

There are numerous devices used in Virtual Reality (VR) that try to simulate haptic feedback. The word *haptic* refers to the ability to experience the environment through active exploration, typically using our hands or for example, to sense an object's shape and material properties. Haptics is important for fast, accurate interaction with our environment. We go about our daily tasks without conscious awareness of haptics. But when in VR, the users are not able to grab an object that is in front of them with controllers. They do not have enough haptic information. Only with time are the users able to figure out how their actions change the state of the environment and are able to grab the object. For this reason, if the interface does not provide meaningful haptic information to learn and perform the task, it might impact users performance [RDLT06]. The users can touch, push, grab and manipulate virtual objects with different levels of similarity of tactile and force stimulation, which users are familiar with in the real world [COB+18]. For example, sometimes these devices do not provide sufficient feedback e.g. for individual fingers, or rely on external structures and cables that disturb the interaction. For this reason researchers are experimenting with real objects which are mapped to virtual ones [SVG15].

1.1 Motivation and Problem Definition

Having similar objects in the real and in the virtual world is not always possible or can be costly. Therefore, researchers often use real objects that are not an exact representation but approximate the virtual one. For example, Simeone et al. [SVG15] have used one mug for mugs with different characteristics together with a sphere, a box and a lamp. To distinguish whether users can tell apart if they were touching something else than what they were seeing, they used questionnaires. However, there are no established questionnaires for haptic experiences in VR and the ones that are used are in dispute whether they ask the right questions.

One of the questionnaires used in the VR field for measuring user experience with real objects is the Presence Questionnaire (PQ) [WS98]. However, Slater [Sla99] argues that questions in the Presence Questionnaire are subjectively defined. According to their findings the questions are fine in themselves, but they elicit users' opinions, none of the questions are directly about object presence and they are asked after the experience. The users can report different responses for the same question due to different experiences, their psychological make-up, skills, etc.

Other questionnaires being used are, for example, the Immersive Tendencies Questionnaire [WS98] and the AttrakDiff Questionnaire [MAO17]. The Immersive Tendencies Questionnaire measures different tendencies of individuals to experience *presence*, which can be defined as a subjective experience of being in one place or environment, even when one is physically situated in another [WS98], and the AttrakDiff Questionnaire measures the attractiveness of an interactive product. These questionnaires are often used, but they still do not measure how the users perceive objects in more depth. Another method commonly used is talking out loud. However, this method distracts people from their actions.

There is no established questionnaire that focuses on the haptic quality of a virtual experience and would allow an objective comparison between different methods for haptic feedback. That is why there is a need to create a questionnaire that fulfills conditions of not disturbing the users while they are in VR, finding out how they experience objects and which properties correspond to their experience. For this purpose, we have created suitable questionnaires and tested them to determine their validity.

1.2 Purpose of the Work

We think that there is a lack of any tools to measure haptic feedback, how properties of an object impact the experience. With all existing approaches, we can only measure impressions where haptics is a part of the experience, e.g. *presence*. If there is a change in e.g. *presence*, we can only guess if the reason for it is the differences of haptic devices, haptic realism or something completely different, like disturbing noises or non-realistic faces.

The aim of this thesis is to develop questionnaires that measure users experience interacting with real objects in VR. To find out whether their virtual counterparts correspond to the real ones, we developed questionnaires, which we believe give more insights into important aspects of object properties. We wanted to design questionnaires that make it possible to get an exact measurement of which property differs (e.g. weight, shape). However, we did not create just the VR questionnaire, but we also used existing questionnaires commonly used for evaluation to create a second questionnaire for comparison.

We named the first questionnaire the *In-Situ-Questionnaire*. It is designed to be used during the experiment while the users are engaged in VR as it is a part of it. The users can fill out the In-Situ-Questionnaire still while they are able to touch the object and

make sure they evaluated it correctly. To have a higher probability of developing a suitable questionnaire, we designed three slightly different In-Situ-Questionnaires, which were all tested. They were developed not only to find out whether our questionnaires allow for more insights into how the users are experiencing touching an object in VR, but also which questionnaire's design is the most suitable.

The reason why we chose to ask questions in VR is that questionnaires like the Presence Questionnaire or System Usability Scale are asked after the users have already accomplished all tasks and are back in the real world. However, the users usually do not remember how they exactly felt while doing the experiment and their memories can introduce more bias to the resulting feedback.

The second questionnaire, which we named *Post-Questionnaire*, is asked after every test session was completed. While the first questionnaire gives us brief information about how the users perceived touching and seeing an object, in the Post-Questionnaire we ask questions from the System Usability Scale and the Presence Questionnaire for comparison with existing questionnaires. As we asked questions from the Presence Questionnaire, we get information about whether users felt immersed during experiment, which is an important predisposition, otherwise it could indicate problems with the virtual environment or a possibility that they did not concentrate on the task.

To test our In-Situ-Questionnaires, we used real objects. These objects were either an identical, a similar object or an object different from what the user saw in VR. Through these differences in object properties, we wanted to test whether our questionnaires can measure the differences in user experience.

For testing purposes, we designed a simple task which the users had to accomplish. We chose the real object to be a chair. We used chair because when we attached tracker to find out position of the chair, it did not noticeably change its weight, we found a suitable place where to attach it and the tracking was the most stable among tested objects. Users' task was to sit down on it in VR. This task engaged the users in the virtual world and motivated their interaction with the virtual objects. In order to touch the object, the users should know where their hands are and where they should move them to touch the object. For this reason, we developed a user environment in VR in which users see their hands and gestures they make. However, to know where the hands are is not enough. We needed object tracking as well to know where the chair is. As people touch and grab the object, they naturally move it. Therefore, we needed to track the object so that the 3D model of the object moves in VR as well as the real object. After the users had touched the object and accomplished the task, they were asked to fill out the In-Situ-Questionnaire, which is in VR. Thanks to which the users did not have to exit the virtual environment to evaluate the chair. For each one of the three In-Situ-Questionnaires users were supposed to evaluate the same six properties.

1.3   Structure of the Thesis

Chapter 2 contains an overview of different questionnaires used for evaluation of user's experience in VR. Various technologies and concepts which we use for our VR system are briefly introduced in chapter 3. Chapter 4 contains the conceptual design of the In-Situ-Questionnaire and the environment for evaluation. In chapter 5 we discuss the implementation of the developed system and give details about specific implementation features. Chapter 6 covers the results of our user research, as well as its execution and interpretation. The outcome of our research is discussed in chapter 7. Finally, we give a brief summary and an outlook into the future of this topic in chapter 8.

Related Work

In this chapter we write about the questionnaires which were used by researchers that studied haptic feedback in VR. However, in many papers describing their experiments, the researchers did not mention which questionnaires were used or whether they used any at all (e.g. Cheng et al. [CRR+15], Chague et al. [CC16]). Others described a questionnaire where they used their own set of questions (e.g. He et al. [HZGP17], Yoshimoto et al. [YS17]), or did not use any questionnaire except a brief background questionnaire (e.g. Han et al. [HSR18]). Another option is to use multiple questionnaires. For example Maggioni et al. [MAO17] studied added value of haptic stimulation, where participants' feedback was measured by using AttrakDiff questionnaire, to study user experience, Self-Assessment Manikin, to study emotions, and they created their own set of questions to be answered by a 7-point Likert scale to find out about the participants' expectations. For instance, to find out about the comfort of the haptic feedback, they asked:

> *I think the haptic feedback will be comfortable while watching a video.* [MAO17]

As there is not an established questionnaire for haptics in VR, in this chapter we will describe questionnaires the researchers were using to evaluate their haptic experiments in VR. We will start by describing the most commonly used questionnaires and their relation to our questionnaires.

2.1 Presence Questionnaires

Not everyone can grasp what the word *presence* means. Even researchers have difficulties agreeing on one meaning. One of the definitions, already mentioned in the section 1.1, is that *presence* can be defined as a subjective experience of being in one place or environment, even when one is physically situated in another [WS98]. Another definition

is by Flach et al. [FH98] where they define presence as being more concerned with action than with the appearance of how things look and sound like. They argue that *being there* is the ability to *act* there. There are multiple other definitions as well.

When we talk about *presence* we not only have to talk about the definition, but what it influences is an interesting factor too. Usoh et al. [UCAS00] summarized these factors in VR. These factors are: only high-resolution information should be displayed to the users, but without indication of a display device. Consistency across all sensory modalities is an important factor too. The users should be able to navigate through and interact with the environment in VR, quickly learn the effect of their action and the users' representation in VR should be similar in appearance and appropriately respond to what the users do.

The first attempts to measure *presence* relied on the VR participants' self-reported sense of presence [RDI03]. The most common method of measuring this subjective presence is the post-immersion questionnaires. In this section we will describe two of these questionnaires. The Presence Questionnaire by Witmer and Singer and the Slater-Usoh-Steed questionnaire were among the first widely used presence questionnaires [CGL10]. However, neither one of them could pass a *reality test* - they both failed to produce significantly greater presence scores for a person in a real environment than in a virtual one and there is a lack of reliable statistical data to support their claims [YP02].

As haptics is only a part of *presence*, it is hard to analyze it separately with these questionnaires.

2.1.1 Presence Questionnaire by Witmer and Singer (PQ)

The first of the two presence questionnaires we will describe is the Presence Questionnaire (PQ) by Witmer and Singer [WS98]. They attempt to look beyond just immersion and try to measure the users' involvement as well. They describe presence as a normal awareness phenomenon requiring direct attention, which is based on the interaction between sensory stimulation, environmental factors encouraging involvement and enabling immersion, and internal tendencies to become involved. Witmer and Singer measure presence with their questionnaire as the sense of presence was linked to the effectiveness of virtual environments (VEs).

3. How natural did your interactions with the environment seem?

|_____|_____|_____|_____|_____|_____|_____|
EXTREMELY   BORDERLINE   COMPLETELY
ARTIFICIAL   NATURAL

Figure 2.1: Question with 7-point scale from revised PQ by the UQO Cyberpsychology Lab [Lab].

PQ uses a 7-point scale format, which is based on the semantic differential principle. However, unlike the semantic differential answering options of each item, they are not

only anchored at each end using opposite descriptors, but there is a midpoint anchor as well. These anchors are based on the content of the question (see Figure 2.1).

The PQ is composed of 32 questions, grouped into four categories [RDI03] Control Factors, Sensory Factors, Distraction Factors, and Realism Factors. The Control Factors examine for instance the ability to control the relation of sensors to the environment or the amount to which a user can anticipate what can happen next in the environment:

26. How quickly did you adjust to the virtual environment experience?. [RDI03]

The Sensory Factors examine, for example, the amount, coherence, and consistency of information picked up by different senses:

5. How much did the visual aspects of the environment involve you?. [RDI03]

The Distraction Factors measure the possible distractions a person may experience in a virtual environment, such as awareness of the real environment, devices used to transmit the virtual environment to the user:

8. How aware were you of events occurring in the real world around you?. [RDI03]

Realism Factors measure factors such as environment realism and meaningfulness, as well as disorientation when returning to the real world:

22. To what degree did you feel confused or disoriented at the beginning of breaks or at the end of the experimental session?. [RDI03]

The PQ is also divided into five subscales [TTLECR16]: involvement/control, natural, auditory, haptics, resolution and interface quality. The haptics subscale is the most interesting subscale for our purposes despite having just two questions as by asking the In-Situ-Questionnaires we acquire information about haptic feedback. However, these two questions are not enough to have fine feedback about the nature of haptic feedback. We also used PQ in our Post-Questionnaire as it is the most common evaluating questionnaire used. The haptic feedback questions are:

17. How well could you actively survey or search the virtual environment using touch?

21. How well could you move or manipulate objects in the virtual environment? [TTLECR16]

The PQ was used together with the SUS PQ by Usoh et al. [UCAS00] in an experiment in a real and virtual environment. However, the scores for the real and the virtual environment were not significantly different. Even though Witmer and Singer [WS98] argue that Presence Questionnaires should be able to pass a reality test, where presence scores should be higher for a real experience in comparison to a virtual one, Usoh et al. [UCAS00] do not think that these questionnaires can be used to compare the experience across different environments. However, they can be useful for evaluating experiences of participants for the same environment.

Slater [Sla99] also disputes whether the PQ is well formulated. He argues that questions in the PQ elicit opinions, measuring the user's perception of system properties, rather than psychological presence. For example, one question asks:

How much were you able to control events? [Sla99]

The users answered this question with different degrees of ability to control events, despite the system being the same. Slater thinks that the differences in answers have nothing to do with the immersiveness of the system, but rather with e.g. differences among individuals, their experience, psychological make-up or dexterity.

The PQ is used after the task is already finished, which is a disadvantage as the users usually do not remember everything and cannot verify their answers. To mitigate this disadvantage, we display the In-Situ-Questionnaires while people are still in VR and can change their answers while still experiencing the sensations. However, we used the PQ in the Post-Questionnaire for comparison and to measure the presence as it was linked to effectiveness of virtual environments. This could give us a hint about, for instance, whether people had problems with evaluation.

2.1.2 Presence Questionnaire by Slater Usoh Steed (SUS PQ)

The second Presence Questionnaire was developed by Slater-Usoh-Steed (SUS PQ). SUS PQ consists of 6 questions [CGL10]. These questions look at three themes to identify a sense of physical presence in an environment. Each question is rated on a 7-point scale [RDI03]. In the study by Usoh et al. [UCAS00], already mentioned in the section PQ, people were searching for a red box hidden in a university laboratory research. They had two versions of the laboratory research. One was real, and the other was a virtual environment model of this space. We will give an example of the questions, which were adapted particularly for this experiment.

The first theme looks at the user's sense of being in the virtual environment:

2. To what extent were there times during the experience when the office space was the reality for you?

There were times during the experience when the office space was the reality for me. . .

(1) At no time. (7) Almost all the time. [UCAS00]

The second theme examines the extent to which the virtual environment becomes the user's dominant environment:

4. During the time of the experience, which was the strongest on the whole, your sense of being in the office space, or of being elsewhere?

I had a stronger sense of. . .

(1) Being elsewhere. (7) Being in the office space. [UCAS00]

And the last theme examines the extent to which the virtual environment is remembered as an actual place:

6. During the time of the experience, did you often think to yourself that you were actually in the office space?

During the experience I often thought that I was really standing in the office space. . .

(1) Not very often. (7) Very much so. [UCAS00]

The SUS PQ was used in the experiment by Nagao et al. [NMN+18] to evaluate the participants' sense of presence. They used small bumps as a haptic cue to imitate edges of the stairs, which people felt by their feet. They compared the presence scores when they used haptic cues and without them. The sense of presence was significantly higher when they used the haptic cues.

Even though the presence correlates with the use of a haptic cue in the Nagao et al. [NMN+18] research, we do not know how the users perceived the small bumps. Maybe they would have liked them to be bigger, and that would make their experience presence even better. As this questionnaire is used after the experience, their answers may not be accurate, which we addressed in the In-Situ-Questionnaires as compared to when it is used during the experience. When it comes to the physical objects, whether what they saw corresponded to what they felt with their feet, with the In-Situ-Questionnaires we would receive information about a particular property that did not feel right.

2.2 Immersive Tendencies Questionnaire (ITQ)

Another questionnaire used in VR is the Immersive Tendencies Questionnaire (ITQ) [WS98]. The ITQ, like PQ, relies on self-reported information. If the users are not immersed, they cannot experience presence. ITQ was developed by the same authors as the PQ 2.1.1. Witmer and Singer developed the ITQ to measure the capability or tendency of individuals to be involved or immersed. They believe that involvement and immersion are necessary for experiencing *presence*. User involvement varies based on their attention. On the other hand, they described immersion as a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences.

The ITQ [WS98] is composed of 29 questions. Most of the questions measure involvement in common activities. A few questions e.g.

> *10. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?* [WS98]

measure immersive tendencies directly. Others measure the ability to focus or redirect one's attention or assess one's current fitness or alertness. ITQ uses a 7-point scale format, which is based on the semantic differential principle like the PQ.

Even though Witmer and Singer's research shows positive correlation of ITQ scores with PQ scores, when Johns et al. [JND+00] attempted to replicate these findings in non-immersive virtual environments (the first VE designed to engender a high sense of presence in users; the second VE to disrupt and decrease the sense of presence felt by users), their findings showed that the PQ scores did not reflect these differences in VEs and the PQ scores were correlated with ITQ scores only in the high-presence environment.

However, both PQ and ITQ questionnaires focus on other aspects of a virtual experience. They are not very suitable to measure the haptic experience, even combining them is insufficient. We would like to find out what the users really experience when touching an object in VR and in more detail.

2.3 Self-Assment Manakin (SAM)

The Self-Assment Manakin (SAM) is used to capture the emotion dimensionality [MAO17]. It is a rating system that has three dimensions: valence (events, objects or situations may posses intrinsic attractiveness or aversiveness [F+86]), arousal and dominance (the feeling of being in control or controlled). For evaluation, the SAM uses graphic figures (see Figure 2.2) depicting different values on a scale indicating an emotional reaction. The users select an appropriate figure for each dimension. To provide more consistent psychometric measurement across 5- and 7-point Likert scales, the rating scale has been modified over the years.
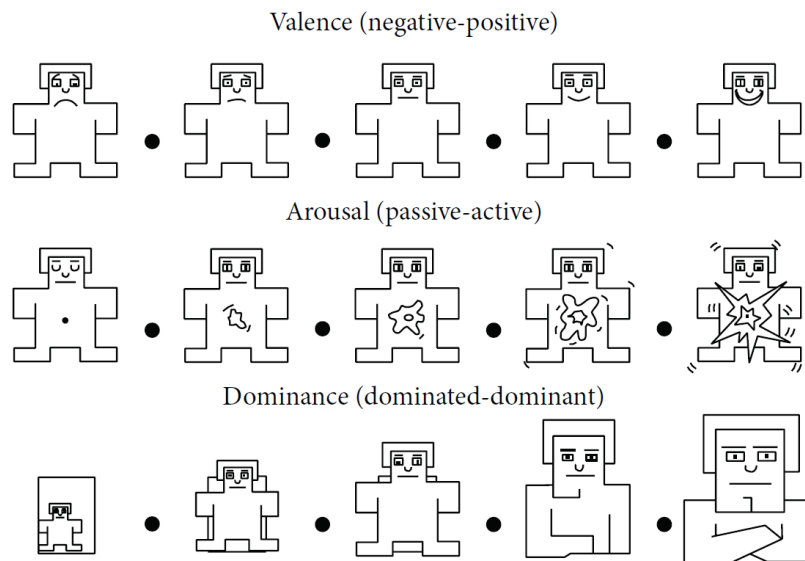
Figure 2.2: SAM from Jirayucharoensak et al. [JPNI14]. Valence is depicted by ranges from a smiling, happy figure to a frowning, unhappy figure. An excited, wide-eyed figure to a relaxed, sleepy figure is arousal. And Dominance is shown from being a large figure (in control) to a small figure (dominated) [MAO17].

Maggioni et al. [MAO17] used the SAM scale for mid-air and vibrotactile stimulation to capture emotional reaction based on the valence and arousal ratings. However, they did not use the dominance dimension as they used rather a passive than an active interaction scenario. The evaluation enabled them to compare added value for the two mentioned haptic stimulation modalities (mid-air and vibrotactile). They used specifically the 7-point Likert scale. They also compare the scenes without the haptic stimulation, using only audio-visual stimuli. They randomized the order of 4 audio-visual stimuli with or without haptic feedback which were randomized in blocks. After each scenario, the participants were given the AttrakDiff questionnaire, SAM, and they captured liking on a 7-point Likert scale.

The SAM represents a good illustration of an easily understandable questionnaire where pictures represent something that would be hard to describe accurately.

We have implemented a similar approach with the In-Situ-Questionnaires where we used pictures. However, they did not represent emotions, but they were a depiction of a corresponding object. However, the SAM captures only the emotional reaction and does not explain why the users evaluated it that way and which properties were important for the rating. It does not tell us about the tactile properties of the haptic stimulation. We believe that with the In-Situ-Questionnaires we get more detailed information about the possible reasons why the users evaluated the haptic stimuli the way they did.

2.4 System Usability Scale (SUS)

The System Usability Scale was developed by Brook [B+96] and brought into attention in 1986. It is a 10-item questionnaire scored on a 5-point Likert scale [LGS+16], 5 meaning agree completely, 1 meaning disagree vehemently. The statements in the SUS alternate between positive and negative. This has to be taken into account when scoring the survey. For each odd numbered question, 1 should be subtracted from the score, for each even numbered question, 5 is subtracted from their value. Everything should be added up and multiplied by 2.5 [B+96]. The final score ranges from 0 to 100. Higher scores indicate better usability. Acceptable scores of the SUS are above 70. A superior product scores more than 90. If the score is less than 70, the product should be improved. If the product scores less than 50, it is than deemed unacceptable. Bangor et al. [BKM08] used a slightly modified version of the SUS as about 10% of participants asked about the word *cumbersome*. Instead of *cumbersome* they used *awkward* as it is a more commonly used word in English than *cumbersome*. They also replaced the word *system* with *product*.

Figure 2.3: First four questions from the System Usability Scale showing alteration of positive and negative items [Bro].

During the user studies we used the original SUS questionnaire where we used the word *cumbersome* and we experienced that some participants asked us what it means. We used the SUS in the Post-Questionnaire as it was developed to quickly and easily collect users' subjective rating of a product's usability [BKM08]. The SUS is a highly robust and versatile tool, which is technology-agnostic. This makes it flexible to assess a wide range of interface technologies. Ranging from traditional computer interfaces and websites, to new technologies like VR. The survey's scale is easily understood providing only a single score. Another advantage is that the SUS is non-proprietary, which makes it a cost-effective tool.

The SUS was used for instance by Sait et al. [SSHR18], where they studied methods allowing physical hand interaction with virtual objects using passive props. Usability of

two passive-haptic interaction methods and a virtual hand approach without haptics was compared. The researchers mapped just one physical prop to multiple different virtual objects distributed at different locations. A game was developed, which participants were asked to play three times, while each trial used a different technique (two were with haptics - physical props, one was without haptics, where participants used only gestures - air grasping). After each trial the participants were given a break and were asked to fill out two questionnaires: the SSQ and the SUS. The SUS was used to give usability feedback on the game experienced and techniques used. However, their results showed that all techniques had no significantly different usability scores. The scores ranged from 60 to 82,5. Nonetheless, the researchers also observed participants, took notes about their behaviors and comments and at the end, the users also answered a semi-structure interview. This gave them more information about how the participants experienced the game and the system's flaws. Even though the SUS scores were similar, the participants reported a higher sense of control and realism when they interacted with a physical prop compared just to the air grasping without tactile feedback. One important observation they mention is that even though the size and scale of the virtual model and physical prop were not accurate, no user commented on it.

The SUS is good for overall assessments of the system usability and is described as a reliable, low-cost usability scale which can be used for global assessments of system usability [WD17]. Nonetheless, the SUS was never intended to diagnose haptic interaction problems. Therefore, we need additional information and explore further to find the culprit if the score is low to identify the problems.

We used the SUS to gather usability feedback from users about our VR environment as when the users are not able to properly interact within the VR environment e.g. with the chair, we would not get relevant results from the In-Situ-Questionnaires. Also, when there is a difference between the virtual model and physical prop, the questions for instance about the ease of use or inconsistency in the system should receive worse score, so the overall score would be lower. Thus, based on the score we should be able to distinguish whether the users did or did not notice the difference between the virtual model and the physical prop.

2.5 AttrakDiff

The AttrakDiff questionnaire is a simple yet effective measuring tool [MAO17] allowing researchers to make decisions either by comparing different versions of the product in an iterative process or comparing it with a competing product. The AttrakDiff questionnaire consists of twenty-eight semantic differentials of opposite adjectives (e.g. ugly-attractive, unusual-ordinary, complicated-simple, impractical-practical). Between opposite adjectives is a 7-point Likert scale which the user can select from.

In the AttrakDiff questionnaire there are four dimensions being evaluated [SAMGB+18], which are color-coded in Figure 2.4. The first dimension is *Pragmatic Quality (PQ)*, describing usability of the product and indicating how successful users are in achieving

their goals using the product. The second dimension is *Hedonic Quality Identity (HQ-I)*, indicating to what extent the product allows the users to identify with it. The third dimension is *Hedonic Quality Stimulation (HQ-S)*, indicating to what extent the product can support the need to develop and move forward in terms of novel, interesting and stimulating functions, contents, interaction and presentation styles. The last dimension is *Attractiveness (ATT)*, which describes the global value of the product based on the quality of perception.

Figure 2.4: AttrakDiff questionnaire with dimensions. From Sanchez et al. [SAMGB⁺18]

Even though Maggioni et al. [MAO17] claim that the added value of haptic stimulation is best captured with the AttrakDiff questionnaire, we would argue that the AttrakDiff does not measure the properties of an object. It evaluates users' subjective feeling of the haptic stimulation, but we would like to measure the users' perception of a specific object, which our In-Situ-Questionnaires can. It measures users' perception of specific properties of the object.

2.6   Simulator Sickness Questionnaire (SSQ)

The Simulator Sickness Questionnaire (SSQ) by Kennedy et al. [KLBL93] lists 16 common symptoms which participants can rate on a 4-point scale, starting from 0 meaning none, to 3 meaning severe. Nausea, oculomotor discomfort and disorientation are subscale

scores, which are formed by combining evaluations of the questions. Total severity score uses all symptoms and should reflect whether there is a sickness problem. Some of the symptoms are for instance sweating, nausea, fatigue or headache (see Figure 2.5).

The SSQ was for instance used for checking any discomfort caused by the techniques used in VR, where Nagao et al. [NMN+18] provided a passive haptic stimuli by small bumps under the users' feet to make them feel they are ascending or descending stairs.

Sait et al. [SSHR18] used the SSQ to evaluate any fatigue or sickness experienced after completion of each game trial. See section 2.4 for more details about the SUS.

Even though the SSQ can not measure the haptic feedback, it could be a good predictor of whether the haptic interaction was well-made and whether the users are immersed, which could influence the results of the experiment. As for example when the model of an object in VR is moving, but not the real object and the users are interacting with it, it could cause discomfort.

Symptom	Severity			
General discomfort	None	Slight	Moderate	Severe
Fatigue	None	Slight	Moderate	Severe
Headache	None	Slight	Moderate	Severe
Eye strain	None	Slight	Moderate	Severe
Difficulty focusing	None	Slight	Moderate	Severe
Increased salivation	None	Slight	Moderate	Severe
Sweating	None	Slight	Moderate	Severe
Nausea	None	Slight	Moderate	Severe
Difficulty concentrating	None	Slight	Moderate	Severe
"Fullness of the head"	None	Slight	Moderate	Severe
Blurred vision	None	Slight	Moderate	Severe
Dizzy (eyes open)	None	Slight	Moderate	Severe
Dizzy (eyes closed)	None	Slight	Moderate	Severe
Vertigo (Giddiness)	None	Slight	Moderate	Severe
Stomach awareness	None	Slight	Moderate	Severe
Burping	None	Slight	Moderate	Severe

Figure 2.5: SSQ from [DAB09].

2.7 Visually Induced Motion Sickness (VIMS)

Visually induced motion sickness (VIMS) is evoked by conflicting motion sensory signals within the brain [HDGP18]. Symptoms such as dizziness, nausea or light-headedness may occur and are similar to those of the motion sickness experienced in cars or boats. When these symptoms are induced by visual symptoms, they are called VIMS.

As motion sickness can influence the users' immersion and experiment results, we decided to use a method of measuring motion sickness with a single question (see Figure 2.6) intro-

duced by Hwang et al. [HDGP18], which is based on the Wong-Baker FACES [WHEW⁺96]. Our participants were asked this question before putting a headset on and after completing each of three tasks in VR, to make sure they feel well enough to be able to fill out the In-Situ-Questionnaires without any problems.
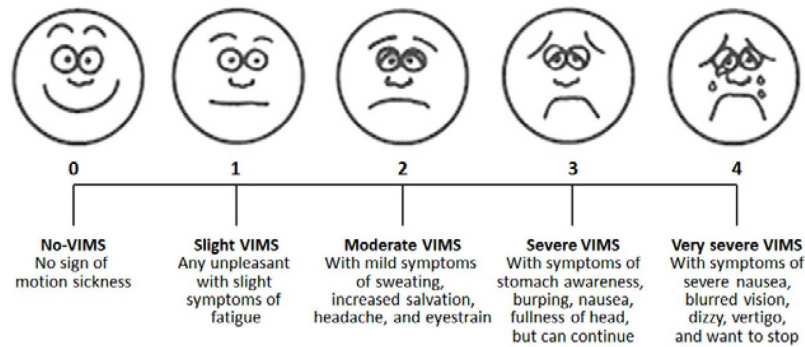

Figure 2.6: VIMS questionnaire from [HDGP18].

2.8 Custom-made Questionnaires

Not all researchers use existing questionnaires. Some created their own set of questions to answer their specific research questions.

2.8.1 Virtual vs. Physical Version

In an experiment by He et al. [HZGP17], the users were among other things playing a 2-player Tic-Tac-Toe. However, players were not sitting opposite of each other, but had their own table with a haptic proxy, which was a mobile robot (see Figure 2.7) used for haptic feedback. For comparison, they used a purely virtual version as well. In the virtual version they tracked the participants' hand. With the haptic approach, the VR users were able to share and manipulate a tangible physical object with remote collaborators. The mobile robots provided a deeper level of immersion by not only imitating the real object, but by applying some force, too. To improve user experience, the researchers added *artificial* latency meaning that when one player was doing the move, the other one saw it a bit later.

For evaluation, they recorded the elapsed time and collected answers to the following questions:

Did you feel like you and your test-mate were sitting at different tables?

Did you feel like your opponent was moving naturally?

Were you more comfortable with the physical or the virtual version of the game?

On a scale from one to five:

How much delay did you experience between your actions and their expected outcomes? (Not much – unendurable)

How well did you feel you could manipulate objects in the virtual environment? (Not well – very well)

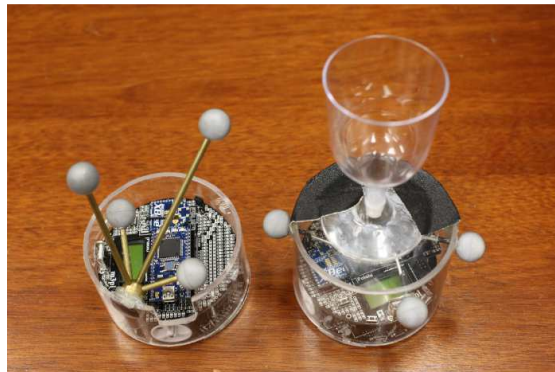*Did you understand the game? (No – enjoyable)* [HZGP17]

Figure 2.7: Mobile robots as proxies [HZGP17].

Based on the answers to these questions, they concluded that the robotic proxies were made in a satisfactory way as for instance 11 out of 16 participants chose the physical controller over no controller, or the participants on average graded 4.02 (the maximum was 5) that they felt they can manipulate objects in VR.

The questions were made particularly for this experiment and the researchers got an answer whether the proposed system is acceptable or not. However, the questions were either yes-no questions or the participants answered them on a scale. They got a hint where an issue may lie, but as the questions were not open-ended they did not get a more in-depth analysis of what they could improve. Although He et al. [HZGP17] used real objects, except for the third question they did not inquire about the users' haptic experience with them. The developed In-Situ-Questionnaires were specifically designed to evaluate the users' haptic experience.

2.8.2 Stiffness of a Virtual Object

Another example of a self-made questionnaire is by Gaffary et al. [GLGM⁺17]. The researchers compared the perception of a piston in augmented reality (AR) and virtual reality using a haptic force-feedback device, enabling to press a virtual piston (see Figure 2.8). The researchers created a questionnaire to ensure the correctness and quality of their setup and to inquire whether the participants felt visual fatigue. They asked a questionnaire with seven statements on a 7-point scale, the last two statements were:

After the experiment, I felt haptic fatigue.

Do you think that the real environment influenced your haptic perception of the virtual piston? If so, how? [GLGM⁺17]

The researchers also asked an open-ended question to receive subjective users feedback concerning a possible influence on stiffness of the piston when the users were in VR and AR:

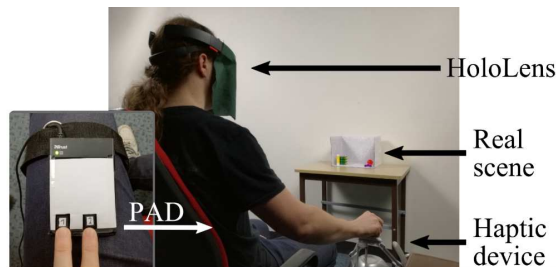*Do you think that the real environment influenced your haptic perception of the virtual piston? If so, how?* [GLGM⁺17]

Figure 2.8: The experimental setup. The participant uses a pad to answer which of the two virtual pistons was the stiffer one. A haptic device, Novint Falcon, is used as a virtual piston [GLGM⁺17].

An interesting outcome of this study was that the piston in VR was perceived as stiffer than in the AR. This is a change in haptic experience, which we capture with the In-Situ-Questionnaires.

2.8.3 Rating Properties

Simeone et al. [SVG15] created a virtual environment, a medieval courtyard, which was based on the layout of a real room. Objects in the real room were paired, with some discrepancy, to their virtual counterparts in the medieval courtyard. In the study they explored how different the proxy of a real object can be before breaking the VR illusion. To investigate this, they used a mug as a baseline object. They showed the users different 3D models in VR (see Figure 2.9). They created the baseline object model. The substituted objects were created by altering their apparent physical properties. For instance, they altered the material of the mug (to be wooden), its perceived surface temperature (the mug looked like it is made out of ice), size and shape.

At the beginning of the test session, the researchers asked the participants questions from the demographic questionnaire. To evaluate the virtual proxies, they asked a set of questions which were answered on a 7-point scale, where 1 was the lowest and 7 the
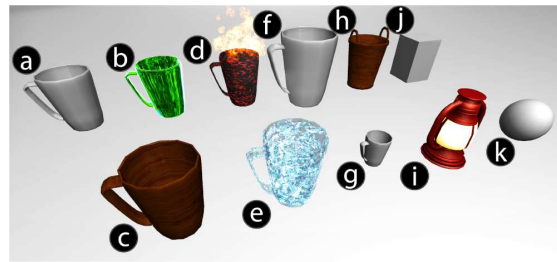
Figure 2.9: Virtual substitution objects for the real mug [SVG15]. The replica mug used as the baseline (a); and the substitutions: a glass (b) and a wooden mug (c); a hot (d) and an ice-cold mug (e); a big (f) and a small mug (g); a basket (h) and a lamp (i); a box (j) and a sphere (k).

highest score. The questions themselves were not specifically mentioned in the research paper, only descriptions of the topics were provided. In five of these questions the researchers asked how similar the virtual object felt to the real one in terms of physical properties (size, shape), temperature, material and weight. Another two questions were about perceived properties of the object (ease of grabbing and manipulation) and how likely they were to believe that what they were actually manipulating the virtual object, considering the overall mismatch between the two objects.

These questions were asked while the participants were in VR, and they placed the mug down after a minute of examining it by touch. In total, the participants examined 11 objects mapped to the real mug. After each session, they were given the SUS PQ to measure the presence.

In this research they studied the differences between the properties in which the participants perceive the proxy, and the real object. This is similar to what we wanted to examine. However, what is different is how they asked the questions. Even though the participants were asked the questions while in VR, they were asked the questions by someone outside of VR and could not see the questionnaire. We want the users to be able to see the questionnaire and fill it out at their own pace, while still experiencing the haptic properties.

CHAPTER 3

Setup

This chapter briefly introduces the reader to the technologies, which were used for creating the VR scenes where In-Situ-Questionnaires were evaluated. This information should give enough background information for understanding the concepts discussed later in this thesis.

We will start by describing one of the VR systems, specifically the HTC Vive, as we used this particular one. Together with the HTC Vive tracker they created a good pair of devices, as the HTC Vive tracker enables real objects to be manipulated and mapped to a virtual object. Even though there is a hand tracking feature in the newer HTC Vive version, we used the LEAP Motion Orion for hand tracking and hand visualization in VR as the HTC Vive hand tracking was still very experimental and not reliable. The VR scene was designed and programmed in the Unity engine; SteamVR was used to facilitate communication with the hardware.

3.1 HTC Vive

The HTC Vive is a virtual reality system. It is comprised of base stations (also called *lighthouses*), two ring-shaped controllers and a headset (see Figure 3.1). The base stations sync wirelessly and synchronously emit infrared pulses at 60 pulses per second, one spinning horizontally and the other vertically [CGAK+19]. Each controller and headset contains infrared sensors, which are hit by the light emitted by the base stations [BSC+18]. The controllers have 24 infrared sensors across the ring. The sensors are used to measure light pulse timing to estimate the horizontal and vertical angles to the base stations and this data is used to calculate position.

The headset has two 1080p AMOLED displays with 2160x1200 resolution and 90 MHz screen refresh rate. It contains an accelerometer, a gyroscope, an infrared sensor and a proximity sensor.

Figure 3.1: HTC Vive.

We had the users put on the headset and with the HTC Vive system we were able to display the scene and track where the users were positioned within the scene. The users had the option to use the controllers to interact with the In-Situ-Questionnaires.

3.1.1   HTC Vive Tracker

Figure 3.2: HTC Vive Tracker

The HTC Vive tracker (Figure 3.2) is designed to be attached to any real-world object and bring it into the virtual environment [CGAK$^+$19]. It has 24 infrared sensors across the ring (as do the controllers) to calculate its position. Using velcro tape, the HTC Vive tracker can be attached to the users to track their body position.

We attached the HTC Vive tracker to chairs to track them and display them in VR. When the users moved the tracked chair, its virtual representation moved as well based on the real chair movement.

We used the HTC Vive system (headset, controllers, base stations, trackers) because it allows precise tracking of the users and objects in a room scale. It is not so prone to occlusions and the integration of hand tracking with LEAP Motion is easily done. These requirements are important for our experiment.

3.2 Leap Motion

Figure 3.3: LEAP Motion Controller.

LEAP Motion enables hand interaction with the virtual environment [DNJ18]. The LEAP Motion controller (see Figure 3.3) is a sensor device using hand and finger motion as input [HHY+17]. It does not require touching a controller. Hand tracking and finger movement are accomplished using two cameras and three LEDs, emitting IR beams from within the controller [DNJ18]. The LEAP Motion can detect gestures like object grabbing [LMb]. The controller has two modes: it can be placed either on a desktop or attached onto the VR headset. The controller is connected to the computer via USB cable. Data is processed on the computer. It can detect an object within a distance of up to 80cm. The device's advantage is high precision, but it has a short range and a small field of view. The LEDs generate a 3D pattern of IR light dots.

Using LEAP Motion, the users in our user study were able to interact with the In-Situ-Questionnaires and touch the chair as they knew where their hands were relative to the chair.

3.3 Unity

Unity (Figure 3.4) is a cross-platform game engine [Tece]. It is used to create 2D, 3D, VR or augmented reality games and simulations, all running on a desktop or a mobile device. It has its own Unity Asset Store where developers can buy assets. Unity uses a simple technique of dragging and dropping objects into the scenes.

We made our 3D evaluation environment (see subchapter 5.3) in Unity by using existing 3D GameObjects like a plane for a wall or floor, UI elements for the In-Situ-Questionnaires or an almost identical 3D model of a chair we used in our experiments.
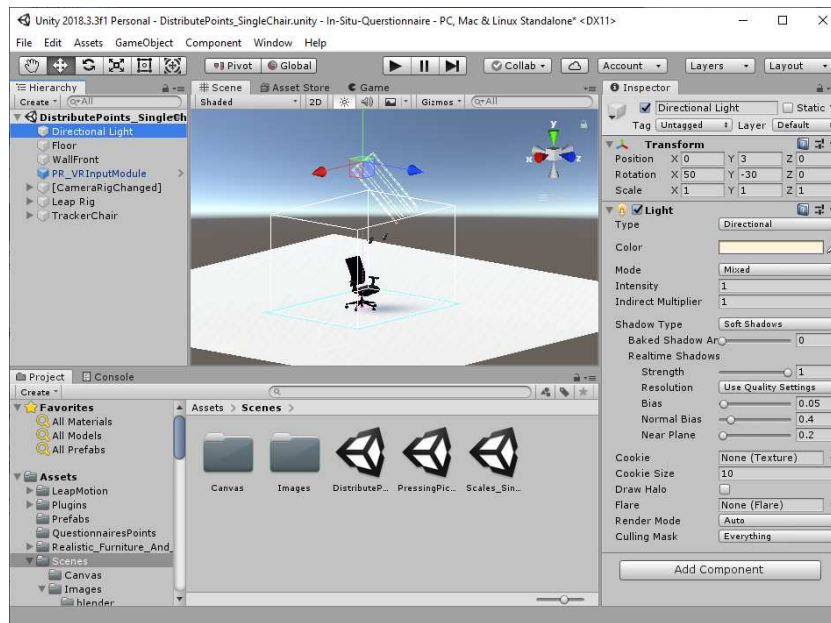
Figure 3.4: Screen-shot of the Unity application.

To program the game object behaviour, Unity's built-in colliders and C# scripts were used.

3.3.1 Colliders

Colliders are invisible components defining the shape of a GameObject for the purpose of simulating physical object collision [Tecb]. However, the shape of the collider does not have to match the shape of the GameObject. In Unity, there are 2D and 3D colliders; the colliders we used can be seen in section 5.2.5.

3.3.2 Scripts

In Unity, scripts are used to respond to input from a player [Tecc] and arrange the timeline of events. Scripts can be used to create a graphical effect or to control physical behaviour of objects. They are written in C# programming language.

We used scripts for instance to show the questionnaire when the users touched the chair (i.e. when hand colliders collided with chair colliders) or when the users interacted with the In-Situ-Questionnaire and the scripts changed what the users saw.

3.4 SteamVR

SteamVR is a suite of tools and services based on the Open Source Virtual Reality libraries [Mur17], providing a platform for VR. Unity communicates with VR hardware,
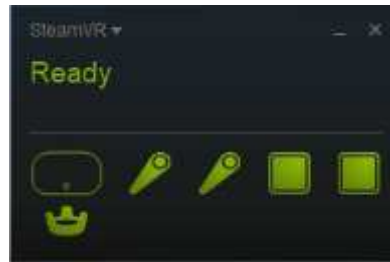
24

Figure 3.5: Screen-shot of the SteamVR status menu.

in our case with the HTC Vive, through the SteamVR system which is an extension of the Steam client. Steam is a video-game distribution service. Figure 3.5 shows a case when HTC Vive headset, controllers, base stations and one tracker are connected to the SteamVR.

An important part of SteamVR, Compositor and Lighthouse tracking, the Chaperone keeps track of where the users are in relation to walls set up by the users. When the users are close to the walls it displays a proximity warning showing a blue grid. The Lighthouse tracks the position of the controllers and the headset.

Even though we do not use the SteamVR directly, it is an important part of the whole system without which the virtual world would not be able to function.

CHAPTER 4

Conceptual Design

Our aim was to develop questionnaires measuring users' experience interacting with virtual objects in VR and finding out whether user perception of the virtual object 3D model corresponds to the real one. This chapter is primarily concerned with the overall design and requirements of the developed questionnaire. The questionnaires were used during user testing while the users were interacting with the real object. We named them the *In-Situ-Questionnaires*. We also developed an environment for evaluation. We tested there the validity of the In-Situ-Questionnaires. The implementation details can be found in Chapter 5.

4.1   In-Situ-Questionnaires Requirements

Objects have different properties, they can differ in e.g. size, weight, shape. These properties can tell us how two objects are different. To study the differences, we chose six properties. Five of them used Simeone et al. [SVG15] to distinguish in which properties real object and its virtual counterparts differed in. We used them in the In-Situ-Questionnaires. The users should be asked whether the real object and its virtual counterpart correspond to each other according to these properties. This is done while they are still engaged in VR and are able to touch and interact with the object repeatedly if they are not sure how to evaluate a property. This option to repeat the experience should ensure the validity of their answers.

The In-Situ-Questionnaires were designed to be stand-alone, which allows them to be used in any other user study to measure user experience interacting with real objects in VR.

In the following sub-sections, we describe the requirements that the In-Situ-Questionnaires must fulfill, and we present arguments on why we think they are necessary. Requirements, which are the same for all three In-Situ-Questionnaires are described first.

4.1.1   Control

How to interact with the VR questionnaire was an important question. The ideal way of interaction is the way how the users naturally interact with the environment. It should be easy to use. This led us to the conclusion that the ideal way is to use hands. It should be the main tool for controlling the In-Situ-Questionnaires. Fingers should be used to push a button, and by that filling them out.

If this hand interaction would not be possible, we should present the user with another option to use VR controllers. However, in the case of VR controllers there is a learning phase, where the users have to learn how to use it and get comfortable using it. This would lead to a distraction of the user from the task. Nonetheless, VR controllers are a sensible alternative as they enable precise selection of any button on the In-Situ-Questionnaires, which changes color when the light beam from the controller hovers over it. To select the button the users then have to pull the trigger.

4.1.2   Size

The size of the In-Situ-Questionnaires is important for the users to easily see that there is a questionnaire, to read labels and task, to understand pictures and interact with the questionnaire in order to evaluate the object. The questionnaire should be readable even when the users are not right in front of it. If its size was like a standard A4 paper, which the researchers usually give to the users, the participants in VR would have problems reading anything written there, as the text would be too small to read comfortably. The same problem would arise if the questionnaire was too big (e.g. the size of a wall) - even if the users were able to read the details, they would not be able to interact with the questionnaire using their hands due to physical limitations in individual user height and reach.

The ideal size of the questionnaire is such that the users are able to move just their arms and hands or slightly step to the side in order to fully reach all UI elements of the questionnaire. The size we found to be ideal is 1 meter width and 0.5 - 0.6 height based on the amount of text and layout of the images.

4.1.3   Evaluated Object Properties

The crucial part of the In-Situ-Questionnaires are the types of object properties which the users evaluate. We were inspired by Simeone et al. [SVG15] where objects were evaluated by their **size**, **shape**, **temperature**, **material** and **weight**. Another property used was **hardness**. The properties and their explanations:

1. Size – the real object can be smaller or bigger than its virtual counterpart.

2. Shape – the overall shape of the object. For instance, the shape would be mismatched if the object is round instead or sharp around the corners

3. Temperature – perceived temperature can change based on the texture or color used [SVG15]. If the user visually perceives an object to be of metallic quality, the intuitive expectation is that it will feel cooler to touch than an object seemingly made out of plastic. Blue-tinted objects seem colder and red-tinted ones are perceived as hotter.

4. Material – haptic sense is essential when a person touches different materials. Glazed stone is smooth, but a wool sweater is bumpy.

5. Weight – the weight of an object can differ for example based on its size or material. If the object looks like it is made out of wood, it would be perceived to weigh more than if it looked like it was made out of plastic.

6. Hardness – the object can be hard or soft to touch. When the object is soft, the users are able to change the shape of the object. Before the pilot studies 6.1, we used the property name *hard/soft*. The name was changed based on user feedback.

4.1.4 Used Colors

In the In-Situ-Questionnaires we used a combination of white, black and gray. These colors are meant to be neutral. If the users had not yet evaluated the properties, we used these neutral colors in the In-Situ-Questionnaires to avoid influencing the user. The color neutrality changes in the Right/Wrong and Scale In-Situ-Questionnaire, which are described later in the section 4.2, when the users start filling out the questionnaire. As the color red is used as a warning, indicative of danger and attention-drawing and generally associated with the meaning *wrong*, we used it as an indicator when the users evaluated that a property was wrong. On the other hand the color green is soothing and generally associated with the meaning *correct*, indicating safety and in our case it means that the property of an object was perceived as correct. Additionally, in the Scale In-Situ-Questionnaire the yellow color on the scale is used as neutral, representing that the property is perceived neither as right nor wrong.

The color red is also used in the Right/Wrong In-Situ-Questionnaire on a label indicating that what the users have done is not allowed. Only neutral colors were used in the design of the Score In-Situ-Questionnaire (Figure 4.4), since the plus and minus signs were used as an indicator of the perception of the correctness of a property.

4.1.5 Types of Pictograms

Figure 4.1: The minus and plus pictogram.

- Minus / plus pictogram (see Figure 4.1)

  The plus and minus sign are commonly used as symbols for addition and subtraction. Therefore, they were used to add or subtract points in the Score In-Situ-Questionnaire.
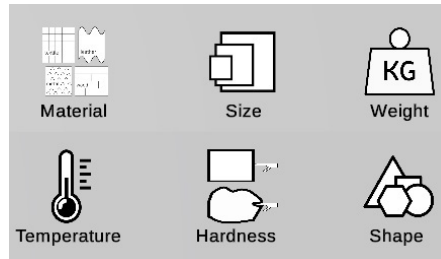

Figure 4.2: User-rated properties.

- Property icon

  The users rate properties described in section 4.1.3 and shown in Figure 4.2. Icons are made in neutral black and white color in order not to evoke any unnecessary feelings in the users. All icons are designed to be a simple and straightforward representation of the property. The material icon is made out of four pictures of different materials (textile, leather, metal and wood) to represent materials the object can be made out of. The rated object can be either smaller, the same size or bigger in reality than seen in the virtual environment. Three differently scaled rectangles represent the size icon. A weight with a shortcut of a kilogram (kg) as a label, which is used to represent how heavy the object is. The temperature property is depicted as a thermometer. If a finger is pushed into a hard object, it will not change its form. However, if the object is soft, it will. This is illustrated in the hardness icon with hard object at the top and soft object at the bottom. The shape icon is composed of a triangle, a rectangle, a circle and a hexagon to represent different types of shape an object can have.

Figure 4.3: Scale.

- Scale

  Figure 4.3 shows a scale used in the Scale In-Situ-Questionnaire (Figure 4.8). It is made out of lines of different lengths and five circular buttons. The color of the circle indicates the extend to which the real object corresponds to the virtual counterpart, with green being the most and red being the least similar. Color selection is described in the section 4.1.4. The five buttons are not shown at the

beginning of the user task session, only the black-line scale is being displayed. Maximum one button is visible at all times.

4.1.6 Saving Users' Ratings

We thought of a submit (or save) button on the In-Situ-Questionnaires, which would save the evaluation data and hide the questionnaire. However, during testing, as touching the evaluated object was used as the trigger to display the questionnaire, the button which previously hid the questionnaire did not prove very useful. Instead, the evaluation from users should be continuously saved into spreadsheet files to ensure preservation of the data when something goes wrong during testing. When the users change the evaluation of an object, it should be automatically saved.

4.2 In-Situ-Questionnaires Design

We developed three styles of the In-Situ-Questionnaire. All use the same set of properties users should evaluate. However, the main difference between them is the description of the questionnaire task and the method of evaluation. In this section we will present the different types and how they differ.

4.2.1 Score In-Situ-Questionnaire

The Score In-Situ-Questionnaire is show in Figure 4.4. At the top of the questionnaire is an introduction text *Compare what you see and feel. Worst rating is -3, the best +3.* It nudges the users to become aware of what they perceive with their senses. The users should touch the evaluated object to realize whether their tactile sense corresponds to what they see. This should provide the users with enough information about what they are supposed to do and about the rating's threshold values. The users should rate properties of a test object. We designed a 7-point scale ranging from -3 to +3, and centered it around 0, which is perceived as a neutral value.

Information about the task was missing in the first version of the Score In-Situ-Questionnaire, which can be found in Figure 4.5. When volunteers read *Distribute 10 points to properties according to their believability. Maximum points per property are 5. Points to distribute: 10*, they did not know what they should rate. As we wanted them to compare what they see with their eyes and touch with their hands, we changed the introduction text. We also changed the length of the introduction text as the initial text was too long for the volunteers. They either skipped reading it or just skim read it without comprehension.

Even though no scale is visually visible in this questionnaire, users should not have problems evaluating the property with the plus and minus icons as they are commonly used as symbols for addition and subtraction, respectively. By using the minus icon, users can give a minimum of -3 point, by using the plus icon, the maximum can be +3. Above the Plus and Minus icon is a label with the current evaluation value, which changes with added or subtracted value, with the text *Points*. The label is located near
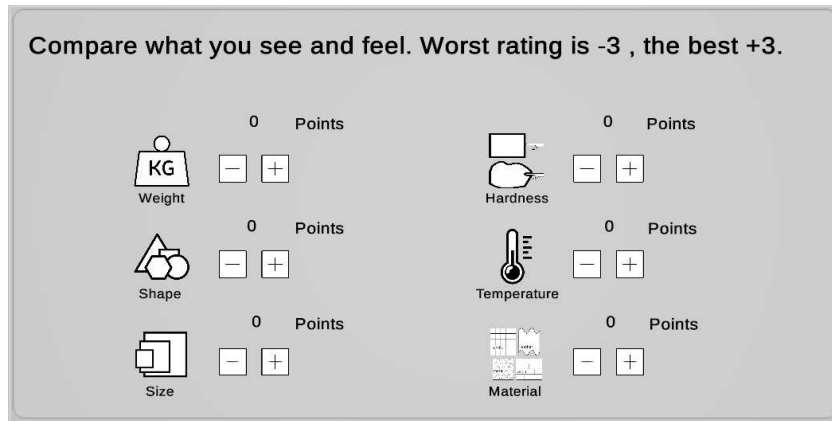
Figure 4.4: Score In-Situ-Questionnaire.

the pictogram icon and plus and minus buttons, based on the proximity principle, where related elements should be placed near each other. This way we visually indicate they relate to the same property.

Pictogram icons are designed to aptly represent individual properties. All six properties are described in section 4.1.5.
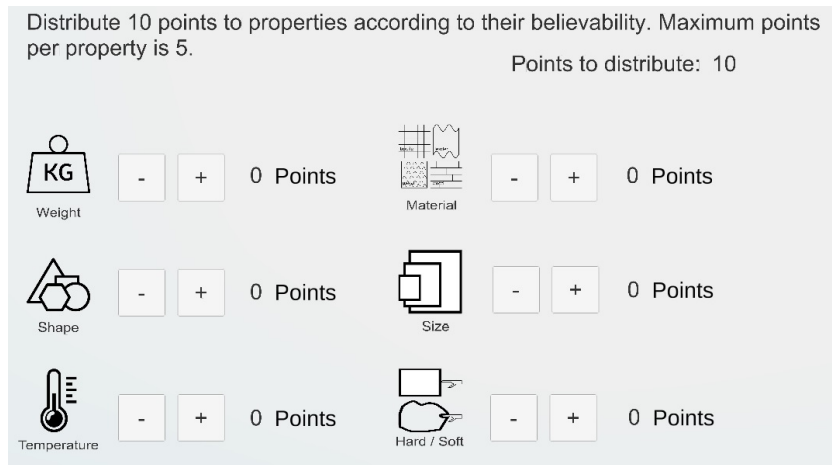

Figure 4.5: First version of the Score In-Situ-Questionnaire.

We used the 7-point scale centered around 0, because we realized the initial idea of having the worst score 0 and the best 5, starting from 0 meant that it was more toilsome to give any property good rating, whereas it started with a bad rating. We changed it to start at a neutral value as we wanted to make it equally hard to rate the object either positively or negatively. With -3 as the worst score, the users had to make mental decision to subtract the points.

We have lifted the restriction of maximum points to asses, which was used in the first

version. The initial idea was to have a constraint so the users would have to weight more, where they assign the points. However, this could lead to users rating not all properties, but only some. Therefore, we changed the way of thinking and preferred to receive ratings of all properties.

4.2.2 Right/Wrong In-Situ-Questionnaire

In Figure 4.6 is depicted the Right/Wrong In-Situ-Questionnaire with introductory text and properties, which were divided into two rows, the *felt right* row and the *felt wrong* row. At first the introduction text was *Press 2 properties that felt right while you were touching an object. Press 2 properties that did NOT feel right while you were touching an object.* depicted in Figure 4.7, which was too long (see Figure 4.7) and the text in the felt wrong row was different only in the added *did NOT* words, which caused difficulties in finding the difference between these two rows. Therefore, we shortened the text to *Comparing what you see and feel. What felt RIGHT, what felt WRONG.* The wording was unified with the other types of In-Situ-Questionnaire and adjusted to the scale we used (e.g. *Comparing* vs. *Compare*).
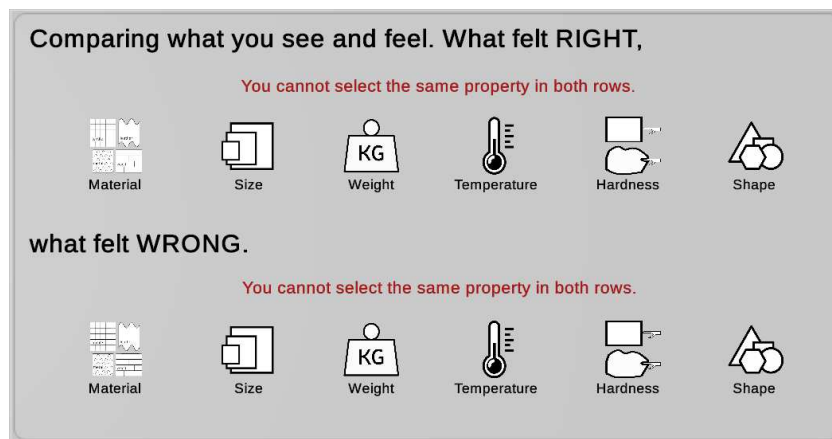

Figure 4.6: Right/Wrong In-Situ-Questionnaire.

Users can select all or no properties per row. The only restraint is that users cannot select the property in both rows as the experience should not feel right and wrong at the same time. If users try to select it, a warning text *You cannot select the same property in both rows.* shows up. We decided not to limit the users in how many properties they could choose in one row (previously two), as in the former case they would not touch upon all the properties, which they considered felt right or wrong. The previous design could have pushed the users to choose a property even if they felt there is no property that felt neither right nor wrong.

If the users choose that the experience of a particular property felt right, the property in the first row should be selected. The property would turn green. If the experience felt wrong, the property should be selected in the second row, it would turn red. The users

do not have to rate any property if they have no particular feeling about that experience. The choice of colors is decribed in the section 4.1.4.
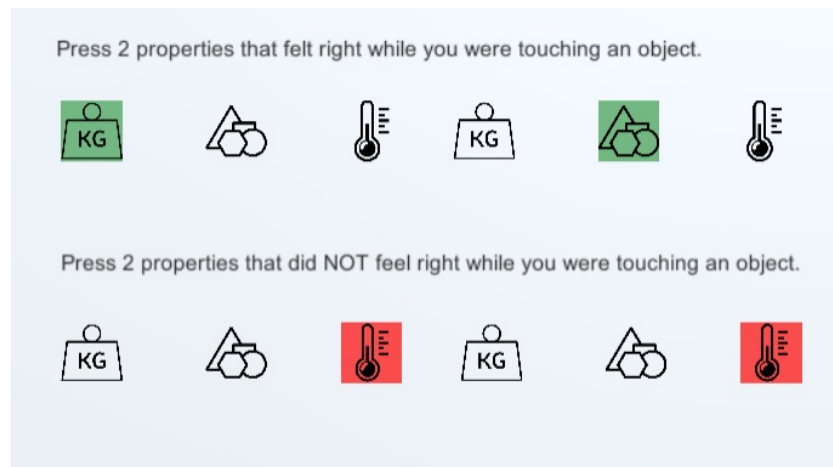

Figure 4.7: First version of the Right/Wrong In-Situ-Questionnaire

We designed the Right/Wrong In-Situ-Questionnaire more dichotomously than the Score In-Situ-Questionnaire as the values they can choose from are either felt right or felt wrong. However, there is a hidden neutral value, when the users do not choose to select anything. The dichotomous decision should give us information whether it felt right or wrong, however, not how much. We wanted to find out whether and how their answers would change based on the change of the scale and its depiction. Unlike the scale in the Score In-Situ-Questionnaire, the scale in the Right/Wrong In-Situ-Questionnaire is not only black and white, but can also be green or red based on the users' selection.

4.2.3 Scale In-Situ-Questionnaire

The introduction text of the Scale In-Situ-Questionnaire is simpler than in the other In-Situ-Questionnaires, as the scales visually nudge the participant how to evaluate the properties, what is the best and the worst rating. The scale is a 5-point scale with circles, a sad smiley representing the worst rating and a happy smiley representing the best rating as people associate them with something that is negative or positive, respectively. Based on the rating, the circles are green, yellow or red. If the best rating is 1 and the worst rating is 5, then 1 and 2 are green, 3 is yellow and 4 and 5 are red. The color choice is explained in section 4.1.4.

Only a black-line scale is displayed when the users had not yet rated a property. One colored circle is shown when the users evaluate the property.

Figure 4.9 shows an earlier version of the Scale In-Situ-Questionnaire. When compared to the final version in Figure 4.8, only two main changes are visible. The first change is the removal of the evaluation numbers above the scale. The pilot users assessed it as
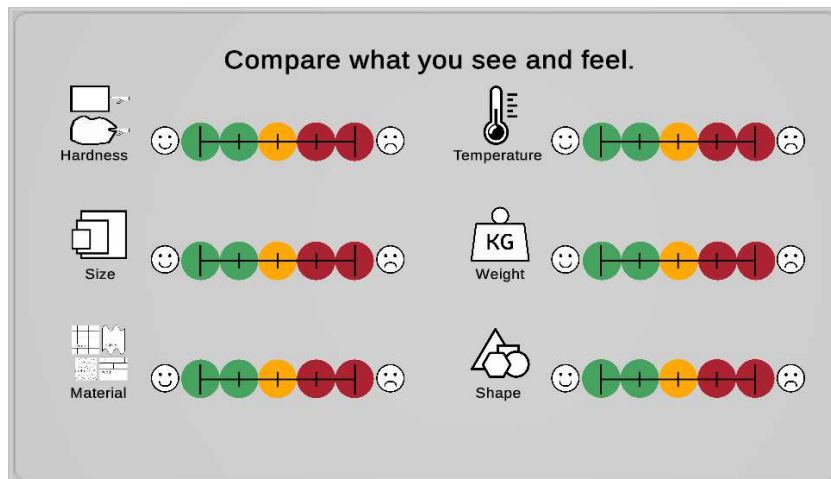
Figure 4.8: The Scale In-Situ-Questionnaire.

confusing and not adding any meaning as 5 can represent both the best and the worst rating.

The second change is the removal of a *submit* button along with a warning text for when the users have tried to submit the questionnaire without evaluating all properties. The users should have pressed on the **submit** button when they were finished with the evaluation, however, most of the time they did not to use it. The reasoning behind why the submit was not used is also analyzed in section 4.1.6.
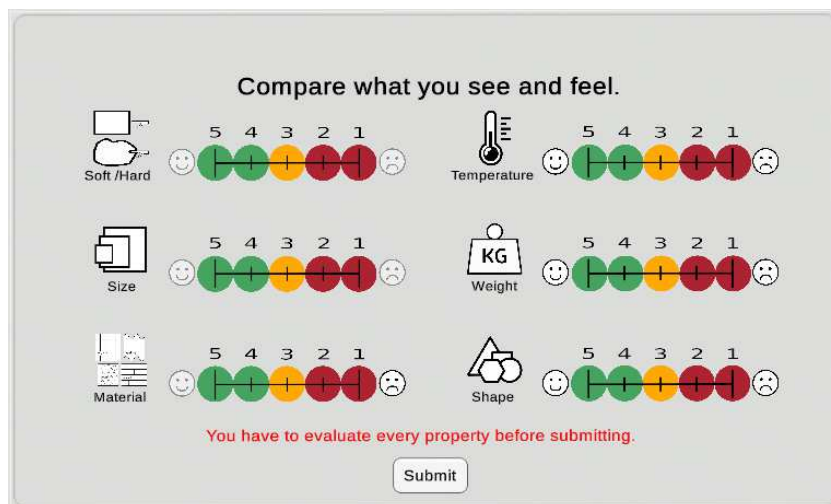

Figure 4.9: Earlier version of the Scale In-Situ-Questionnaire

Although colored circles are not displayed at the beginning, the users can anticipate that something will appear after rating an object. Except for the introductory text, everything

is shown visually. With this design we wanted to see if a mostly visual depiction changes the way the users rate properties.

4.3 Environment for Evaluation

To evaluate the In-Situ-Questionnaires, we needed to test it in VR. This constituted the need to create a VR environment, where the users would interact with real objects and their virtual counterparts and later evaluate this experience using the In-Situ-Questionnaire. This section describes the requirements for the environment and its objects.

4.3.1 Environment Design / Scene

As the main part of our study was to gather data about an object to evaluate the validity of the In-Situ-Questionnaires, we focused our attention on these two components of the environment. Therefore, we wanted to make the VR scene as simple as possible with as little distractions as possible. The scene should contain only the object the users evaluate and one of the In-Situ-Questionnaires. Another requirement was for the space where the users found themselves to be large enough, so they could move freely and would not feel constrained by the size of the environment.

The room and the scene where the user study takes place, have to be large enough so that when the objects are being changed, nothing would bump into the users. When the users are observing and touching the object, they need to have enough space around it so they can move without restrictions. This restriction concerns both, the real and the virtual world.

4.3.2 Interactive Object

To evaluate the validity of the In-Situ-Questionnaires, three different objects of the same type should be used. A 3D model of one of these objects should be shown in VR. It should resemble the actual object as close as possible. The other two objects are either similar or different from the object that has a 3D model. For all objects, the same 3D model is shown in VR. If the object is the same as its virtual counterpart, we expect the evaluated properties to get the highest score, the similar object should have a slightly worse score and the different object should have the worst score. This should prove the validity of the In-Situ-Questionnaires.

An earlier approach would have all three objects in the same scene next to each other. However, later we decided to use just one object and one questionnaire in one scene in order to keep the evaluation straightforward and avoid influencing the results for the different questionnaires.

The most suitable object for our purpose proved to be a chair. More details on why we chose the chair can be found in the section 5.3.5.

4.3.3 Tracking

For interaction, tracking was necessary, ideally real-time with minimum delay and fluid movement.

4.3.3.1 Object Tracking

As the users were asked during the user study to interact with the object, they were allowed to move it. The object could not stay at the same spot in the VR scene the whole time. We needed to ensure that the position of the real and virtual counterpart is the same. Therefore, the real object needed to be tracked as robustly as possible, so the virtual model could move along with it. The tracking should not only show just where the object is, but also its rotation depending on each user's actual interaction. This can be accomplished with the HTC Vive tracker described in the chapter 3.

4.3.3.2 User Tracking

To show what the users are looking at and where they are in the scene, we needed to track the users as well. This can be done with the HTC Vive headset, which communicates with the computer and sends information about its position.

4.3.3.3 Hand Tracking

The optimal VR solution for an interaction between the user and the object is to visualize the user's hands, as people are used to seeing their hands while interacting with their environment. Looking at an object in VR to interact with and not seeing their hands would feel out of place, and they could have problems with hand-eye coordination e.g. when grabbing an object. Therefore, we would like to track the users' hands, so they could see them visualized in VR. The virtual representation of their hands does not need to be a perfect representation - a simplified model is sufficient to understand where their hands are. We can do this by using the LEAP Motion, which tracks not only the position of hand, but also the position of each finger.

4.3.4 Position of the In-Situ-Questionnaires in the Scene

At first, we have tried to position the In-Situ-Questionnaires in the middle of the room where it would be located the whole time. However, when we interacted with the object, we moved it and the result was that when we wanted to fill out the questionnaire, we had to approach it. If the questionnaire was too far away or too close, users would be unable to read the details and to easily evaluate the object. Therefore, the questionnaire was restricted to be positioned near the object the users are interacting with. Users should be able to reach out and fill out the questionnaire for the sake of ease of use and minimal distraction. When the object, which users are evaluating is moving, the questionnaire should move with it and be positioned within the same distance and angle from the object. Vertical positioning and the questionnaire facing users are recommended. However, when

the users are moving around the object the questionnaire should stay anchored to the object. In this way users can view the questionnaire under different angles.

4.3.5 User Task

A task is needed in order to engage the users in VR and to fill out our questionnaires. However, with the description of the task, we should not give away the aim we want to accomplish, we just want to make them familiar with the object. Not just to look at it, but also to touch it, play with it in order to make them familiar with its properties, so they could answer the In-Situ-Questionnaire to the best of their knowledge. Even if after examining the object, they still cannot answer the questionnaire, they have the object nearby and can examine it again at their leisure.

The task should not be too difficult, as our goal is to test the In-Situ-Questionnaires and not to test whether users understand what we want from them. As objects we used are chairs, to make them interact with the object, the task could be to sit down on them. The users first have to touch the chair to make sure it is there and could carry their weight and after that they could sit down on them.

4.3.6 Motion Sickness Problem

Some people may experience motion sickness when in VR, which could influence the result of the experiment. Therefore, during user studies, there is a need to make sure the users are feeling well and the results would not be influenced by motion sickness. This could be accomplished by a simple question like in 2.7 after each test session.

CHAPTER 5

Implementation

This chapter starts with the description of the architecture of the whole system to give a brief overview of how the individual parts of this thesis come together. Later on, a description of the implementation of the In-Situ-Questionnaires and the virtual environment where we tested the In-Situ-Questionnaires are provided.

5.1 Architecture

This section briefly describes how the hardware, the software, the In-Situ-Questionnares and real objects are put together, which is shown in Figure 5.1.

The physical objects were used to track users (HTC Vive), their hands (LEAP Motion controller) and the real object (HTC Vive tracker). In our case by HTC Vive we mean a set, which consist of HTC Vive headset, HTC Vive base stations and, in case the LEAP Motion would not work, HTC Vive controllers. The HTC Vive tracker is attached to the real object, in our case, the chair. It sends information about where the chair is positioned through StreamVR to the Unity, which shows the virtual representation of the chair in the corresponding position. Unity also displays the In-Situ-Questionnaire and the environment. The users can interact with the In-Situ-Questionnaire either with their hands (LEAP Motion controller) or with the HTC Vive controller. The LEAP Motion controller streams image data to the computer, where the LEAP Motion Service processes them to reconstruct a 3D hand representations [LMb].

Unity, in our case specifically Unity version 2018.3.3f1, is a cross-platform game engine which we used for creation of the In-Situ-Questionnaires and environment for evaluation. In Unity, we have created three scenes which differ only in which In-Situ-Questionnaire is present. To be able to use the HTC Vive with Unity, we needed to use a SteamVR plugin which ensures the loading of 3D models for VR controllers, and handles their input and

interaction with Unity UI elements thanks to its interaction system [Ste]. The SteamVR also communicates with HTC Vive base stations, controllers, the headset and the tracker.
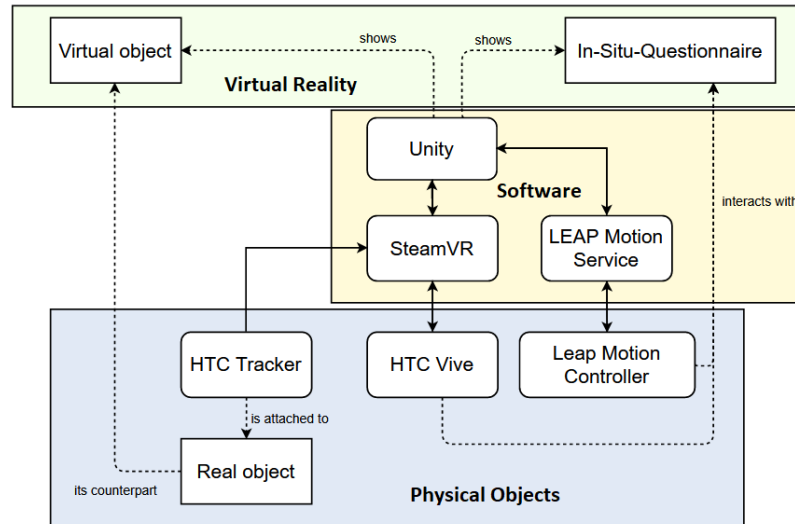

Figure 5.1: Architecture of the system.

We used the HTC Vive headset to display the virtual scene to the users. Thanks to the LEAP Motion, the users are able to see the position of their hands, as their fingers can interact with buttons on the In-Situ-Questionnaire. There are multiple modules of the LEAP Motion Orion SDK [LMa] for Unity, such as the Core, the Interaction Engine, the Graphic Renderer or the Hands module. For our purposes, the Core module was sufficient as we did not require the Interaction Engine's object grasping functionality. The Core is a necessary module, as it provides the foundation for VR applications. It renders a basic set of Leap hands and attaches objects to hand joints. For our project, only hand rendering was implemented.

For more information about HTC Vive, HTC Vive tracker, LEAP Motion, SteamVR and Unity see chapter 3.

5.2   In-Situ-Questionnaires

The In-Situ-Questionnaires are designed to be used in the VR, therefore, we implemented it in the Unity game engine. We used Unity's components such as Canvas, Labels, Images, and Buttons to assemble the In-Situ-Questionnaires designed in chapter 4.
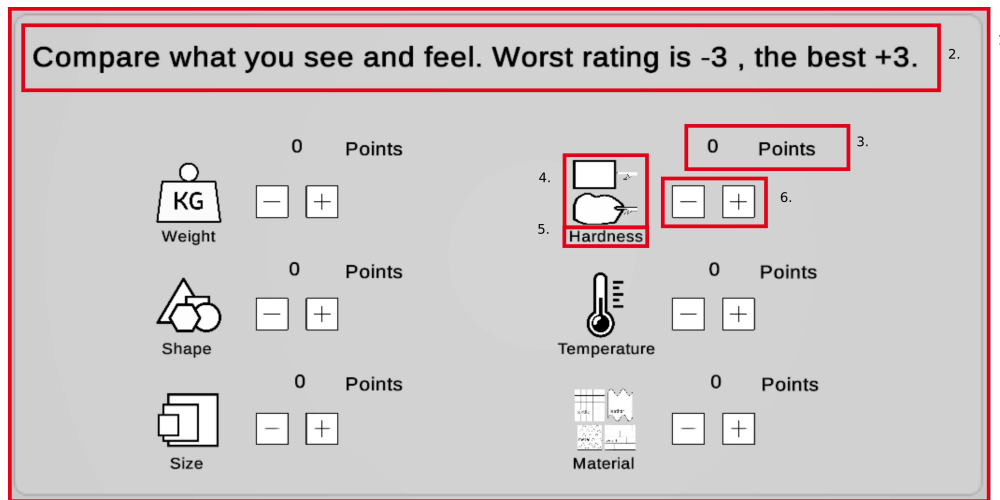
Figure 5.2: Score In-Situ-Questionnaire - 1. Canvas, 2. Introduction text, 3. Points counter, 4. Property icon, 5. Property label, 6. Buttons for adding or subtracting points.
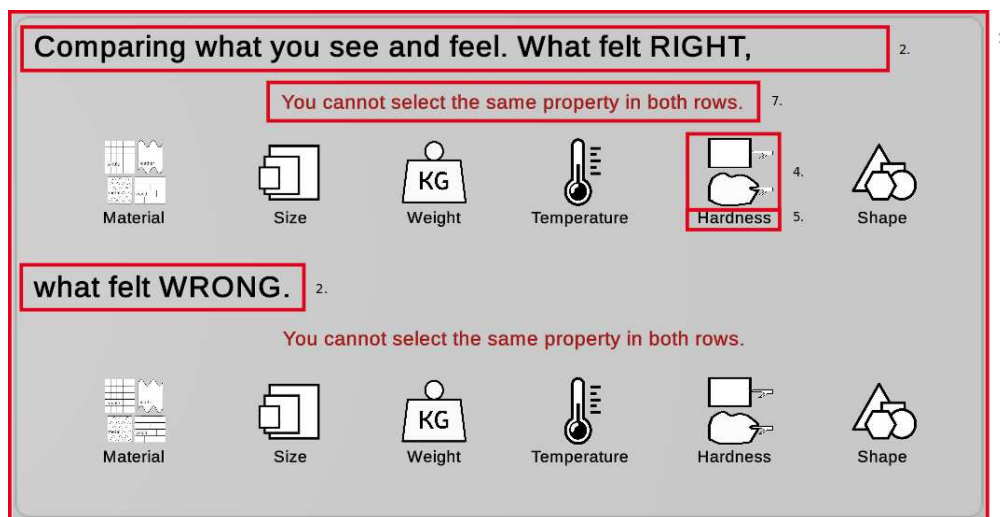

Figure 5.3: Right/Wrong In-Situ-Questionnaire - 7. Warning label.

5.2.1 Canvas

Canvas (see Figure 5.2, 5.3 or 5.4, 1. Canvas) is a mandatory Game Object [Teca] for every UI element that is placed inside as a child GameObject. As in our case, the Canvas is rectangular. To create the In-Situ-Questionnaires we added text, images and buttons inside the Canvas. A simple C# script is attached to the Canvas, which hides the whole In-Situ-Questionnaire at the beginning of the VR session. The In-Situ-Questionnaire is shown later when the users' hands collide with the chair. Details of the testing process can be found in section 6.2.2
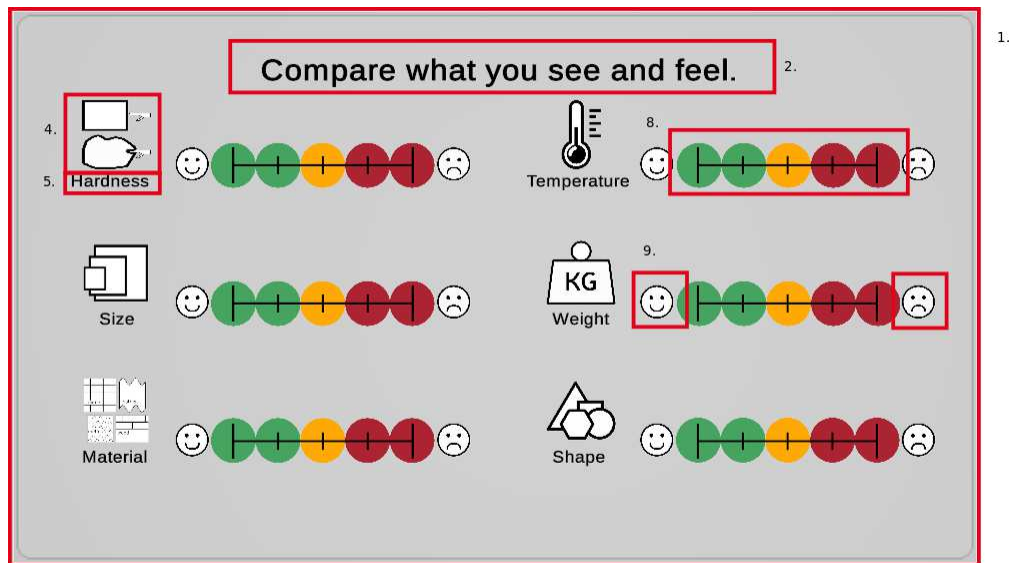
Figure 5.4: Scale In-Situ-Questionnaire - 8. Scale 9. Emoticons.

5.2.2 Text / Labels

We used the UI TextMesh Pro [Tecd] element to display text on the canvas. Not all text or labels are used for all the In-Situ-Questionnaire types. We reference the In-Situ-Questionnaires figures based on whether they use the specified text or label. We used the TextMesh Pro for:

- Introduction text (see Figure 5.2, 5.3 or 5.4, 2. Introduction text)

  Each In-Situ-Questionnaire contains written information about the task what the users should evaluate (e.g. *Compare what you see and feel.*) and also contains a hint on how to evaluate (e.g. *Worst rating is -3, the best +3*) except for the Scale In-Situ-Questionnaire, which uses scale where the users can discern at a glance that the scale has five level and the worst and best evaluation is marked by sad and happy smiley, respectively.

- Property label (see Figure 5.2, 5.3 or 5.4, 2. Property label)

  To make the meaning of the property icons more understandable, we labeled the properties underneath these icons.

- Points (see Figure 5.2, 3. Points)

  The point count serves as information for the users on how they evaluated a specific property. The count starts at 0 and can have a maximum value of 3 and a minimum value of -3.

- Warning label (see Figure 5.3, 7. Warning label)

Warning label is hidden unless the user tries to select the same property in both rows, which triggers a C# script to show a message (*You cannot select the same property in both rows.*). The message is shown for 5 seconds to inform the user.

The introduction text and the icon description do not change during the user studies. They are the same for every user. However, the point value changes when the users click either plus or minus button. If they click on the plus or the minus button, one point is added or subtracted to the points value respectively. The warning label does not change, but is shown only when triggered.

5.2.3 Property Icons / Images

The Unity element Image (see Figure 5.2, 5.3 and 5.4, 4. Property icon) was used to display all property icons described in the chapter 4.

5.2.4 Buttons

Buttons are interactive, which for instance were used to change the score by attaching C# scripts to the click (e.g. Figure 5.2, 6. Buttons for adding or subtracting points – the plus button has a C# script attached that gets the number of points of a corresponding property, adds a point and rewrites the number of points). One of the button properties in Unity is the ability to have images placed on top of them. We used this property on all buttons to make the interface more user-friendly.

- Minus / plus button (see Figure 5.2, 6. Buttons for adding or subtracting points)

  To add or subtract points in the Score In-Situ-Questionnaire, the users have to click on the buttons with a plus and a minus sign respectively. The updated point score is shown above the plus and minus signs.

Figure 5.5: Image on the left is a property image, which is not selected. In the middle and on the right is a selected image of a property that felt either right or wrong to the users, respectively.

- Property icon (see Figure 5.5)

  If the properties, which are implemented as buttons, in the Right/Wrong In-Situ-Questionnaire are selected, their color changes either to green or red, depending on whether this property felt right or wrong to the users, respectively. If the button is selected again, it turns to its original (black and white) color and is not considered

selected. Properties in the Right/Wrong In-Situ-Questionnaire do not have to be selected before submitting the questionnaire. In case a property is highlighted neither green nor red, this indicates that users had neither a positive nor a negative experience with the particular property.

- Scale (see Figure 5.4, 8. Scale)

  The scale used in the Scale In-Situ-Questionnaire consists of five circular colored buttons, which are not shown at the beginning of the user task session. Only the black-line scale is being displayed. When the users collide their point finger with one of the points on the scale, the respective colored circle appears. We take the position of where the users collided with the In-Situ-Questionnaire and we display the button based on which one is the nearest. Maximum one button is visible at all times. So when the finger collides with another button, the first disappears and the second becomes visible.

5.2.5 Colliders
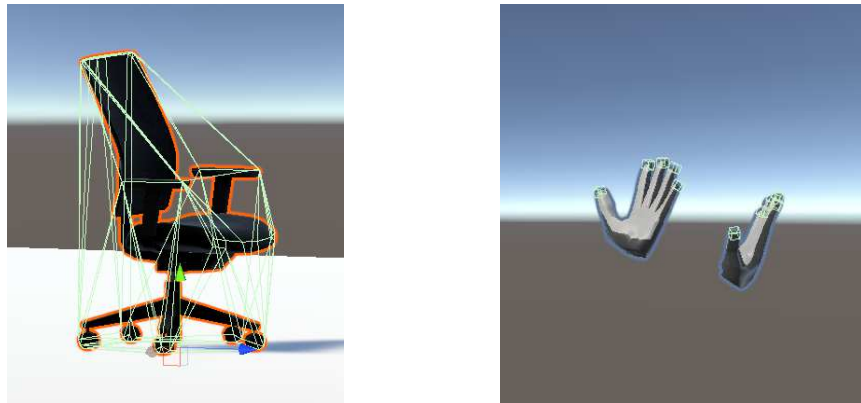
(a) Chair with multiple colliders of different size and shape depicted with green color. (b) On each finger there is a green cube which represents one collider.

Figure 5.6: Chair and hand colliders.

Colliders were a crucial component of Unity for our implementation. In our case colliders are attached to our test object (see Figure 5.6a), to all fingers of both hands (see Figure 5.6b) and to buttons (see Figure 5.7). Colliders on the fingertips have a shape of a small cube. If any fingertip collider collides with another collider (specifically with colliders attached to the chair or to the In-Situ-Questionnaire) they trigger a C# script in Unity, which further processes the interaction.

Not everything in the In-Situ-Questionnaire is a collider. For our purpose, we chose buttons to trigger events. We attached colliders to them, so the users could select one button when the finger intersects with button boundaries. The collider is as wide and tall as the button it is on (see Figure 5.7). At first, we gave almost no depth to the collider,
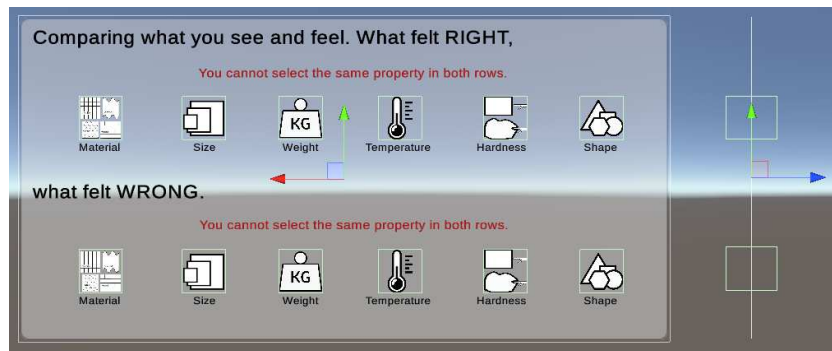
Figure 5.7: In-Situ-Questionnaire colliders from the front and side-view. The colliders depicted as green cuboids can be seen as rectangles from the front and also from the side.

but even though the buttons are just 2D, the colliders did not fire. When we increased the depth to be as deep as the height or width, colliders started to fire and in this way the users could click the button.

Similar collision as with buttons and fingers can occur with the chair and fingers. However, the chair has multiple colliders applied to it. The reason for multiple colliders is the chair being constructed from multiple parts. We needed colliders for the seat, back-rest and arm-rest. We were able to match the various chair parts using only a handful of colliders which resulted in a sufficient collision model approximation. We needed the colliders to reveal the In-Situ-Questionnaire when the user's hand was close to the chair.

5.2.6   Size

The size described in the section 4.1.2 proved to meet the conditions when implemented.

5.2.7   Control

Buttons on the questionnaires can be controlled either by the hand, by the means of LEAP Motion, or using the HTC Vive controller. Using the hand for evaluation is more natural than using the controller. However, we implemented also the HTC Vive controller for use cases where no hand tracking is available, for instance in case of failure of the LEAP Motion.

- Hand

  To use hands for interaction with the In-Situ-Questionnaire one can use any finger for pressing a button on the canvas. The finger and button need to collide in order to register an interaction.

- Controller

  We implemented a laser pointer, which is a virtual light beam coming from the controller visible for 1.5 meters with a small sphere at the end. If the beam intersects

with any button on the questionnaire, the button changes color and the users can click the button by pressing the controller's trigger. Because of the reach of the light beam the users do not have to be near the questionnaire.

5.2.8   Property Evaluation

This sub-chapter describes how to fill out the In-Situ-Questoinnaires, how to select the properties, evaluate them on a scale or assign points based on the In-Situ-Questionnaire used. For simplicity, we presume that the users are using their hands to fill out the In-Situ-Questionnaires. If users want to use the controller, the difference is that instead of the hand they would use the light beam from the controller to aim at the desired button (icon of property/ scale / +-) they want to click on and press the trigger on the controller.
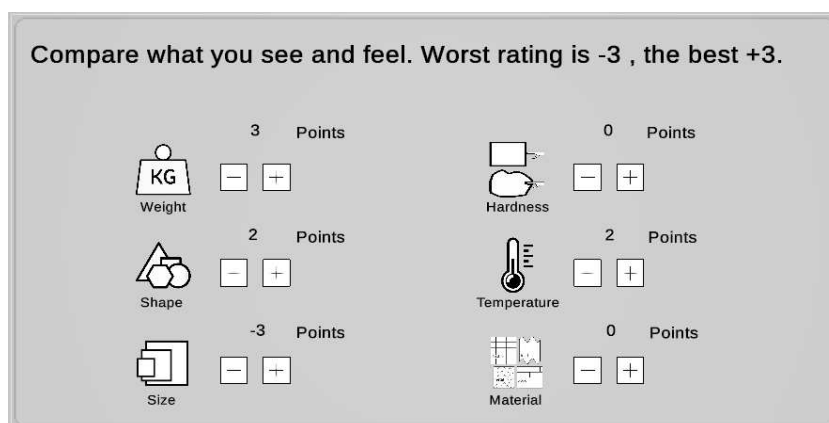
1. Score In-Situ-Questionnaire

Figure 5.8: Possible property evaluation where Weight has the maximum score and Size has minimum score.

The worst rating is -3 and the best rating is +3 for the Score In-Situ-Questionaire. The score starts at 0 which is a neutral score and should not influence the users. To subtract or add a point, the users should click on the minus button or plus button next to the property icon, respectively. The score changes based on the button clicked. If the score is minimum or maximum and the users click on minus or plus button respectively, the score will not change as the scores are on the limit of possible values (see Figure 5.8).

2. Right/Wrong In-Situ-Questionnaire

One property can either be chosen that it felt right, wrong or if it is not selected, it felt neutral. If it felt right, the users should select the property in the first row. This property would turn to green. If it felt wrong, the users should select the
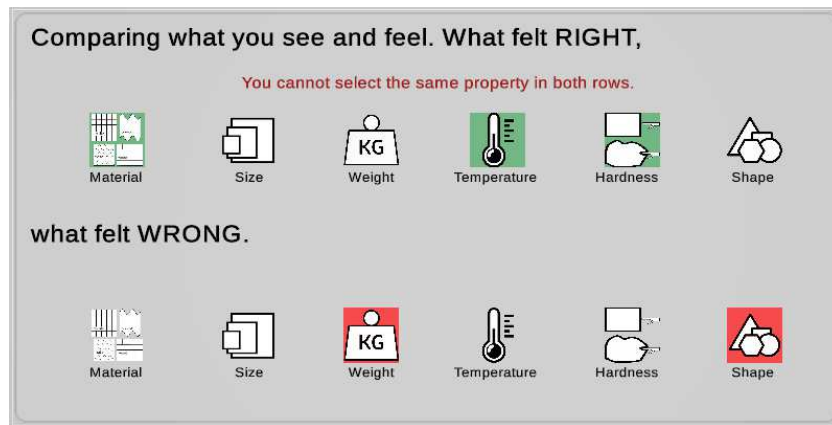
Figure 5.9: Material, Temperature and Hardness were selected that they felt right. Weight and Shape felt wrong. As the message *You cannot select the same property in both rows.* is shown in the first row, the users have tried to select either Shape or Weight in the first row.

property in the second row. This would turn the property to red color. The users may not select the property in both rows. If the property is already selected in one row and the users try to select the same property in the other row, a message would show up for 5 seconds. The text in the message is *You cannot select the same property in both rows.* (see Figure 5.9, as the message is shown above the first row, the users could have tried to select either Shape or Weight in the first row. However, they were not able to as these properties were already selected in the second row).

3. Scale In-Situ-Questionnaire

   To evaluate properties in the Scale In-Situ-Questionnaire, the users should touch the desired point on the scale (see Figure 5.10). A colored circle would show up. If the users select 1 or 2 on the scale, the circle will be green, if 3 yellow, 4 or 5 red. If the users change their mind, they can select another button. The selected circle will show up and the old one will disappear.

5.2.9 Saving Ratings

The In-Situ-Questionnaires are automatically saved during the evaluation with every score change and at the end of every VR task session. The score is saved in .xls files. However, for every type of the In-Situ-Questionnaire, we needed to slightly change the format. As e.g. the Score In-Situ-Questionnaire has 7-point scales ranging from -3 to +3 and the Scale In-Situ-Questionnaire has 5-point scales ranging from 0 to 4.
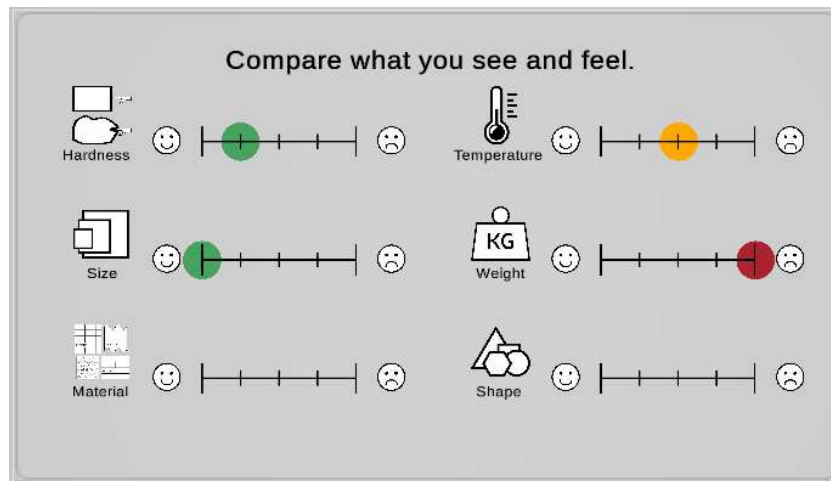
Figure 5.10: On the 5-point scale Hardness has 2, Size has 1, Temperature has 3 and Weight has 5. Material and Shape have not been evaluated yet.

5.3 Environment for Evaluation

Even though the purpose of this thesis was to develop the In-Situ-Questionnaires, we needed to create a VR environment for evaluation as we needed to test the questionnaires we created. The VR environment was made in Unity. We made sure not to add any additional content like furniture or textures, in order to avoid disturbing the users so they can concentrate on the object and the In-Situ-Questionnaire. In this chapter we will describe how the environments were made and how they behave.

5.3.1 Environment Architecture

The scene where the users used In-Situ-Questionnaires, is simple, with a white floor and one white wall. Users can see the Unity's default skybox. In the middle of the scene are the LeapRig and the CameraRig positioned. LeapRig is a part of Unity Core Assets 4.4.0 for Leap Motion Orion Beta. It is a prefab which can be dragged into the scene to be used immediately. The LeapRig displays the hands, enables the users to see how they move their hands and interact with objects.

CameraRig is a VR prefab which enables the users to see the virtual world based on where they are standing. Both LeapRig and CameraRig need to be positioned in the middle of the scene, and have their default position and scale. Otherwise, there would be a discrepancy between real and virtual movement. For instance, if the users' hand size was enlarged, meaning that the scale of the LeapRig was increased, and if the users moved their hand 1 cm left, the virtual hands would move for instance twice as far.

Figure 5.11 shows together with the CameraRig and the LeapRig also a 3D model of the rated object together with the In-Situ-Questionnaire. If the virtual chair moves, the In-Situ-Questionnaire moves along with it. In the Unity scene the virtual HTC Vive

Figure 5.11: VR environment. LeapRig is shown as hands and a purple dotted line. The CameraRig is the turquoise rectangle on the floor together with the white lines forming a cube.

tracker is placed in the middle of the scene and attached to the virtual chair. However, the position and rotation of the virtual chair in the Figure 5.11 is just illustratory as the virtual tracker is in the middle of the scene and the virtual chair is positioned based on where it is attached to the virtual chair. Nonetheless, when the scene is started, the 3D model of the chair together with the 3D model of tracker and the In-Situ-Questionnaire change their position based on the position of the real tracker attached to a real chair.
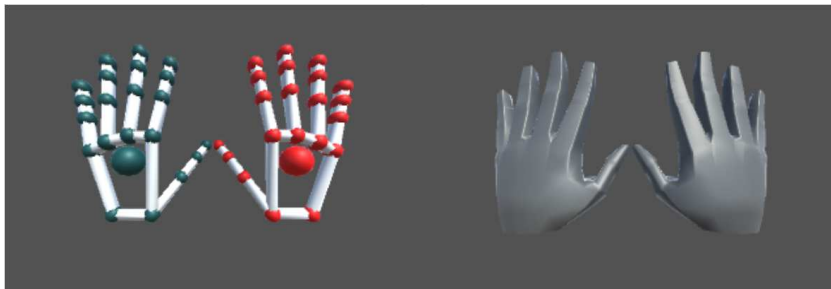
5.3.2 Hands

Figure 5.12: Capsule hands (left) and rigged hands (right) [LMa].

To simulate hand-object interaction, we chose to use the LEAP Motion for hand-tracking. This allows the users to see their hands in the VR and then move their hands according to where the object is positioned to touch it.

With the LEAP Motion hands, users are able to move each finger separately and make gestures in VR. Even though sometimes the wrong finger is bent in VR, most of the time

it is a good representation of the users' hands. We used the LEAP Motion to display just the hands, not the object itself. For interaction, we used colliders which are part of the Unity.

When everything is set up, the LEAP Motion hands work well. However, there are moments when the users can not see the representation of their hands in the VR. When the hands get occluded from LEAP Motion's sensor vision, they either disappear from the VR or get stuck and behave unpredictably. The users have to keep their hands close to the LEAP Motion controller as it has a limited view angle (150°and 120°)[LMb] and depth (80 cm).

To represent users' hands we could have used either Rigged or Capsule hands (see Figure 5.12). Capsule hands are made up from multiple parts and when used by some test users, they were described as a bit foreign. For that reason we chose to use the Rigged hands. In contrast to the Capsule hands, the Rigged hands felt more natural to our test users.

5.3.3   In-Situ-Questionnaire in the VR Environment

Figure 5.13: In-Situ-Questionnaire placement relative to the chair.

The In-Situ-Questionnaire is not visible in the environment at the beginning as we want the users to concentrate on the object first. They should look at the object and interact

50

with it. Once they touch the object, the In-Situ-Questionnaire will show up. Then the users can fill it out.

The choice of position of the In-Situ-Questionnaire was not straightforward. Even when we positioned the In-Situ-Questionnaire near the object, it was not always easy to fill it out. At first, we have tried to position the questionnaire to the right of the object. But when the users in the pilot study 6.1 tried to fill it out, it still did not feel natural, because they had to move themselves to the right and had difficulties to touch and interact with the object to properly evaluate it.

As we have chosen our object to be a chair, the best option offered itself. When the users sit down on the chair, the In-Situ-Questionnaire appears in front of them (see Figure 5.13). However, the object does not have to be a chair. The task can be changed and the In-Situ-Questionnaire can be easily relocated. Apart from the chair we experimented with a bottle. The task for the bottle could be to pretend that the users are drinking from it and the In-Situ-Questionnaire's position could be fixed in a way that the angle at which the bottle is hold would not influence the angle of the In-Situ-Questionnaire, but the In-Situ-Questionnaire would be shown to the right of the bottle at eye level.

5.3.4 Tracking

At the beginning we intended to use LEAP Motion not just for hand representation but also for object tracking. There was an option to attach the object to the hand using attachment hands, but after attaching the object to the hand the virtual object did not behave as the real one. For example when we used a cube, and we grabbed it, we naturally rotated it a bit, but with LEAP we could not tell to which side the cube rotated. In this case we could not change the virtual representation accordingly. Another problem we experienced was when we attached the object to one hand, we had troubles detaching it. We have tried detaching it when the users extended their fingers, but it did not work as expected either. Another problem was how to determine where the object is when not attached to the hand, but has been moved e.g. accidentally or with another body part. Because of the multiple problems, we decided to try a different approach.

To track the object and see it in the VR we could have used markers, optical tracking or tracking based on a color. However, when the users touch an object the marker is on, users hands occlude the marker and the object becomes invisible or stays frozen as it can no longer be tracked. The optical tracking changes how the real object looks as researchers place small balls all over the object. We needed the object to look and feel approximately the same. If we used tracking based on color, we may have had problems with rotation. As we needed the users to interact with the object and they could rotate it, the tracking could be problematic.

Discarding these possibilities, we chose the HTC Vive tracker. When the users touch the object but do not occlude the tracker from both base stations, the object is still tracked unlike when using markers. With the tracker we get information, not just about the position, but also about the rotation angle. Even though the tracker is relatively big

and heavy, it does not have to be in multiple places all over the object like the optical tracking.

However, there is a disadvantage of changing the weight and shape of the object. Nonetheless, this can be minimized when the tracked real object is bigger and the tracker is placed where there is a minimal probability of users touching it.

Another problem we experienced with the HTC Vive tracker was finding a proper place where to attach it. If we would have attached it not parallel or perpendicular to the floor, there would be a need to measure the angle of the leaning part and adjust the model in the Unity scene accordingly. However, the problem with attaching was not only in the angle. We tried to attach the HTC Vive tracker with a tape. However, the attachment was not ideal as it caused some occlusions and there was a possibility that the tracker would fall down. What proved to be suitable were glue pads. We used a small piece, heated it up a bit, sticked it on the tracker and with small force, attached it to the object. There was an advantage of not worrying about occluding the light sensors on the tracker like with the tapes.

5.3.5   Chairs

The main disadvantage of the HTC Vive tracker is a loss of tracking because of occlusion when the LED light from the base stations can not reach the HTC Vive tracker. This limits the choice of possible objects for interaction. Even in case of a transparent object with the tracker placed inside of it, the tracked position of the object was not correct or it was not tracked at all.

In order to find a suitable object for interaction for our evaluation, we tried several alternatives. A very important limitation was, that we had to find not only one object, but a physical object which could also be used in three variants of different *realism*, meaning that one variant should look the same, the other variant would be similar and the last variant would be different to the 3D model. We tried smaller objects like a bottle or a food container and bigger objects such as a chair or a table. We also had to make sure, the object was suitable to be used together with the tracker and had to try whether the tracker was tracking or not.

The object was positioned in the VR environment based on the position of the real counterpart, taking into account the position of the HTC Vive base stations and the CameraRig in the scene. Another important step was to find a good 3D model of the object. We have found a quality model of the chair and made one for the bottle.

We chose to use a chair as during test with a bottle tracking was not stable when users moved their hands. Users kept occluding the tracker during Pilot studies (described in chapter 6). The chair occludes the HTC Vive tracker only when the user is behind the chair and this happens not often. For every object we found a spot which would not be occluded most of the time.

(a) Chair model.

(b) Identical chair.

(c) Similar chair.

(d) Different chair.

Figure 5.14: Chairs used for evaluation.

In Figures 5.14 are real chairs and 3D model, which were used in the user study (section 6.2). The 3D chair model was displayed to the users while they were in the VR. The users evaluated all three real chairs with one of the In-Situ-Questionnaires. The identical chair was an office chair on wheels with a back-rest, arm-rest and was made out of textile and plastics, the similar chair was a wooden chair without arm-rest and the different chair had three legs, adjustable weight and a seat made out of synthetic leather.

CHAPTER 6

Evaluation

User studies are an important part of this thesis as they are used as a verification of the In-Situ-Questionnaires and to distinguish how good they measure user experience in the VR.

In this chapter we will first describe two pilot studies done before the actual user studies, which will be described afterwards. We will summarize a course of the test sessions, describe the task the participants had to accomplish, the testing process, and we will give an account of the findings from the questionnaires' evaluation.

6.1 Pilot studies

We conducted pilot studies to test and improve the design of the In-Situ-Questionnaires and evaluation environment. We tested the In-Situ-Questionnaires together with the evaluation environment continuously during development. However, to asses design, usability and user-friendliness, we needed people who had no knowledge of the purpose of the questionnaires. Therefore, we conducted two pilot studies, where we wanted to find out how people react to our questionnaires when they had never seen them before.

6.1.1 First Pilot study

At this stage we had decided for two possible objects to use, a chair and a bottle. We wanted to test for which object the tracking worked better and to which object interaction would the users respond better. We were able to use a work chair (see Figure 6.1b) with an arm-rest whose 3D model can be found online [Mil] (see Figure 6.1a). The bottle was chosen and modeled in Blender [Ble] (see Figures 6.2).

During the first pilot user study we experienced difficulties during calibration of the chair, as there was no ideal spot on the chair which was seen by the base stations, would not

(a) Sayl work chair model.

(b) Sayl work chair used as the identical chair.

(c) Chair used as the similar chair.

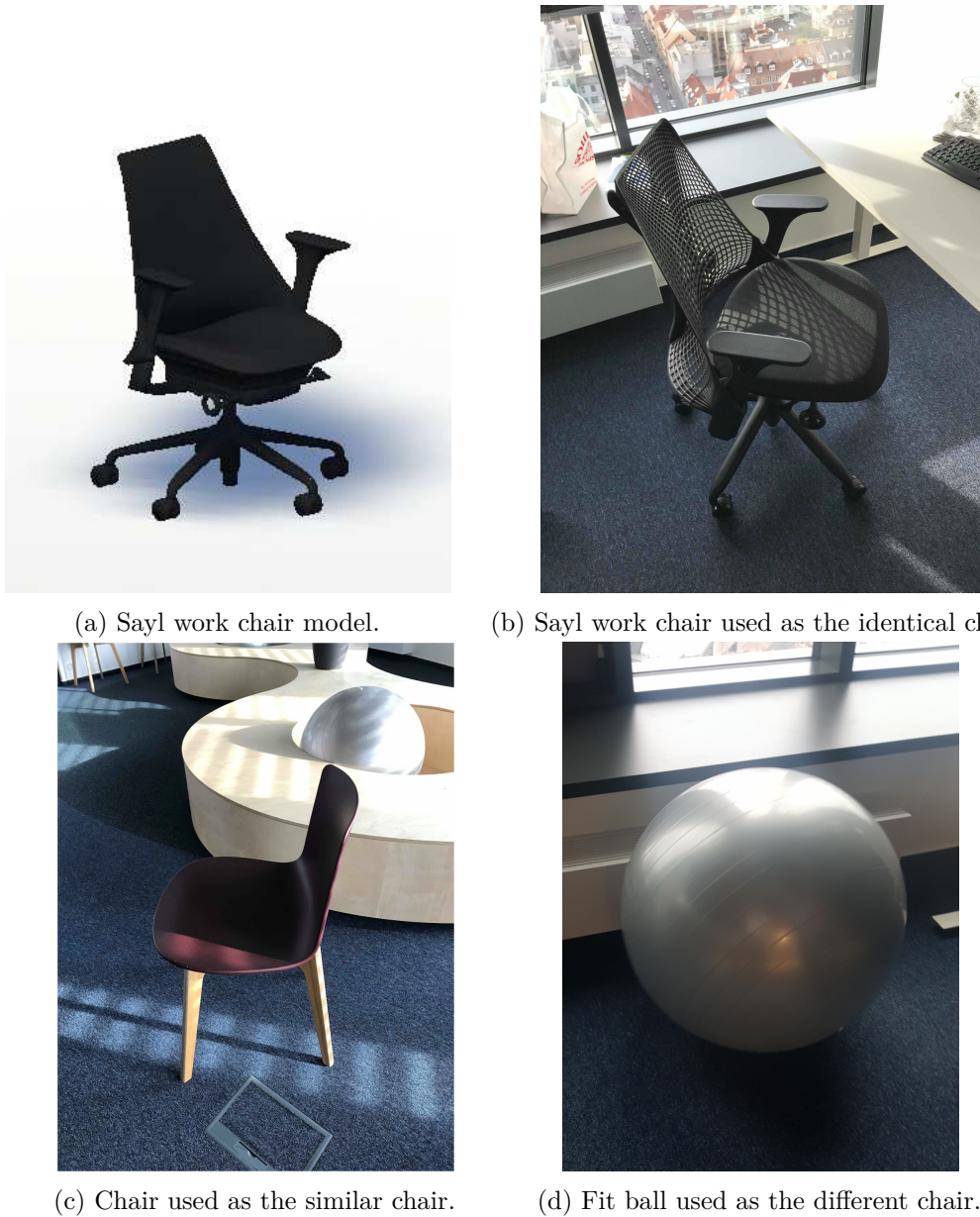(d) Fit ball used as the different chair.

Figure 6.1: Chairs used for evaluation.

be in the users' way and would have an angle, which we could set up in Unity. This caused the chair to be a bit tilted even after adjustments. Despite the discrepancies, the users gave the properties of the identical chair a very good rating. The second chair (the similar chair) lacked elbow-rests, was made out of a different material and the backrest was different (see Figure 6.1c). Even though it was still a chair and had similar properties, it was rated that it had almost no similarities to the model of the chair the users saw in

VR. Just two out of six properties did not get the worst possible point value. As the different chair we used a fit ball (see Figure 6.1d). It was rated with the worst possible point values for every property.

We thought that the second chair was a similar chair and users would give it a better rating. The fit ball rating was as expected, but proved not suitable for the user studies, as it is round and due to rolling motion attaching of the tracker is not possible. The participants were asked to fill out the Post-Questionnaire after the study.

One finding from this pilot study was the unwillingness of the participants to sit down on the chair although they had already touched it. We did not ask them to sit down explicitly, but if they wanted to fill out the questionnaire it seemed self-evident to do it while sitting. This inspired us to give users the task to sit down on the chair. Another observation from this pilot study was that the participants did not understand what was meant by the property Temperature. During the pilot study they asked questions about their task while they were in VR. We answered them in order to learn more about their understanding of the In-Situ-Questionnaires.

(a) The bottle used for evaluation.   (b) 3D model of the bottle used in VR.

Figure 6.2: Bottle and its 3D model.

After we tried the chair, we asked some users to test the second scenario where we used the bottle. They liked that they could move the bottle more freely than the chair. It was more precise. However, the chair worked better because it was tracked more robustly. The bottle often disappeared or showed up where it was not supposed to be. The LEAP Motion tracking was supposedly interfering with the base station tracking when directly looking at the tracker.

Despite choosing the chair, we knew that there was still a position where the users could occlude the tracker. The occlusion happens when the users are between the chair and the tracker behind the chair.

6.1.2 Second Pilot study

(a) Balcony chair used as an identical chair. (b) Wooden chair used as a similar chair. (c) Model of a balcony chair [Sto].

Figure 6.3: Used chairs and 3D model used in the VR.

The second pilot study was conducted to test all three In-Situ-Questionnaires and to find a suitable place to attach the tracker. We used a balcony (see Figure 6.3a) and a wooden chair (Figure 6.3b). We did not use a fit-ball nor the bottle as we have already decided to use chairs. Four users tried out all In-Situ-Questionnaires. As most of the time the users were in front of the chair (or sitting on it), the most suitable place to stick the tracker on was a vertical spot at the back of the chair, which was perpendicular to the floor. After exactly measuring where the HTC Vive Tracker was placed on the chair, the position of the chairs and its angle corresponded to the model the users saw in VR. When the chair was in the correct position, the users had no problems sitting down and had a "wow"-moment, as they have not yet experienced anything like it in VR. Nonetheless, to position the similar (wooden) chair was harder. The seat was almost at the same height, but its width, length and depth were different. The same applied to the backrest. The 3D model of the balcony chair was adjusted so that the front of the chair was in the same position. As the users were searching for the back-rest with their hands during the study, we decided to place the wooden chair's backrest to be in the same position as on the 3D model instead of on the front of the chair.

As the chair was offset from the model, users were a bit afraid and used their sense of touch more in order to determine where the chair is and to sit down on it. Another thing they disliked was the Scale In-Situ-Questionnaire where they had to click on the + or – button in order to evaluate the chair. One participant did not realize they can assign a negative value to the chair. He thought it was only possible to have positive numbers. This was solved by talking to the participant. Another one said it was cumbersome to

click on the plus or minus more than once. However, this may be good as they will give +3 (the best score) to the properties they really want to rate the best. There was a possibility that everything would converge to zero in this case, however in the previous pilot studies we saw they gave a +3 or -3 score when they filled out the questionnaire, so we did not consider their disliking of multiple clicking as a problem.

Unlike in the first Pilot study, the participants had no problems with the property Temperature and understood what it meant. They suggested introducing another property – Cleanliness. They described this property as when something is oily or there is dust on it but it should not be there. The reason behind it was a chair grabbed from a balcony where it stayed for some time gathering dust. To introduce a new property could be a possible future improvement. However, in the user studies we intended to use just indoor chairs and as a result we did not consider it necessary to add the property Cleanliness to the In-Situ-Questionnaires for the user studies.

Even though the users in the second pilot study did not have problems with the Temperature property, they had problems with the property Weight. Despite the fact that the real chair corresponded to the 3D model, the users thought the chair should have a different weight. Besides Weight, they had problems with the name of the Soft/Hard property. The users thought it suggested that on one side of the scale means soft and on the other side means hard. As we meant it as one property, we changed the name to Hardness, which they claimed was more understandable.

A second thing we changed based on the pilot studies was the scale design of the Scale In-Situ-Questionnaire. The scale had numbers from 5 to 1. However, the users were not sure whether 5 was the best or the worst. After the pilot studies we removed the numbers, leaving just smileys as an indicator of the best and worst rating (see Figures 4.9 and 4.8).

We asked the users to talk out loud during the pilot study and asked questions what they should do and could do, even though they received instructions at the beginning. The instructions were based on questions from participants in the first pilot studies. For instance, their task was to sit down on the chair and fill out the In-Situ-Questionnaire. As the LEAP Motion showed difficulties initially recognizing hands and showed them in unusual postures, we changed the instructions after the study by adding a part where we advised the users to lift their hands at the beginning and make sure they can see them. Another reason was that the participants had to realize they were able to see their hands and touch the chair with them. We thought we should leave people to play with their hands first, just to get used to seeing them in VR.

After the study the users evaluated their experience as good. However, they did not like the look of the VR test environment. They wanted to be placed in an environment that is cozier. Because of the representation of the sun in the VR scene, they thought the chair temperature should have been higher. One of the participants was colorblind, and reported problems seeing the In-Situ-Questionnaires because of the environment surrounding him. At that time there was just one wall and floor all white using the

default Unity environment. Even though the users did not like the environment and its temperature, we did not change these aspects as in the real world there will always be external factors, which we cannot control. If we made the environment cozier it could distract the users in other ways, so we decided to keep it the way it was.

6.2   User study

The goal of the user study was to evaluate whether the point score from the designed In-Situ-Questionnaires can distinguish the chairs together with the properties they differed in based on the evaluation from the users, who experienced interacting with the real chair and its counterparts in VR. We also wanted to find out how differently the users evaluate chairs based on which of the three In-Situ-Questionnaires was used. Depending on their answers, their reactions to interaction with chairs and In-Situ-Questionnaires, their remarks and Post-Questionnaire answers we wanted to discover individual advantages and disadvantages of the In-Situ-Questionnaires. If one of outperforms the others, mark it as the most suitable one.

A total of 19 people participated in the user study. The everage time spent was 36 minutes, ranging from 25 to full 40 minutes. The study was conducted during three consecutive days. We set the time slots to be 40 minutes with a 10 minute break in–between and after every four participants we had a 30 minutes break. The OBS Project [Stu] was used to capture the screen of the Unity Editor to record what the users saw. We used a camera secured on a tripod to record what the users did and said during the task session.

6.2.1   Task

At first, we thought of a complex task with more steps, which the users would do in the virtual environment. We would have used a robotic arm which would hold different parts of a coffee machine and bring the objects to the users. The task would have been to make a coffee. However, we realized touching different parts of the machine would not allow us to ask the users about their individual experience with every object they have touched. Therefore, we simplified the object, how the users interact with it and their task. We used an ordinary chair, a chair somehow similar to the model and a totally different chair. The task became simpler: one object for evaluation in one scene. When we made the decision to use a chair as the object, the task came naturally as was described in the previous section. The users were supposed to sit down on the chair and fill out the In-Situ-Questionnaire. While staying in VR, the users could do anything with the chair. They were even encouraged to, as it would help the users fill out the questionnaire. For instance to evaluate the weight of a chair, they would have to lift the chair at least a bit to feel it.

6.2.2 Design

We designed the user studies in order to gather chair ratings to allow us to make a statement on whether the In-Situ-Questionnaires are well-made. We tested them on three chairs. Each chair had its own HTC Vive Tracker attached in order to make the process of changing the chair efficient and to display them in VR. Each In-Situ-Questionnaire had its own VR scene, where we could hide the chairs, which we did not want to display and show the one the volunteer was intended to interact with.

Before the user studies, we had to check whether the VR scene was up and running together with the LEAP Motion and if the trackers were placed on the chairs. The chairs were covered with a blanket, so when the users came into the VR lab, they could not see them. When the users entered the VR lab, we seated them behind a table and gave them a one-page consent paper with the description of what they can expect from the experience in VR, together with instructions and their task. By signing it, they gave consent to be recorded. Before we put a headset on their heads, they were asked the VIMS question, where they evaluated their motion sickness symptoms, which will be described in the next sub-section. Afterwards, we put the headset on their head, started video recording with a camera and started capturing the screen with the OBS Project. The users did not yet see the scene as we had to place the chair first and start the scene.

Even though the task (sit down on the chair) and the aim of their stay in VR (fill out the In-Situ-Questionnaire) was written on the paper, the participants had to be reminded of what to do and asked whether they understood what they were supposed to do in VR. We added this step as during the pilot studies we saw that the volunteers did not thoroughly read the paper, were not sure what their task was, how they should interact with the In-Situ-Questionnaire and what they should evaluate.

Each participant should evaluate the same three chairs. However, the order is changed for each participant. We used counterbalancing to ensured that every combination was tested. This scheme was applied to all In-Situ-Questionnaires.

After all chair orders for a particular In-Situ-Questionnaire were tested, we showed another In-Situ-Questionnaire. The scene stayed the same for all chairs and In-Situ-Questionnares. The only thing that changed was the shown questionnaire and the representation of the chair.

After placing the chair next to the user, we started the scene made in Unity. As we used the LEAP Motion for hand tracking and it is not a common experience to see your hands in VR, we wanted the participants to realize they can see them. So when the users were in the scene, we asked them whether they can see their hands and to make themselves familiar with them. After that we asked them to come to the chair, interact with it, then sit down on it and fill out the In-Situ-Questionnaire. If they did not know how to evaluate a property, they could reexamine the chair by touching or standing up and lifting the chair. When the users were finished, we stopped the scene, asked them the VIMS question and changed the chair. The same was done for all three chairs, except that after the third chair, we only stopped the scene and took the headset

off the users' head. We seated them at the desk again where they filled out the rest of the Post-Questionnaire. During the time they were filling it out, we stopped recordings. After they filled it out, the user session was over.

6.2.2.1 Post-Questionnaire

The Post-Questionnaire was created for comparison with the In-Situ-Questionnaires. It consists of multiple parts. One part of the Post-Questionnaire is a method measuring Visually Induced Motion Sickness (VIMS) with a single question. Even though we consider it as a part of the Post-Questionnaire, it was asked before the users put the headset on and after every scene (four times in total). The rest of the Post-Questionnaire was asked after the users have accomplished their tasks and took of the headset. They were given the Post-Questionnaire to evaluate their VR experience. The Post-Questionnaire is described in greater detail in the rest of this section.

Visually Induced Motion Sickness

This method (see Figure 2.6) was introduced by Hwang et al. [HDGP18] which is based on the Wong-Baker FACES [WHEW+96] and it measures Visually Induced Motion Sickness (VIMS) with a single question. It is asked before the users are in VR and after each of the three tasks. We used this method, as when the users experience motion sickness, it is a good predictor for immersion. If the users are not feeling well in VR, they are usually distracted by this uneasy feeling and the fulfillment of the task is made harder. Furthermore, it is also an indicator for other problems that might come up, e.g. with the technical setup.

General data

The next part of the Post-Questionnaire is about general information of the users. We wanted to gain some background information about our users. They were asked about their prior experience. It could have influenced the experience they had during the experiment. We asked standard demographic questions such as their age and gender, Likert scale questions about how often they use computers, how much experience they have with VR and LEAP Motion, whether they have previously used VR and LEAP Motion and we asked them open-ended question about their experience.

System Usability Scale (SUS)

SUS is a *quick and dirty* method to asses system usability and has ten questions with a 5-point Likert scale [B+96]. Questions asked cover effectiveness (the ability of the users to do the tasks and the output quality), efficiency (how many resources were needed to perform the task) and satisfaction (users subjective reactions). Output from this part of the

Post-Questionnaire gave us information about the usability of our In-Situ-Questionnaires and the environment for testing. It could give us hints on what to improve in the future.

Presence Questionnaire

In the fourth part of the Post-Questionnaire we used 21 out of 24 questions from the Presence Questionnaire by Witmer and Singer [WS98]. We did not use the part concerning sounds as we did not use any sounds in VR. However, we kept the part about the use of touch as it is an important part of the participants' experience during our user study.

End notes

At the end of the Post-Questionnaire we asked the users to describe their most positive incident, most negative incident and to write about their experience to evaluate what the users liked and did not like. This could help us improve the questionnaires and the virtual environment for the future use.

6.3 Results

In this section we will present gathered data from the In-Situ-Questionnaires and the Post-Questionnaire from the user study. We will start with the Post-Questionnaire.

6.3.1 Post-Questionnaire

Results from the Post-Questionnaire will be presented in order and in sections based on how it was designed.

6.3.1.1 General data

A total of 19 people, five being female and 12 male, between the age of 26 and 49 volunteered for to the user study (mean = 31.2, SD = 5.2). All participants stated that they use computers very often, three people had no prior experience with VR, others had at least some experience with an average of 2.3 on a scale from 0 - *Not at all* to 4 - *Very much*. The LEAP Motion was new to all except three participants, which experienced it in another study (mean = 0.4, SD = 0.9 with the same scale as for the VR experience).

6.3.1.2 Visually Induced Motion Sickness

The users were asked whether they were experiencing any motion sickness four times in total - in the beginning and after each task. They answered on a scale from 0 to 4, 0 meaning *no sign of motion sickness*, 4 meaning *severe motion sickness symptoms*. Only four people evaluated their motion sickness symptoms other than 0 - no sign. Two users experienced slight symptoms (evaluated as 1) after accomplishing all tasks. This was due

to a jittery In-Situ-Questionnaire, which originated from unstable sticking of the tracker to the chair, and unstable chair tracking. Even though they were not the only ones with the jittery In-Situ-Questionnaire or the flying chair, other participants still evaluated their motion sickness with 0.

One participant started with mild symptoms of motion sickness. However, after the first task, he changed it to *no sign of motion sickness.* Another participant evaluated his symptoms as *slight* from the beginning to the end.

6.3.1.3 SUS

The SUS questionnaire and how it's evaluated was described in the section 2.4. The experience with the environment and the interaction within was evaluated with the score of 78.33. This is according to Sauro [Sau11], an acceptable score as it is above 68.

As we mentioned in the section 2.4, there is an alteration of positive and negative statements. Sometimes our system is evaluated better if the value is 1, sometimes when it is 5. To compare the results, we inverted positive statement evaluations in this chapter to match the negative statements, so if the positive statement had 5 points, after inverting it received 1 point, which was *strongly agree* and 5 - *strongly disagree.*

The statement *4. I think that I would need the support of a technical person to be able to use this system.* received the worst rating (mean = 2.32) with the highest standard deviation (SD = 1.38). Another statement *5. I found the various functions in this system were well integrated.* received the second worst mean score (mean = 2.21, SD = 0.85).

The following two statements received the highest score: *10. I needed to learn a lot of things before I could get going with this system* with mean = 1.42, SD = 0.61 and *3. I thought the system was easy to use* with mean = 1.47, SD = 0.61.

We think the two worst evaluated statements reflect an unstable environment, where the chairs flew away or the users' hands disappeared. The two best evaluated statements describe our system as easy to use and straightforward.

6.3.1.4 PQ

The PQ had a scale from 1 to 7. Mostly the 7 was the best score. However, e.g. the question:

14. How much delay did you experience between your actions and expected outcomes? No delays — Long delays [WS98]

was asked in such a way that 1 point was best. Unlike e.g. the SUS, we are not interested in the overall score and compare it since it was not an objective to develop a highly immersive environment. Nonetheless, we can look at the evaluation of the individual questions. We can start with the questions about haptics and the questions with the lowest or the highest score.

64

The users were asked two questions about haptics:

20. How well could you actively survey or search the virtual environment using touch?

21. How well could you move or manipulate objects in the virtual environment? [WS98]

The first haptic question received a mean value of 5.16 (SD = 1.30). The second question received a mean of 5.45 (SD = 1.01). This score indicates a positive experience with touching and manipulating the chairs.

Based on the best mean score, the users adjusted quickly to the virtual environment (mean = 6.26, SD = 1.19), experienced low delay between the action and the expected outcome (mean = 1.92, SD = 0.85) and they stated they were able to examine the objects closely (mean = 5.95, SD = 1.22). This indicates an accurate implementation of our VR environment.

We also identified questions with the worst score. The users did not completely agree that the virtual environment seemed consistent with the real world experiences (mean = 4.32, SD = 1.29) and that they were involved through visual aspects (mean = 4.58, SD = 1.84), which had the highest standard deviation. The question about how compelling the sense of moving around in VR was, had the second highest SD (mean = 5.37, SD = 1.71).

6.3.1.5 End notes

At the end of the Post-Questionnaire the participants answered three open questions. We will start by writing about what they found most positive during their experience. Nine people claimed that controlling virtual object with hands was fun, immersive, or it felt natural. Seven volunteers liked playing and seeing how their virtual hands reacted to movement, three praised hand and chair tracking, two users liked sitting on the chair, another two participants claimed that the 3D model of the identical chair reflected the reality well and expectations were subverted for one participant.

On the other hand, eight users disliked incorrect chair position tracking, four disliked when the real chair shape did not match the virtual model, another four found it hard to sit down or interact with the different chair. Two participants disliked disturbing cable, jittery In-Situ-Questionnaire and the fact that the VR fingers did not match what their fingers did.

Three users experienced a jittery In-Situ-Questionnaire which we think was due to unstable attachment of the tracker on the chair. Despite the jittery questionnaire, the users were able to fill out the In-Situ-Questionnaires. This happened at the end of the second and third day of testing. However, when we reattached the tracker, the jitter stopped.

To the last question, four people answered that ability to feel and see details of virtual objects or hands aroused their interest. One volunteer was surprised how fake the identical chair felt despite anticipating that it would match the expectations the most, another

one was a bit scared to move around with no feeling of how big the scene is and wanted a reference object. One participant found interesting the clash of expectations and real sensory input when the chairs were exchanged.

6.3.2   In-Situ-Questionnaires

When the users filled out the In-Situ-Questionnaire, their rating was saved to a spreadsheet file, which we then processed. The results will be presented in this sub-section.

6.3.2.1   Score In-Situ-Questionnaire

We gathered data from six people about three chairs with the Score In-Situ-Questionnaire. The users rated the properties on a scale from -3 to + 3. We transposed it to a scale from 0 to 6 to allow easier comparison to the other questionnaires. The chair corresponding to the virtual 3D model received the best mean score (mean = 5.03) and also the lowest standard deviation (SD = 1.07). The similar chair received a mean score of 2.44 (SD = 1.76), the different chair received 1.22 (SD = 1.24).

Except for the Temperature property, the similar chair had a higher mean score than the different chair, which can be seen in the Figures 6.4. The Shape property of the different chair was evaluated as what matched the least (mean = 0.17). It also had the lowest standard deviation (SD = 0.41). The Weight of the identical chair received the worst score of all properties (mean = 4.33, SD = 1.21). The similar chair (mean = 3.67, SD = 2.16) did not differ with the identical chair in the Weight property as much as in the other properties. Two out of six people assigned more points to the similar than to the identical chair, two assigned the same number of points and two people assigned 1 point, which increased the standard deviation and lowered the mean score for the similar chair.

However, the standard deviation score of the similar chair was the highest, except for the Temperature property.

6.3.2.2   Right/Wrong In-Situ-Questionnaire

Another six people evaluated chairs with the Right/Wrong In-Situ-Questionnaire, which is different from the others in its scale. The scale is divided into two rows, *felt right* and *felt wrong*. The same property could not be selected in both rows, but it also does not have to be selected at all, which we can interpret as something in between. To create a chart, we transposed the data collected from the users. When they selected *felt wrong*, we evaluated it as 0 points, when they selected nothing, we assigned it with 1 point and when they selected *felt right*, it received 2 points. The average mean score was 1.69 for identical chair (SD = 0.44), 0.53 for similar chair (SD = 0.58) and 0.67 for the different chair (SD = 0.73), which shows us that the different chair received higher point scores than the similar chair.

All participants chose *felt right* for the properties Shape and Material for the identical chair (see Figure 6.5). On the other hand the property Shape felt wrong for all participants

for the similar and different chair. All participants selected *felt wrong* for the Material of the similar chair. The only property in which the identical chair did not have the highest mean score was Weight. The participants rated the weight of the similar chair better.
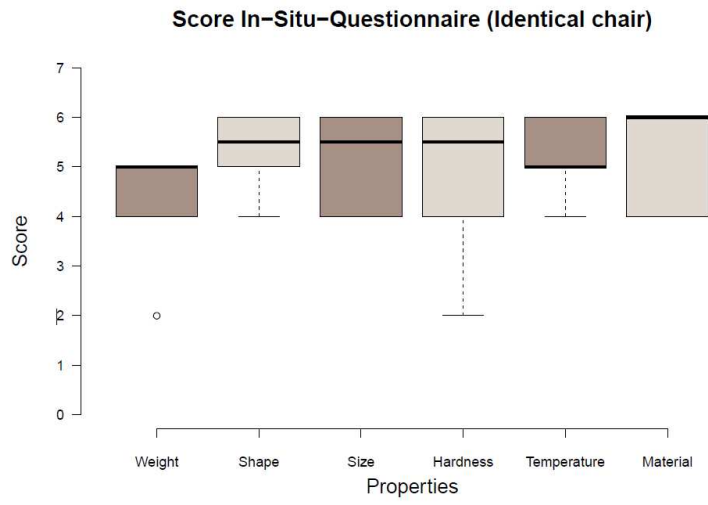
Standard deviation of the Shape property for all chairs and the Material for the identical and similar chair is 0. However, the property Material of the different chair received the highest standard deviation (SD = 1.1) of all properties.

6.3.2.3 Scale In-Situ-Questionnaire

The Scale In-Situ-Questionnaire was translated to data by assigning values to the circles. The best evaluation, which was the green circle, received 4 points. The worst evaluation, which was the red circle, received 0 points.

The different chair had the highest average standard deviation across all properties (see Figure 6.6). However, this is due to one person, who evaluated for instance the different chair with the highest score of 4 for the property Shape, when the other participants evaluated the properties with 0 or 1.

The Size property of the identical chair received the highest possible scores from all participants (mean = 4, SD = 0). The average mean score was 3.48 (SD = 0.66) for the identical chair, 1.67 (SD = 0.9) for similar chair and 1.52 (SD = 1.42) for the different chair.

Score In−Situ−Questionnaire (Identical chair)

(a) Identical chair score.

Score In−Situ−Questionnaire (Similar chair)

(b) Similar chair score.

Score In−Situ−Questionnaire (Different chair)
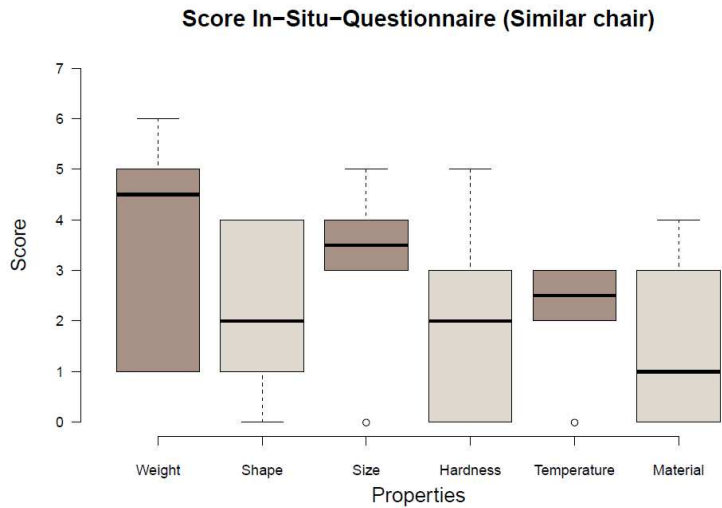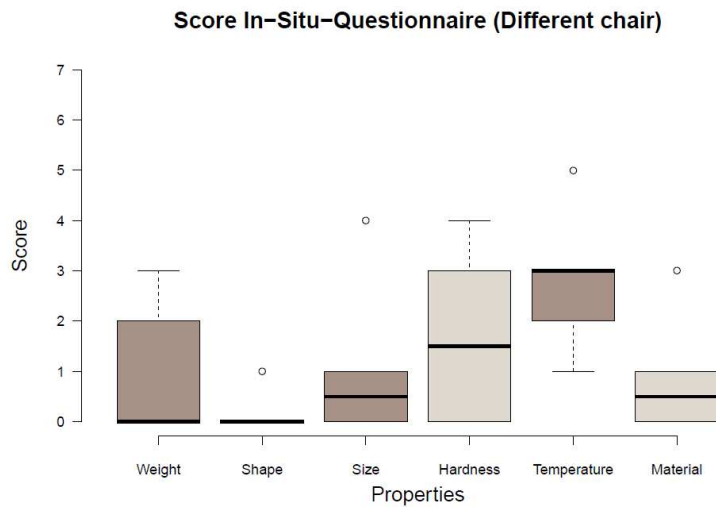
(c) Different chair score.

Figure 6.4: Score In-Situ-Questionnaire. Evaluation of the chairs.
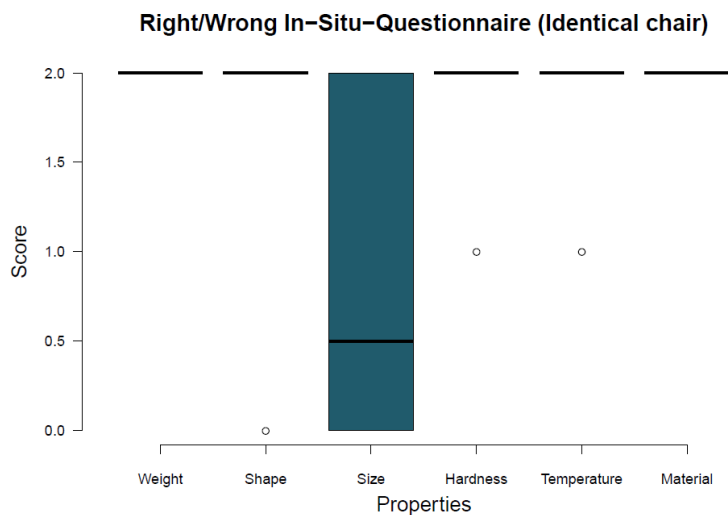
(a) Identical chair score.
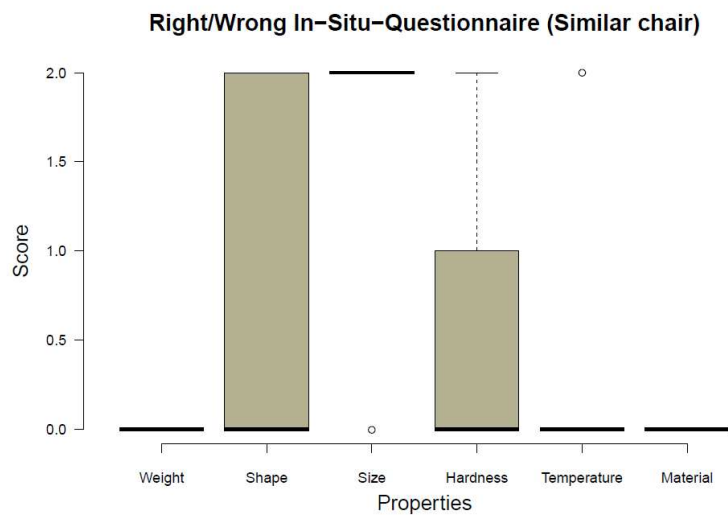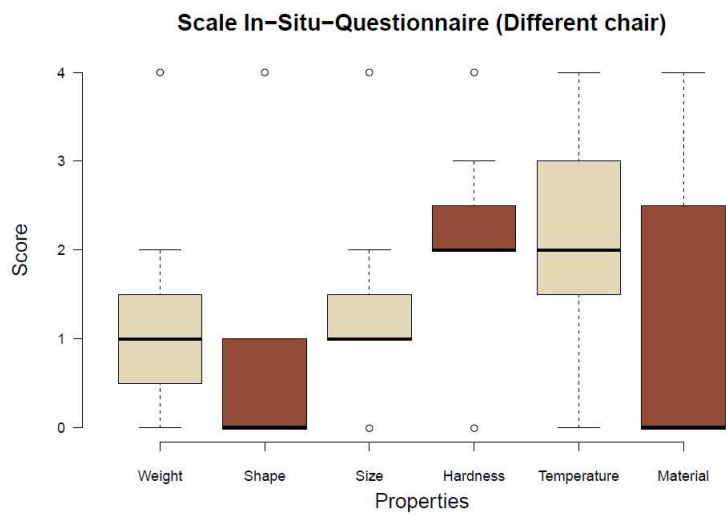

(b) Similar chair score.
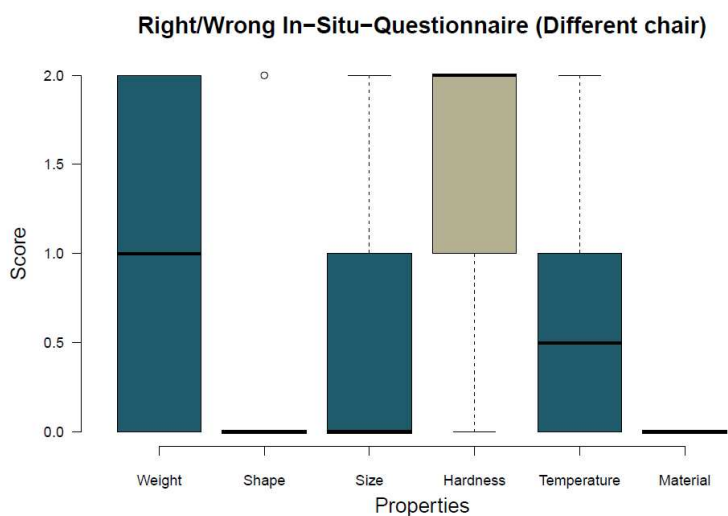
(c) Different chair score.

Figure 6.5: Right/Wrong In-Situ-Questionnaire. Evaluation of the chairs.
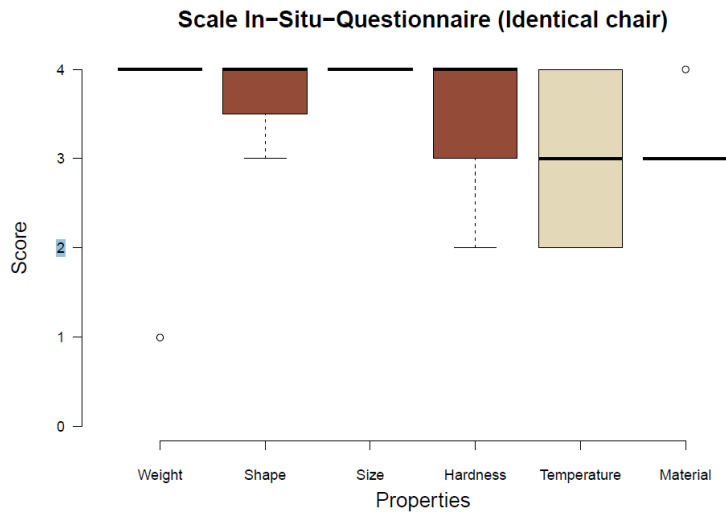
(a) Identical chair score.
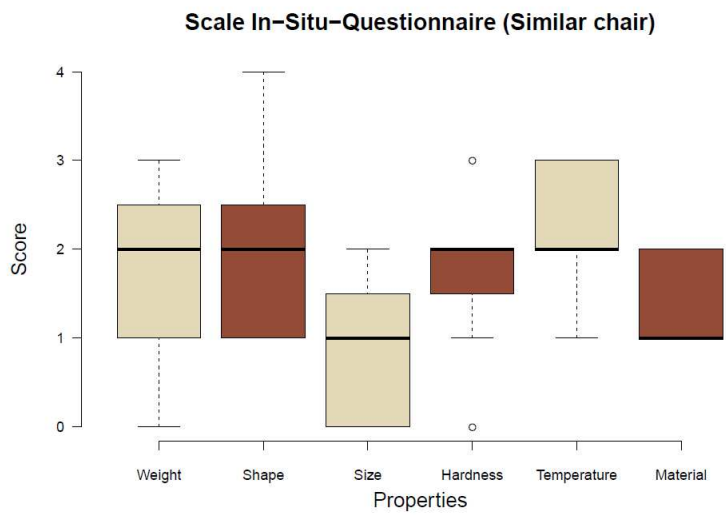
(b) Similar chair score.

(c) Different chair score.

Figure 6.6: Scale In-Situ-Questionnaire. Evaluation of the chairs.

CHAPTER 7

Discussion

In the previous chapter we presented the analysis of the study results, in this chapter we will discuss the results of the VIMS, the SUS and the PQ and compare the results of the three In-Situ-Questionnaires. Afterwards, we will briefly discuss possible improvements of the In-Situ-Questionnaires.

7.1 Post-Questionnaire

7.1.1 Visually Induced Motion Sickness

Only two people noted worsened motion sickness, which they rated as slight symptoms (1 on the scale). This was due to the jittery In-Situ-Questionnaire and unstable tracking of the chair. The VIMS ratings suggest the interaction with the chair and the In-Situ-Questionnaire was properly made and despite the fact that two out of three chairs were not identical representations of the chair they saw and thus the haptic feedback was different, it did not make them feel sick.

7.1.2 System Usability Scale

The acceptable score from the SUS indicates, that the In-Situ-Questionnaires together with the environment for testing are adequate and can be easily used. The users found it neither complex nor hard to use. They needed a support to make themselves familiar with how the environment worked, but after a few minutes they were able to interact with it on their own.

The SUS was used by researchers also to assess haptic experience with a system as mentioned in section 2.4. But it can only tell us about the system as a whole, not give any insights into the haptic experience. Therefore, the results from the SUS mean that the environment for evaluation was well integrated. The participants should not have

been disturbed by it. They had suitable conditions for immersion, interaction with the chairs and for filling out the In-Situ-Questionnaire without many disturbances.

7.1.3 PQ

The users filled out the PQ once as a part of the Post-Questionnaire. We thought of asking the PQ after each task to be able to distinguish their experience with different chairs and compare it with results from In-Situ-Questionnaires. However, to answer 32 questions after each task would take too much time, distract the participants from the VR experience and it still would not give us any fine detailed feedback about their haptic experience.

Despite asking the PQ only once, we can still examined results of each question as there is no score key for overall impression.

Haptic questions already mentioned in the 6.3.1.4 received 5.16 and 5.45 out of 7, which we assume means the users were able to search the environment using touch and manipulate the chairs. We think the users did not evaluate the questions with a higher score is that they were rated all three chairs at once. And when they remembered touching the similar or the different chair, what they were looking at did not correspond to what they were feeling and sometimes they had problems finding the chair. The drawback of the PQ is that we do not get rating for a single experience, but a sum of experiences.

The highest standard deviation of 1.84 (mean = 4.58) had the question:

How much did the visual aspects of the environment involve you? [WS98]

A comment of one user, who stated he would have liked more objects in the environment to use as anchor points could give a hint to the reason for this score. We have made the environment plain in order to not disturb the user. However, based on the standard deviation score, we think not all participants liked it.

Based on the highest and the lowest mean score, the users adjusted quickly to the virtual environment (mean = 6.26, SD = 1.19), experienced low delay between action and expected outcome (mean = 1.92, SD = 0.85) and they were able to examine the objects closely (mean = 5.95, SD = 1.22).

Based on the score of the questions, we believe the users were quick to familiarize themselves with the environment, did not have problems interacting with it and the environment behaved as they expected, which indicates a good design.

However, as we used the PQ only at the end of the user session, it is not possible to determine how the differences of the chairs would reflect on the answers. Nonetheless, we believe that using PQ after each task would not have given us information about the properties of the objects as most of the questions were about the environment and even the haptic ones were about searching and manipulating the environment and not about the object itself.

7.2 In-Situ-Questionnaire

In this section we will start by discussing the individual In-Situ-Questionnaires, and later we will provide a summary.

7.2.1 Score In-Situ-Questionnaire

Based on the evaluation of the three chairs, we think the Score In-Situ-Questionnaire can distinguish whether the chair is alike or different. The similar chair received a higher mean score for all properties (mean = 2.44) than the different chair (mean = 1.22). Even though, the similar chair was resembling the 3D model more closely, there were slight differences. We can see it, when we compare it to the average score, which the identical chair received (mean = 5.03). However, when we look at the properties one by one, the identical chair was made out of textile and plastic, and the similar chair was made out of wood and metal, hence the score for the property Material and Hardness should be lower for the similar chair. And this can be seen in the Figure 6.4. We believe this shows us that the Score In-Situ-Questionnaire can distinguish, whether the object is different or alike based on the received score. If the object differs, the score tells us in which properties.

The scores for the different chair would have been even more distinct. However, they could have been influenced by one participant, who evaluated the Size, Hardness with 4 and Weight and Material with 3, which were the highest scores the different chair received. This could be a possible outlier as maybe the participant misinterpreted the question.

We think that the perceived temperature of an object in our case depends on the color and the material. The virtual lightsource of the scene was not shining directly on the chairs, but as the same and different chairs were black, it could have *heated it up* a bit. As the same and the different chair had a material made out of fabric or plastic, the similar chair, which was made out of wood, should have felt different to touch. However, for the score for Material of the similar and the different chair, the users gave more points to the similar chair than to the different chair. We believe the reason for that is the overall difference between the 3D model perceived by eyes, and the real chair, which the users touch.

7.2.2 Right/Wrong In-Situ-Questionnaire

As the users had only three possibilities, to select in the Right/Wrong In-Situ-Questionnaire (felt right or wrong and when the users did not select anything) the users had to choose from fewer options. This structure seems to have forced them to pick either *felt right* or *felt wrong*, as the users did not chose to select no property only in 8 out of 108 property evaluations.

Results of the Right/Wrong In-Situ-Questionnaire seem to be different in comparison to the Score and the Scale In-Situ-Questionnaire, as for instance the overall score for

the similar chair is higher than the overall score for the different chair. This is not true for the Right/Wrong In-Situ-Questionnaire when we compare just the similar and the different chair, where the overall score is better for the different chair. The properties Hardness, Temperature and Material of the different chair received more points. Weight and Size turned out better for the similar chair and Shape of both of them felt wrong to all participants. This could be due to the material of the chairs. The similar chair was mostly wooden and the different chair was covered with synthetic leather. When we compared it to the 3D model, which was shown to the user, it looked like it was made out of textile and plastics, which made the different chair look more similar. We believe the material of the chairs also explains the evaluation of the Hardness and Temperature properties as the wood is harder and colder to touch than textile or synthetic leather, which are softer and warmer.

The properties Weight and Size corresponded to the reality, where the similar chair weights more and the size is similar to the identical chair. On other hand, the different chair is a lot smaller and weighs less.

The Shape of the chair felt right only for the identical chair for all participants, which corresponds to how the chairs look like and the fact, that the participants considered the missing arm-rests as a noteworthy difference.

Examining the evaluated properties one by one, the only property, which did not receive the highest score for the identical chair was Weight, where the similar chair received the highest score. The reason for the evaluation could be the anticipated weight of the identical chair, which according to participants comments was lighter and they considered the weight of the similar chair to be more appropriate. However, this result can be seen only in this In-Situ-Questionnaire.

7.2.3   Scale In-Situ-Questionnaire

The Size and Hardness have higher point score for the different chair than the similar chair. The score of the property Size is higher due to one participant, who assigned 4 points to the different chair when other participants assigned 2 or fewer points. Based on the texture and material of the chair, we think the property Hardness of the different chair should have a higher mean score than that of the similar chair.

We think the Score and the Scale In-Situ-Questionnaire brought similar results, where identical chair received the highest score, similar chair received fewer points, but still more than the different chair. Probably more participants would have yielded clearer results.

7.2.4   Comparing all In-Situ-Questionnaires

We believe that all designed In-Situ-Questionnaires met the expectation. The identical chair had the highest score and the In-Situ-Questionnaires were able to distinguish different properties of the objects, which based on our research, no questionnaire used in VR

could do. The PQ and the SUS, which we used in the Post-Questionnaire, focused on the environment, its visual aspects, interactions happening inside, but not on haptic qualities of an object. Other questionnaires we presented in chapter 2 consisted of questions about physical presence (SUS PQ), how immersed the users (ITQ) are, emotional response of the users (SAM) or evaluation of users' subjective feelings (AttrakDiff). However, we wanted to measure users perception of a specific object, which we believe the In-Situ-Questionnaire can do.

When we compare the data for all three In-Situ-Questionnaires, they do not give us identical results (e.g. the worst evaluated property of the similar chair is Material for the Score In-Situ-Questionnaire, Shape for the Right/Wrong In-Situ-Questionnaire and Size for the Scale In-Situ-Questionnaire). From the evaluation results, we consider the Right/Wrong In-Situ-Questionnaire to represent the reality the most as it clearly states that Shape was right only for the identical chair and despite the fact that we thought of the second chair as similar, the third chair had (based on user feedback) more properties in common with the 3D model than the similar chair had. However, the other two In-Situ-Questionnaires could represent more of an overall similarity of an object, where size could be a more important property than material. For instance even when we chose which chair was similar, we did not look at the temperature or hardness of the chair, but its shape and size. Therefore, we think we can split the questionnaires into two groups and use them based on what we need to accomplish.

7.3 Possible Improvements

Even though the tracking was stable most of the time, some users experienced occasional problems with shaky In-Situ-Questionnaire or chairs flying off. We believe the shaky In-Situ-Questionnaire was due to an unsteady attachment of the tracker, which could be addressed with more durable adhesive options. When it comes to the tracking problem, where the chair was flying off, we could add two base stations at an angle, which could locate the tracker even when the original base stations could not see it as it was occluded by the user.

Another possible improvement concerns the task written in the introductory text *Compare what you see and feel* in the In-Situ-Questionnaires, which was not well understood, even when we explained to the users what to do at the beginning of the session. We had to repeat what they should do multiple times in a different way. Maybe a more accurate text like *Compare what you see with your eyes and feel with your hands* would help. However, this is not a common comparison, which people make regularly, therefore, we believe additional explanation would always be needed.

Some participants proposed a new property *Cleanliness*, which could be added to In-Situ-Questionnaires. Other properties could be also identified later, which are noteworthy for object comparison. Also, we could identify, which properties are essential and matter the most, when it comes to the similarity based on the users perception.

In the Right/Wrong In-Situ-Questionnaires not all properties have to be selected. Not all users realized this. We could tell this to the participants beforehand or include it in the questionnaire, to make it more apparent. Other changes could include the way of how the users change their answer. Some users would like to just click on the property in the other row, whereas now they have to unclick the selected property and click on the same property in the other row. However, even if it was made this way, the users would still need to click on the selected property to deselect it, when it felt neutral. Still, the design of the Right/Wrong In-Situ-Questionnaire allowed this particular result and by changing the design, the results could change as well. Further developing and testing different questionnaires with this setup is highly recommended.

As we had one colorblind participant during the pilot studies, who reported problems seeing the In-Situ-Questionnaire, we could look at how people with colorblindness see the In-Situ-Questionnaires in the environment and decide how we could improve the color scheme so that they can see the environment and the In-Situ-Questionnaire clearly.

CHAPTER 8

Conclusion and Future work

Our aim was to develop a VR questionnaire, which would not disturb the users while they were in VR interacting with real objects. We wanted to get an exact measurement of which examined object properties differ. None of the existing questionnaires we have found during our research would fit this purpose as they do not examine object properties specifically.

The In-Situ-Questionnaires were designed and developed to be stand-alone for future use. It was implemented and tested in an environment for evaluation with three scenes. The users were able to move in the scene, see their hands and interact either with the chair or the In-Situ-Questionnaire.

For testing purposes, we designed a simple task, to sit down on a chair while users were in VR and to fill out the In-Situ-Questionnaire. Even though, we could see score point differences between the three chairs during user studies, we had a sample of 19 participants and for further verification we would need a lot more participants. More evaluation would give us a clearer insight into the difference between the In-Situ-Questionnaires, which would enable us to decide which In-Situ-Questionnaire gave us the most precise point score.

As people can adapt to small visual and proprioceptive mismatches, we could use the In-Situ-Questionnaires to find out how much simpler the haptic representation could get e.g. whether a pumpkin in the VR could be represented by a ball in reality. With the In-Situ-Questionnaires it would be possible to not only find a replacement for a simple single object, but also to put together multiple objects to represent a larger, more complex one. The In-Situ-Questionnaires would simplify and make the process of finding either the 3D representation or real object cheaper.

Even though the In-Situ-Questionnaires allows score point comparison of multiple properties of tested objects, we cannot yet distinguish based on the point score, which objects can be used instead of an identical object shown to the user in the VR. Therefore, the

future work includes limits of when an object is sufficiently similar to the shown 3D model in VR and can be used as a substitute for an exact copy of the 3D model.

List of Figures

Bibliography

[B⁺96]   John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[BKM08]   Aaron Bangor, Philip T Kortum, and James T Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

[Ble]   Blender. Free and open source 3d creation suite. `https://www.blender.org/`. Accessed March 4, 2020.

[Bro]   John Brooke. System usability scale. `https://hell.meiert.org/core/pdf/sus.pdf?/`. Accessed March 4, 2020.

[BSC⁺18]   Miguel Borges, Andrew Symington, Brian Coltin, Trey Smith, and Rodrigo Ventura. Htc vive: analysis and accuracy improvement. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2610–2615. IEEE, 2018.

[CC16]   Sylvain Chagué and Caecilia Charbonnier. Real virtuality: a multi-user immersive platform connecting real and virtual worlds. In *Proceedings of the 2016 Virtual Reality International Conference*, pages 1–3, 2016.

[CGAK⁺19]   Polona Caserman, Augusto Garcia-Agundez, Robert Konrad, Stefan Göbel, and Ralf Steinmetz. Real-time body tracking in virtual reality using a vive tracker. *Virtual Reality*, 23(2):155–168, 2019.

[CGL10]   Dustin B Chertoff, Brian Goldiez, and Joseph J LaViola. Virtual experience test: A virtual environment evaluation questionnaire. In *2010 IEEE virtual reality conference (VR)*, pages 103–110. IEEE, 2010.

[COB⁺18]   Inrak Choi, Eyal Ofek, Hrvoje Benko, Mike Sinclair, and Christian Holz. Demonstration of claw: A multifunctional handheld vr haptic controller. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–4, 2018.

[CRR⁺15]   Lung-Pan Cheng, Thijs Roumen, Hannes Rantzsch, Sven Köhler, Patrick Schmidt, Robert Kovacs, Johannes Jasper, Jonas Kemper, and Patrick Baudisch. Turkdeck: Physical virtual reality based on people. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pages 417–426, 2015.

[DAB09]   Christina Dicke, Viljakaisa Aaltonen, and Mark Billinghurst. Simulator sickness in mobile spatial sound spaces. In *Auditory Display*, pages 287–305. Springer, 2009.

[DNJ18]   KS Dhanasree, KK Nisha, and R Jayakrishnan. Hospital emergency room training using virtual reality and leap motion sensor. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 924–928. IEEE, 2018.

[F⁺86]   Nico H Frijda et al. *The emotions.* Cambridge University Press, 1986.

[FH98]   John M Flach and John G Holden. The reality of experience: Gibson's way. *Presence*, 7(1):90–95, 1998.

[GLGM⁺17]   Yoren Gaffary, Benoît Le Gouis, Maud Marchal, Ferran Argelaguet, Bruno Arnaldi, and Anatole Lécuyer. Ar feels "softer" than vr: haptic perception of stiffness in augmented versus virtual reality. *IEEE transactions on visualization and computer graphics*, 23(11):2372–2377, 2017.

[HDGP18]   Alex D Hwang, Hongwei Deng, Zhongpai Gao, and Eli Peli. Quantifying visually induced motion sickness (vims) during stereoscopic 3d viewing using temporal vims rating. *Electronic Imaging*, 2018(14):1–9, 2018.

[HHY⁺17]   Ridho Rahman Hariadi, Darlis Herumurti, Anny Yuniarti, Imam Kuswardayan, Nanik Suciati, and Tikva Mooy. Virtual sasando using leap motion controller. In *2017 International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA)*, pages 161–164. IEEE, 2017.

[HSR18]   Dustin T Han, Mohamed Suhail, and Eric D Ragan. Evaluating remapped physical reach for hand interactions with passive haptics in virtual reality. *IEEE transactions on visualization and computer graphics*, 24(4):1467–1476, 2018.

[HZGP17]   Zhenyi He, Fengyuan Zhu, Aaron Gaudette, and Ken Perlin. Robotic haptic proxies for collaborative virtual reality. *arXiv preprint arXiv:1701.08879*, 2017.

[JND⁺00]   Cathryn Johns, David Nunez, Marc Daya, Duncan Sellars, Juan Casanueva, and Edwin Blake. The interaction between individuals' immersive tendencies and the sensation of presence in a virtual environment. In *Virtual Environments 2000*, pages 65–74. Springer, 2000.

[JPNI14]   Suwicha Jirayucharoensak, Setha Pan-Ngum, and Pasin Israsena. Eeg-based emotion recognition using deep learning network with principal component based covariate shift adaptation. *The Scientific World Journal*, 2014, 2014.

[KLBL93]   Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[Lab]   UQO Cyberpsychology Lab. Presence questionnaire. `https://marketinginvolvement.files.wordpress.com/2013/12/questionnaire-sur-l_c3a9tat-de-prc3a9sence-pq-f.pdf`. Accessed April 17, 2020.

[LGS+16]   Danielle Levac, Stephanie MN Glegg, Heidi Sveistrup, Heather Colquhoun, Patricia A Miller, Hillel Finestone, Vincent DePaul, Jocelyn E Harris, and Diana Velikonja. A knowledge translation intervention to enhance clinical application of a virtual reality system in stroke rehabilitation. *BMC health services research*, 16(1):557, 2016.

[LMa]   Inc. LEAP Motion˙ A basic set of leap hands. `https://leapmotion.github.io/UnityModules/core.html#a-basic-set-of-leap-hands`. Accessed March 4, 2020.

[LMb]   Inc. LEAP Motion˙ How does the leap motion controller work? `http://blog.leapmotion.com/hardware-to-software-how-does-the-leap-motion-controller-work/`. Accessed March 4, 2020.

[MAO17]   Emanuela Maggioni, Erika Agostinelli, and Marianna Obrist. Measuring the added value of haptic feedback. In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6. IEEE, 2017.

[Mil]   Herman Miller. Chair model. `https://www.hermanmiller.com/en_eur/resources/models/3d-models/?q=Sayl%20Work%20Chair`. Accessed March 4, 2020.

[Mur17]   Jeff W Murray. *Building virtual reality with Unity and Steam VR*. CRC Press, 2017.

[NMN+18]   Ryohei Nagao, Keigo Matsumoto, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. Ascending and descending in virtual reality: Simple and safe system using passive haptics. *IEEE transactions on visualization and computer graphics*, 24(4):1584–1593, 2018.

[RDI03]   G Riva, F Davide, and WA IJsselsteijn. 7 measuring presence: Subjective, behavioral and physiological methods. *Being there: Concepts, effects and measurement of user presence in synthetic environments*, pages 110–118, 2003.

[RDLT06]   Gabriel Robles-De-La-Torre. The importance of the sense of touch in virtual and real environments. *Ieee Multimedia*, 13(3):24–30, 2006.

[SAMGB⁺18]   Luis Martín Sánchez-Adame, Sonia Mendoza, Beatríz A González-Beltrán, José Rodríguez, and Amilcar Meneses Viveros. Aux and ux evaluation of user tools in social networks. In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 104–111. IEEE, 2018.

[Sau11]   Jeff Sauro. Measuring usability with the system usability scale (sus). In *Measuring Usability*, pages 1—5, 2011.

[Sla99]   Mel Slater. Measuring presence: A response to the witmer and singer presence questionnaire. *Presence*, 8(5):560–565, 1999.

[SSHR18]   Mohamed Suhail Mohamed Yousuf Sait, Shyam Prathish Sargunam, Dustin T Han, and Eric D Ragan. Physical hand interaction for controlling multiple virtual objects in virtual reality. In *Proceedings of the 3rd International Workshop on Interactive and Spatial Computing*, pages 64–74, 2018.

[Ste]   SteamVR. Steamvr unity plugin. `https://valvesoftware.github.io/steamvr_unity_plugin/articles/intro.html`. Accessed March 4, 2020.

[Sto]   Unity Asset Store. Balcony chair. `https://assetstore.unity.com/packages/3d/props/furniture/summer-open-air-table-and-chair-94677`. Accessed March 4, 2020.

[Stu]   OBS Studio. Obs project. `https://obsproject.com/`. Accessed March 4, 2020.

[SVG15]   Adalberto L Simeone, Eduardo Velloso, and Hans Gellersen. Substitutional reality: Using the physical environment to design virtual reality experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3307–3316, 2015.

[Teca]   Unity Technologies. Manual - canvas. `https://docs.unity3d.com/Packages/com.unity.ugui@1.0/manual/UICanvas.html`. Accessed March 4, 2020.

[Tecb]   Unity Technologies. Manual - colliders. `https://docs.unity3d.com/Manual/CollidersOverview.html`. Accessed April 4, 2020.

84

[Tecc] Unity Technologies. Manual - scripting. `https://docs.unity3d.com/Manual/ScriptingSection.html`. Accessed April 4, 2020.

[Tecd] Unity Technologies. Textmesh pro. `https://docs.unity3d.com/Manual/com.unity.textmeshpro.html`. Accessed March 6, 2020.

[Tece] Unity Technologies. Unity. `https://unity.com/`. Accessed April 4, 2020.

[TTLECR16] Katy Tcha-Tokey, Emilie Loup-Escande, Olivier Christmann, and Simon Richir. A questionnaire to measure the user experience in immersive virtual environments. In *Proceedings of the 2016 virtual reality international conference*, pages 1–5, 2016.

[UCAS00] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. Using presence questionnaires in reality. *Presence: Teleoperators & Virtual Environments*, 9(5):497–503, 2000.

[WD17] Rustin Webster and Joseph F Dues. System usability scale (sus): Oculus rift® dk2 and samsung gear vr®. In *2017 ASEE Annual Conference & Exposition*, 2017.

[WHEW⁺96] Donna Lee Wong, M Hockenberry-Eaton, D Wilson, ML Winkelstein, and P Schwartz. Wong-baker faces pain rating scale. *Home Health Focus*, 2(8):62, 1996.

[WS98] Bob G Witmer and Michael J Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.

[YP02] C Youngblut and BM Perrin. Investigating the relationship between presence and performance in virtual environments. *IMAGE*, 2002.

[YS17] Ryota Yoshimoto and Mariko Sasakura. Using real objects for interaction in virtual reality. In *2017 21st International Conference Information Visualisation (IV)*, pages 440–443. IEEE, 2017.