Institute for Visualization and Interactive Systems

University of Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Bachelorarbeit

Evaluation of “In-VR-Questionnaires”

Nico Haas

Course of Study:	Informatik
Examiner:	Prof. Dr. Niels Henze
Supervisor:	Dipl.-Ing. Valentin Schwind, Pascal Knierim, M.Sc.
Commenced:	15. Februar 2017
Completed:	15. August 2017
CR-Classification:	H.5.1

Abstract

Within the past ten years virtual reality has grown up and the topic became omnipresent. As a result, accompanying research in the realm of VR has increased, too. The two most popular terms describing the quality in VR are “immersion” and “presence”. Presence is mainly defined as the sense of “being there”, i.e. the feeling of being in the virtual world (instead of the real world). Measuring the level of presence by evaluating presence-specific questionnaires is a a well-established technique.

In a previous project my colleagues and I used a questionnaire that has been completed inside the VR. This was the first time to our knowledge an integrated questionnaire was applied in a Virtual environment (VE)) to this extent. While the project was conducted, the question arose whether these “In-VR-Questionnaires” can even improve Presence because the user inevitably can not see the real world during the whole experience.

In the scope of this thesis we carried out a study and conducted a factorial analysis variance to compare the main effects of the two different appearances of questionnaires (in VR vs not in VR) with other factors on presence. We did not find evidence that In-VR-Questionnaires (In-VR-Qs) have a significant effect on the experience in the virtual world in terms of presence. Yet, applying qualitative feedback like the “think aloud protocol” showed that there might be an increased focus completing the questionnaire in virtual reality (VR) and that In-VR-Qs may lead to more truthful answers of the participants. Furthermore, the collected data gave us more insight in the 3 chosen questionnaires and the correlations of their respective items. One result is the fact that a new presence questionnaire can be build on the basis of these data. Also we found item six of the *IGroup Presence Questionnaire* showing extremely high correlation in respect to the questionnaire itself which means this item may be used alone to achieve almost the same results as the whole 14-item questionnaire. Besides, the fact that both questionnaire appearances do not show significantly differences, also implies that you can choose which one you want to use depending on the present situation. Especially for research and gaming development like games during a testing phase, participants can give immediate feedback filling integrated questionnaires without leaving the VR.

Kurzfassung

In den letzten 10 Jahren ist die Entwicklung von Virtual Reality rasant gestiegen und wurde so zu einem mittlerweile allgegenwärtigem Thema. Infolgedessen ist ebenso der Anteil der Forschung, die sich mit dieser Thematik auseinandersetzt, größer geworden. Die zwei meist angewandten Begriffe in diesem Zusammenhang sind „Immersion“ und „Presence“. Mit Presence wird im Grunde beschrieben inwiefern sich der Nutzer in der virtuellen Umgebung wiederfindet und diese als „echte“ Welt wahrnimmt („the sense of being there“). Um den Grad an Presence zu ermitteln haben sich insbesondere auf Presence spezialisierte Fragebögen bewährt. In einem Projekt haben meine Kollegen und ich Fragebögen innerhalb der virtuellen Umgebung ausfüllen lassen. Nach unserer Kenntnis war es das erste Mal, dass in diesem Ausmaße ein integrierter Fragebogen in einer virtuellen Umgebung eingesetzt wurde.

Während der Durchführung unseres Projekts, stellte sich uns die Frage, ob diese „In-VR-Fragebögen“ die Presence sogar erhöhen können, da der Proband während der ganzen Studie nicht ein einziges Mal die reale Welt zu sehen bekommt. Im Rahmen dieser Thesis führten wir eine Studie durch, um die Auswirkung von den zwei verschiedenen Fragebögendarstellungen (in VR vs nicht in VR) auf die Presence zu untersuchen. Eine faktorielle Varianzanalyse wurde angewandt um die Hauptauswirkungen von den vorliegenden Faktoren, inklsuive Fragebogendarstellung, auf Presence zu untersuchen.

Wir fanden keine Beweise, dass In-VR-Fragebögen einen signifkanten Einfluss auf die Presence haben. Dennoch zeigte qualitatives Feedback, dass Probanden beim Ausfüllen der Fragebögen in VR fokussierter zu sein scheinen und deren Antworten wohl der Wahrheit entsprechen könne im Vergleich zur anderen Gruppe. Darüber hinaus ergab die Auswertung der gesammelten Daten , dass aus diesen Daten ein neuer Presence-Fragebogen erstellt werden. Zudem wies die sechste Frage des *IGroup Presence Fragebogens* eine extrem hohe Korrelation auf den Fragebogen selbst auf. Die Tatsache, dass beide Fragebogendarstellungen keinen signifikanten Unterschiede aufweisen, hat auch zur Folge, dass man sich je nach Situation die besser geeignete Variante aussuchen kann, ohne Verluste einzubüßen. Gerade im Forschungsbereich oder in der Spieleentwicklung kann diese Tatsache ausgenutzt werden um zum Beispiel integrierte Frageögen in Spielen, die sich in der Testphase befinden, angewandt werden um unmittelbares Feedback der Spieler zu bekommen, ohne dass diese die virtuelle Welt verlassen müssen.

Contents

List of Figures

List of Tables

1 Introduction

Applying questionnaires is a common method to gather information from respondents. With this way it is possible to get quite easily a large set of comparable data and evaluate these, eventually. In case of user studies questionnaires can be used for evaluating the quality of the experiments and therefore, making reliable statements. So it is hardly surprising that in studies related to VR such questionnaires are also being used fairly often. The importance and development of VR application has grown in recent years. So did research in this field as well. There are various concepts for describing requirements and measuring VR experiences. Two well-known terms are immersion and presence. Immersion describes the process of removing the perception of the real world and, metaphorically speaking, *immerse* in the virtual environment.

Whereas immersion can be seen as the technological point of view to get the user into the virtual world (defintion of Slater and Wilbur [SW97]), the term presence is mostly used as a result of this process - thus, it is a way of measuring the quality of the VR *experience*, also known as "the sense of being there". By now, measuring presence is well established. One way to achieve meaningful conclusions is the think-aloud protocol deployed by Turner et al. [TMTC03]. However, to get quantitative results the common approach is applying post-test questionnaires as for example proposed in the "Presence Questionnaire" paper [WS98]. Most of the time, qualitative measures are applied in addition to the quantitative ones gaining extra information which can not be collected through questionnaires.

In our previous work, we investigated the effect of using different hand styles on presence. Each participant had to test 6 different hand models and therefore, to complete the same the same questionnaire 6 times. So we implemented a questionnaire inside the virtual environment (In-VR-Q) for the first time (see **??**). The main reason for this was simply saving the time we had to spend for switching between the virtual and the real world. Moreover, we assumed that staying in the VE would maintain or even improve immersion and presence because the participants did not see the real world during the entire study.

Yet the question remained open, whether there is a noticeable difference in terms of effect on presence between the common questionnaires outside the VR and the proposed In-VR-Qs.

The aim of this thesis was to conduct a study using different levels of realism for the VE (task scenes) and both styles of questionnaires to find support for the assumption that the conclusion would be in favor of staying in VR to complete the questionnaires.

Figure 1.1: Completing questionnaire inside the virtual world

1.1 Vision

In the past years the VR standards have significantly improved. Virtual reality applications can be much more immersive due to progress in PC and VR technology. Along with this trend, the demand of immersive VR applications increases as well and so does research in this area. When measuring the experience in the VE, questionnaires are always a powerful tool to achieve meaningful results. So why should we not introduce questionnaires which will be completed inside the VE when we can easily build relatively virtual questionnaire environments (like in a lab in this study) with powerful and well-documented game engines like Unity. With the proposed approach, using In-VR-Q we could create an overall better experience in user studies and applications in general (which include some kind of questionnaires) with respect to VR in terms of getting a higher "presence score".

Document Structure

The chapters break down as follows:

Chapter ?? – ??: Basic variables and principles concerning experiences in VR and determining variables.

Chapter ?? – ??: Research in similar areas.

Chapter ?? – ??: Description of the study structure and project.

Chapter ?? – ??: Description of the conducted user study and results.

Chapter ?? – ??: Full summary and description of potential further research.

The following chapter "Theoretical Foundation" will present an overview of the most important terms and concepts with respect to VR and this thesis delivering the reader the necessary background knowledge . **??** "**??**" will firstly give an overview about the beginning of VR and its covered areas. Subsequently, the reader will be given a deeper look into the basic terms described in the chapter before. In the following, the wide variety of presence measurements will be presented and eventually, previous work in similar areas showing the reader how and why this thesis emerged in the first place. The next chapter "Design" will provide the setup and functionality of our project. After a small introduction about the study design its subdivided sections will give an Overview about the apparatus used in our project, the process of determining the three chosen questionnaires and finally, the section "Stimuli" will provide the information how the VE of the task scene is built up and shows which tasks the participants had to complete. The **??** "**??**" will initially present the measurements used in this study. The subsequent section "procedure" will explain the structure of the conducted study. Afterwards demographic information about the participants will be given. The last two sections will be presenting and discussing the collected data through quantitative and qualitative measures. The final chapter "Conclusion and Future Work" will provide a summary of the entire work we have done and at last, give insight about potential future work.

2 Theoretical Foundation

This chapter will present a theoretical overview of the main terms related to this thesis: “Immersion” and “Presence” as well as “questionnaires” in general and related to “Presence”. The following section will discuss the two most well-known terms describing the quality of virtual worlds. Differences in understanding of “Presence” and “Immersion” and thus, varying explanations emerged in the past and a clarification is need how these terms are used in this thesis. Also, in the second section, we will give an introduction in questionnaire in general and describe its usage with respect to “Presence”.

2.1 Immersion and Presence

To be able to discuss about VR and made experiences in it there should be universal terms describing it. Nowadays, the two most popular and frequently used concepts are *immersion* and *presence*. Still, there is a problem of misunderstanding: People use them in different ways. (Sometimes they are even treated the same way.) We regard the idea of immersion from a technological standpoint: To what extent the hardware and software can provide a virtual world in which the user will *immerse*. It is about removing the perception of the real world and let the virtual world become the current “real world”. Slater and Wilbur characterize immersion as “the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant” [SW97]. From this perspective immersion is an objective property of a virtual environment: So in theory you can differ applications in sense of immersion independently of peoples personal preferences. When you have two TVs which are the same, except for one having color television this one would be considered “more immersive”. Another way to increase immersion can be achieved when the system provide better graphics or higher resolutions. One important component of an immersive environment is the sound. That is also the reason why cinemas can create a greater immersive experience than a simple flat-screen TV without without an external sound system. Yet, Witmer and Singer defined immersion not only in a objective manner. It is a “psychological state” which is influenced of the given stimuli by the VE [WS98]. Thereupon, Slater wrote his paper “Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire” in which he clarify his view of immersion and also presence which is directly related to immersion. This already indicates the main problem in this field of research: The ambiguous notion of immersion and consequently, the significant confusion about the notion of presence.

So computer displays need not necessarily be a Head Mounted Display (HMD) or costly constructs like Cave Automatic Virtual Environment (CAVE) but with the progress in VRs technology HMDs seem to be the way to go in the near future. In contrast to the first HMDs they are now much more wearable, mainly because they become much lighter. In sense of sight you could create a perfect immersion with these HMDs because no matter where the user is looking he or she never sees the real environment around him (when the HMD fits and is adjusted correctly).

Apart from the visual and auditory aspect, interaction plays an important role creating a big immersion. The visual and auditory channels only can give you feedback in the virtual world, but to increase the immersion further you can let the user do things in the VE, manipulate it. One method is using controllers such as ones from consoles like Xbox. But HMD developer now have their own controllers like the VIVE Controller or the Oculus Touch Controller from HTC VIVE and Oculus, respectively. In addition to common controller like the ones you know from the PC or consoles, their position in the room is also tracked for example through a connection to appropriate sensors in case of the Oculus Touch Controllers. So when the user move his hands while being in VR the movement can be mapped into the VR experience so that he has the feeling that the arms and hands or whatever really belong to him. With realistic avatar models in the VR this can lead to big improvement in immersion. There is also another possibility to track hands even without using any controller at all: a Leap Motion attached to your HMD allows you to be in the VE and using your hand to manipulate the world around you through specific hand gestures. Wearing a HMD and still knowing where which part of the body is when we move our movable parts of the body is due to our ability of “proprioception”. When using a virtual avatar and tracking devices like mentioned above the key is a “consistent, predictable and complete relationship between proprioceptive information and sensory data” q. Slater and Wilbur also pointed out that “The greater the degree of body mapping [...] the greater the potential match between proprioception and sensory data”.

To sum up, immersion refers to the objective level of sensory transportation a VR can provide. The movie Matrix shows an example for a fully immersion: You are transported (immersed) with all of your five senses (sight, hearing, taste, smell and touch) into a virtual world. The only way to distinguish this world from the real world is by knowing that it is, in fact, only a virtual world.

This “knowing” or better say believing that you really are in this immersed world and not in the real world is expressed by presence. Lombard and Ditton regard presence as “an illusion that a mediated experience is not mediated” [LD97]. Presence refers to the subjective psychological response to the VR, the sense of “being there” (In contrast to technological point of view of immersion speaking in terms of the definition of Slater and colleagues). Schubert et al. describe it as “the outcome or a direct function of immersion” [SFR01]. But that does not mean that this “function” is the same for each person. Given the same immersive system, different user may experience different levels of presence. Additionally, different systems (i.e. different functions) can lead to the same level of presence.

Nowadays, authors distinguish presence in three dimensions depending on the kind of environment and interaction presence: "Physical presence", "Social Presence" and "co-presence". Most of the the time when people only speaks of "presence" they mean physical presence. It can also be named as "place presence" which already give a suggestion what it means. I There is also the term "tele-presence" which is the origin of presence. When tele-presence is not meant as presence itself it refers, in contrast to (physical) presence, to the sense of "being at a distance (real) place" in terms of remote control. In this context, presence alone means the transportation of a person to a virtual place, interacting and manipulating this virtual environment. Yet, many authors do not bother about this difference and see presence as a combination of both. Once again, you can see that the concepts describing VR are not consistent.

Social presence and co-presence describes the sense of "being together in the shared space". Whereas social presence refers to the *medium*, co-presence is seen as a more psychological factor, the possibility to perceive others and to know that they perceive them, too. This mutual perceiving is also the reason it sometimes defined as "shared presence".

To measure presence the kind of presence is mostly subdivided in different factors which are responsible for the characteristics. Factors can be *involvement, realism/realness, interaction, spatial presence, distraction* and so on. It depends on various factors like the dimension of presence, the environment or measurement method and of course, on the view of the researcher with his own understanding of what is important and shall influence presence.

2.2 Questionnaires

Questionnaires are an extremely effective method for collecting plenty of information from much people. Typically a questionnaire consists of questions with mostly predefined answers opportunities such as "yes" and "no" or in form of scales where you put in your mark in a range of x possible choices. The most common form of questionnaires were paper-and-pencil surveys for a long time. But with time web-based (or computer-based in general) questionnaires become much more popular because they are easier to create, change and evaluate, eventually.

Especially for evaluating the experience in VR many "presence questionnaires" emerged. Due to to the above-mentioned facts that presence is a multi-dimensional construct and the ambiguous understand of its terms and concepts there are also many different presence questionnaire to address presence.

In the following chapter we will discuss previous research on these terms and how they emerged and mainly give an overview about work with presence questionnaires that we considered using in this thesis.

3 Related Work

It is difficult to say who really the father of VR is. There was more than one person claiming this special title. Many say it may Ivan Sutherland for inventing the first HMD which was attached to a computer system, the “head-mounted three dimensional display” [Sut68]. But the first person who did research and and inventions in this area was Morton Heilig. 1960 he invented the first example of an HMD, the “Telesphere Mask” which was able to provide stereo sound and stereoscopic 3D. Furthermore he patented his own “experience theater”, the Sensorama Stimulator [62]. Although 3D movies did not exist back then, he invented a motion picture theater where all senses were engaged - which are the essential factors to create an immersive experience as we discussed in **??**. For example, he used an oscillating fan producing the wind in the hair to create this illusion. He stated that this kind of stuff - today known as VR - could be used for training armed forces and students. Especially the former is applied for a longer time now, back then he stated that “there are increasing demands today for ways and means to teach and train individuals without actually subjecting the individuals to possible hazards of particular situations.”

This point is also the reason for increasing usage of VR in therapy, lately (cf. [DCJ+07; PR08; RHS+00]). Rothbaum, Barbara Olasov, et al. showed that Virtual reality exposure therapy (VRET) can be quiet effective as normal exposure therapy with lower costs for arising extended sessions. Although, you can not verify the validity for VRET in general for now, other studies show that it can be helpful and further research is required.

The concept “telepresence” was introduced firstly in 1980 by Marvin Minsky [Min80]. It is the feeling of social presence without being physical present: Being in one place physically and nevertheless, acting through a remote machine in a different location. His notion was remote-controlled robots that could improve work conditions and address problems like“energy, health, productivity, and environmental quality” in locations like nuclear plants. He differ from already existing terms like “teleoperator” and “telefactor” in a way that “telepresence emphasizes the importance of high-quality sensory feedback and suggests future instruments that will feel and work so much like our own hands that we won’t notice any significant difference.” In contrary, according to Sheridan the term “presence” is used similar but, however, it refers to the ability to remote-control another body or object in a simulated world like in virtual reality and not a physical one. He introduced this extension of telepresence related to VEs in 1992 [She92]. Heeter describes presence as the sense of “being there” and divide presence in subjective personal presence , social presence and personal presence[Hee92]. Another term in the realm of VR is “immersion”. Witmer and Singer describe immersion as “a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provide s a continuous stream of stimuli and experiences” [WS98]. Furthermore, they pointed out that immersion directly affect presence: “ a greater sense of immersion will produce higher levels of presence”. They explicitly pointed out that they are not sharing Slater’s opinion of an objective immersion which we discussed in the previous chapter: “immersion, like involvement and presence, is something the individual experiences”. In Slater’s “Response to the Witmer and Singer Presence Questionnaire” he clarifies his point of view and stated that “metrics can be established which are descriptions of the system” differing levels of immersion. Also he sees Witmer and Singer’s definition of immersion “as part of my understanding of the meaning of presence” [Sla99]. In 1997, Lombardi and Ditton classified presence based on previous literature in social richness, realism, transportation, immersion, social actor with medium and medium as social actor [LD97]. The most applied conception of presence is *transportation*, the sense of “being there”.

As you can see, there are lots of different definitions and opinions about presence. It is therefore not surprising that measuring presence also leads to different approaches. One distinction of presence measurements is that between “subjective” and “objective”. The objective variant is characterized by measuring automatically produced user responses that are not invoked intentionally. In subjective measurements the user is asked to give information about his personal state during or after an experience in the VE.

One way to achieve objective results are psychophysiological measures by controlling physiological processes. Observing heart rate and blood pressure fall under the category of *measures* (cf. dillon2000aroused, dillon2002s, slater2003physiological). For example, Slater, Brogni and Steed ascertained with a “Break in presence (BIP)” that heart rate provide important information about the presence. Dillon and colleagues found that the user’s emotion has an effect on presence: Whereas the change of visual angles of the display had “no clear overall effect” but content (amusement, sadness, neutral) delivered a significant effect on the heart rate so the authors came to the conclusion that “further investigations of the relationship between presence and emotion are needed”. Checking the Validity of physiological measures Meehan et al. pointet out that “Among the physiological measures,Heart Rate correlated best with the Reported Presence” [MIWB01].

Another approach is measuring the Skin temperature (ST) and Skin conductance (SC). In further research Meehan et al. found out that ST also highly correlate with reported presence but ST did not. In the BIP experiment Slater et al. recorded the SC response as well which “indicates a similar type of arousal as caused by an induced anxiety”.

Although psychophysiological Measures are objective and continuous they also have obvious disadvantages such as expensive equipment, produced noise of the measurement itself or the need of wearing special equipment like electrodes can also disturb the experience.

In his research “Musings on Telepresence and Virtual Presence” Sheridan already annotated that due to the fact presence is a subjective phenomenon its measurements should be this way as well: “‘Presence’ is a subjective sensation [...] Therefore, as with mental workload and mental models, subjective report is the essential basic measurement.”

Before we will come to the most common post-test presence questionnaires we now first discuss alternative analyzing methods in the area of subjective measures. In contrast to the post-test variant, one apparent technique is to ask after information during the VR experience. IJsselsteijn and colleagues investigated in several experiments the change of the feeling of presence over time . In one experiment, the participant should estimate his feeling of presence during experience with help of a hand-operated slider (cf. [ID98; IJs04; IRH+98]). One clear advantage compared to other measures is getting time-variant information because feeling present can change during an entire VR session. However, this method gets directly part of the experience and thus, can affect the outcome of presence. In addition, participants can not estimate multiple factors during one experience.

Slater and Steed introduced a new interesting approach with their paper "Virtual Presence Counter". With this counter the number of transitions from the virtual world to the real world was measured. Because participants had to report their transition it only could be this direction and not vice versa. The results of the measured time being in the virtual world were correlated with other validated presence questionnaires.

Qualitative measures are widely used for gathering information you can not get from quantitative measuring methods and are normally applied in addition to other research methods like presence questionnaires or objective investigations. Often they are the reason for new uprising hypotheses. Freeman and Avons used in their study "Focus Group Exploration" to investigate presence in a qualitative manner. In different groups they watched video sequences either on a normal flat-screen TV or a stereoscopic display. In the 3DTV condition all groups described the feeling of "being there" without specifically using the word presence. They found support that presence seem to be a multidimensional construct: realism, involvement and naturalness enhance that feeling of "being there". They also come to the conclusion that "interactivity is not an essential facet of presence engendering systems" although, at the same time, they decline that this also means interaction would have no positive effect on presence. Heeter also questioned his participants after their "subjective experience of how much they feel like they are and what makes them feel that way" and with this approach, dissociate himself/herself from the "traditional objective concerns". Another efficient way to gain qualitative insights in presence is the use of the thinking aloud method. Basically, a participant should just feel free to express verbally his opinion so the examiner can record and afterwards evaluating them [TMTC03]. For example, Turner et al. gained by using this method additional results like that "the real world stimulated much more reference to significance and memory" [TTC+03].

The most conventional method for measuring presence is the use of so-called "presence questionnaires". They are applied post-testing wise which means they are administered after a participant had experienced a VE. As mentioned above, there is no universal definition of presence out there. Thus, there are many different presence questionnaires. Another reason for the high amount of available presence questionnaires is the fact that presence is mostly seen as a multidimensional construct and so this leads to a large set of possible questionnaires addressing different kinds of presence, of course also depending on the respective settings. Yet, we only present research whose questionnaires are reported and accessible.

In 1993 and 1996 Barfield and colleagues interpreted presence as “ego-presence” and include both telepresence and virtual presence (remote and virtual environment, respectively). The applied questionnaires based on criteria of mental workload measures consisted of 2 or 3 presence items which was found to be highly intercorrelated. In case of the 2 item questionnaire the changing variables were spatial landmarks and abstractness whereas the questionnaire with 3 items was used in an experiment with changing stereoscopy (mono vs. stereo), head tracking (yes/no) and field of view [BW93; HB96].

Based on the work of Hendrix and Barfield, Parent developed a new 10-item questionnaire. It addresses presence and realism but a usage was never reported [Par98].

Slater and colleagues deployed questionnaire in several studies based on following three topics: the subjects sens of “being there”; the extent to which the VE becomes more “real or present” than everyday reality; and the “locality”, the extent to which the VE was already visited instead of created of a set of images. They distinguish “ between external and internal factors that contribute to presence”. It is similar to the immersion and presence subject discussed in **??**: External factors are directly given from the “Immersive virtual environment (IVE)” and internal factors “internal factors determine the responses of different people to the same externally produced stimuli”, i.e. the same given external factors.. This so-called “SUS” questionnaire finally consisted of 6 items (initially of 3 items), was rated on a seven point rating scale and the final presence score arise from the number of 6 or 7 points answers. [SUS94; SUS95; UAW+99]. It is one of the most well-known and used questionnaires in presence research.

Taken 3 items of the SUS which directly addresses presence Nichols et al. build a 9-items questionnaire. Apart from the questionnaire they used 2 other methods to measure presence: The first was background awareness which had no big success but secondly, they used a reflex response in form of a “randomly timed ‘startle event’” which was rated with “no reaction, a verbal report of a reaction, or a physically noticeable reaction”. Interestingly, the response method correlated with two of the SUS items. With the exception of one all others items did not correlate at all. Besides, the experiment showed that the HMD lead to higher SUS results than the desktop display [NHW00].

Kim and Boccia created a questionnaire of 8 items which can be subdivided in two categories: “departure” and “arrival”. In this context arrival means the transporting in another world than the real word and departure describes the situation that the real world was never left. Yet, the results of the applied questionnaire did not reveal the hypothesized outcomes from the manipulations on presence [KB97].

Apart from the SUS, the questionnaire created from Witmer and Singer is probably the most popular questionnaire in the realm of presence and is by far the most cited one in Google Scholar [1]. It is called the “Presence questionnaire (PQ)” and contains 32 items accompanied by a seven point rating scale. They argue that presence arise from involvement and immersion. Based on this idea they grouped their items in “control factors”, “sensory

[1] 3157 citations, August 9th, 2017

factors", "distraction factors" and "realism factors" which shall influence immersion and involvement and thus, presence [WS98].

3 years after the release of the PQ, Gerhard et al. created a new one based on the work of Witmer and Singer. Adjusting the factors to a multi-user environment with animated agents they also added *awareness* and *communication* to address presence [GMH01].

The first questionnaire from Baños and colleagues with 77 items also contains many items of the PQ. The authors disagreed with the fact that reality judgment was only seen as part of presence at that time. With reality judgment they mean "to what extent is this experience real?" and not detailed graphics ("We are not talking of realism but of reality."). They admitted that presence and reality "are very close and related but keep being different". Their questionnaire was created to assess 9 concepts, but presence and reality judgment provided the largest part with 17 and 14 items, respectively. After applying factor analysis in the final version of the questionnaire only 18 items were left including 10 items addressing reality judgment and presence (5 each). But they also admitted that measuring these concepts also depends on the setup and can vary in each case (which also counts for the multidimensional concept of presence as well): "both concepts of reality judgement and presence are relevant for VR, but it is feasible that their importance was determined by the aim of the particular VR application [...] Therefore, different assessment instruments are needed, according to the specific field of interest". Unfortunately, the validity of the questionnaire was never reported.

Schubert et al. constructed developed a questionnaire containing items of previous work like the presented SUS and PQ and also created some new items. After a first study with 246 participants the 75-item questionnaire, they used in a second one only items related to presence and interaction factors and build the final version of the IGroup Presence Questionnaire (IPQ) consisting of 14 items [SFR01].

In an extensive study (16 conditions and 322 participants) Dinh et al. investigated the influence of 4 variables on presence and memory. Their 14-items questionnaire consists of one item which directly asked after the "level of presence" on a scale from 0 to 100 and 13 other presence related questions on a 1-5 scale. The main results were that auditory and tactile aspects had a great impact on presence whereas olfactory cues and level of detail (vision) had no significant influence [DWH+99].

Based on previous work Murray et al. investigated the supposed large effect of auditory aspects on presence and developed a questionnaire containing 5 presence items. All the presence items were taken from previous research and containing the subject hearing loss. Yet, just the presence related items are listed in the paper and the results of the questionnaire itself was never reported [GW95; MAT00].

In a first study Larsson et al. used their self-developed Swedish Viewer-User Presence (SVUP) questionnaire (developed in previous work) to investigate different perceived levels of presence between actors which interacted with the VE through a HMD and observers which were only watching the actors on a projector using headphones [LVK01]. After the experiment results showed that actors experienced higher presence and realism than the observers which were in contrast easier distracted by the surrounding (real) environment they did a another study which indicated that "the addition of sound to a virtual environment can significantly affect sense of presence, engagement and emotions of the users" [TS12]. The conclusion of a further experiment was that presence even more increased with better sound rendering. Thus, this experiments support the assumption of the important role of auditory aspects in a virtual experience.

Unlike the questionnaires above, Gunawardena and Zittle address social presence and not physical presence. Physical presence is about being in another place whereas social presence describes the "degree to which a person is perceived as 'real' in mediated communication". They tested their 61-item questionnaire with a virtual video conference and found high correlations with other researches with social presence [GZ97].

Similarly, other studies as well used Shared virtual environment (SVE) conducting studies to address social presence (or co-presence). This sense of "being together (with another)", as most authors describes social presence, was investigated for example in an experiment which studied the influence of mutual gaze (by making eye contact) [BBBL01], with a collaborative task with haptic feedback [BHSS00] or in a pilot test "the Networked Minds Measure of Social Presence" addressed social presence as multidimensional construct consisting of co-presence, psychological involvement and behavioral engagement [BHG01]. As discussed in the work with physical presence the authors researching social presence also came to the conclusion that realistic environments is net crucial factor in a VR experience but rather interaction: "Realistic gaze behavior seems to be more crucial than photographic realism in Personal Space in Virtual Environments 22 establishing the social presence of an agent" [BBBL01].

The Nowak&Biocca questionnaire addresses physical presence, social presence and "Co-Presence" (discussed in chapter **??**). The authors found that interacting with an image (of a head) increases presence (regardless of being a human avatar or an agent) rather than the partner has no image at all. But in general, there was no validity for increasing one of the three investigated dimensions of presence.

Schroeder et al. developed a 11-item questionnaire also including both presence items (three) and co-presence items (two items). The rest 6 items address collaboration and "contribution to task". They investigated the effect on these factors with different environments: "face-to-face, IPT-to-Desktop and IPT-to-IPT" where IPT stands for "Immersive Projection Technology" which nowadays is something like the CAVE. Even though they stated that both presence and co-presence resulted in different scores in each environment the validity of the questionnaire itself never was reported.

In previous work we investigated the effect of using different hand styles on presence in an interactive environment. Because we applied a questionnaire consisting of the PQ and 4 own items on 6 different hand models (human models: male, female androgynous and non human hands: abstract, cartoon, robot) which is very time-consuming we decided to implement an questionnaire in the VE. With this approach the participants stayed in the VR during the entire experiment without leaving it for filling out the 6 questionnaires. This was the first time to our knowledge that In-VR-Qs were applied, at least to this extent.

[SKT+17]. In further research on avatar hands (this time removing up to 3 fingers) Schwind and colleagues successfully applied the the usage of In-VR-Qs one more time.

Recently, Frommel et al. investigated the same concept on “serious game prototype with an integrated survey”: In a combination of racing and Spanish learning game they observed if a questionnaire which is integrated in the game provides better effects on presence than overlay questionnaires which they assumed to “interrupt the experience”. Indeed, evaluating the presence questionnaires (from Witmer and Singer) showed that the integrated survey lead to “significantly higher” presence scores. However, the study was not concluded in a VR setup but nevertheless, shows potential in the application the same construct in the realm of VR.

In the experiment of Bailenson et al., the participants were immersed once more in the VE to rate their social experiences with the two agents. The experimenter ask the 5 questions from the social presence questionnaire and the participants answered on a 7 point Likert scale which was shown above the agents head [BBBL01]. Unlike our procedure, participants had to left the VR for a recall test and only afterwards going back in the VE.

4 Design

In the following, we will introduce the basic setup of our project. We decided to use a 2 x 2 mixed factorial design. The between subject variable is the “questionnaire appearance” consisting of the levels “in VR” and “not in VR”. “Task scene” is the name for the withing subject component and its accompanying levels are “immersive” and “abstract”. These are the VEs where the participants can interact in terms of shooting drones and are the bases for completing the questionnaire. After the following section “Apparatus” in which the used hardware and software are presented we will discuss the choice of questionnaires in the second section of this chapter. Since authors have different understandings of the term presence and its multidimensional components we use three different presence questionnaires obtaining more different data for the purpose of evaluating. This combination of the three questionnaires shall give us a complete analysis on the subjective measures of Presence. The design of the project and the exact tasks for the participants will be discussed in the last section “Stimuli”.

Because we want to avoid the effect off participants being not so concentrated at the end of their questionnaires we used a “balanced Latin-square design” for the sequences of these three questionnaires. Therefore, we used the following three orders: Questionnaire 1- Questionnaire 2 - Questionnaire 3 (Q123), Q231 and Q312. One participant will first go into the VE and interact with, either in the immersive or the abstract scene. After that he will be filling out a questionnaire. Depending on Condition he will either be in VR (If ID number is odd) or not in VR (if ID number is even). Then the participant complete one more time a task scene but this time the other task scene. Finally, he will fill out the same questionnaire a last time. Due to our balanced Latin-square design, after 4 participants the order of the questionnaires will be changed. Note: the participants only sees one big questionnaire which consists of the items in the order depending on the condition. Table **??** shows one iteration of the study design to meet all conditions one time.

With this choice of variables and their accompanying levels we need 12 participants to meet each condition once. Thus, we need a multiple of 12 participants so each condition will be done from the same amount of participants. We finally went for recruiting 36 participants which is fairly doable in the scope of this thesis and particularly, quite enough getting viable data.

ID	Task 1	Questionnaire	Task 2	Questionnaire	Order
1	Immersive	IQ	Abstract	IQ	123
2	Immersive	OQ	Abstract	OQ	123
3	Abstract	IQ	Immersive	IQ	123
4	Abstract	OQ	Immersive	OQ	123
5	Immersive	IQ	Abstract	IQ	231
6	Immersive	OQ	Abstract	OQ	231
7	Abstract	IQ	Immersive	IQ	231
8	Abstract	OQ	Immersive	OQ	231
9	Immersive	IQ	Abstract	IQ	312
10	Immersive	OQ	Abstract	OQ	312
11	Abstract	IQ	Immersive	IQ	312
12	Abstract	OQ	Immersive	OQ	312

Table 4.1: Conditions

4.1 Apparatus

Primarily, our setup consisted of one desktop computer, one iMac11,3 (Mid-2010) and Oculus components. The desktop computer was equipped with a NVIDIA GeForce GTX 1060-6GB video card, an Intel i7-4790 CPU with 3.60 GHz and 16 GB of memory. The operating system was Windows 10 Education. The Hardware responsible for immersing in the VE consisted of Oculus devices: As HMD we used the Oculus Rift consumer version. As input devices we finally opted for the Oculus Touch Controller (first idea and implementation was with an wireless Xbox 360 Controller). And finally, we placed two Oculus Sensors on the table in front of the user - with approximately 1 meter distance between them. They are necessary for tracking the right position of the participants head and hands by means of communicating with the HMD and the Oculus Touch Controller, respectively. Due to the limited “mobility place” we had to put the participants chair 1 meter behind so he can move his hands in front without leaving the so-called “guarding system?” from Oculus. If you leave this area, correct position tracking is not guaranteed anymore and even worse, the Oculus guarding system shows this behavior with a grid-based wall. It’s safe to say seeing this wall would/will have a negative impact on immersion and presence. Between the two sensors the iMac is placed, also running the 64 bit version of Windows 10 Education. For our virtual world we used the Unity engine and the associated? Unity editor. The entire setup is shown in figure **??**.

4.2 Presence Questionnaires

In our research shown in **??** we figured out that many post-test questionnaires were created and applied in the last two decades. There, we listed most of the work which we

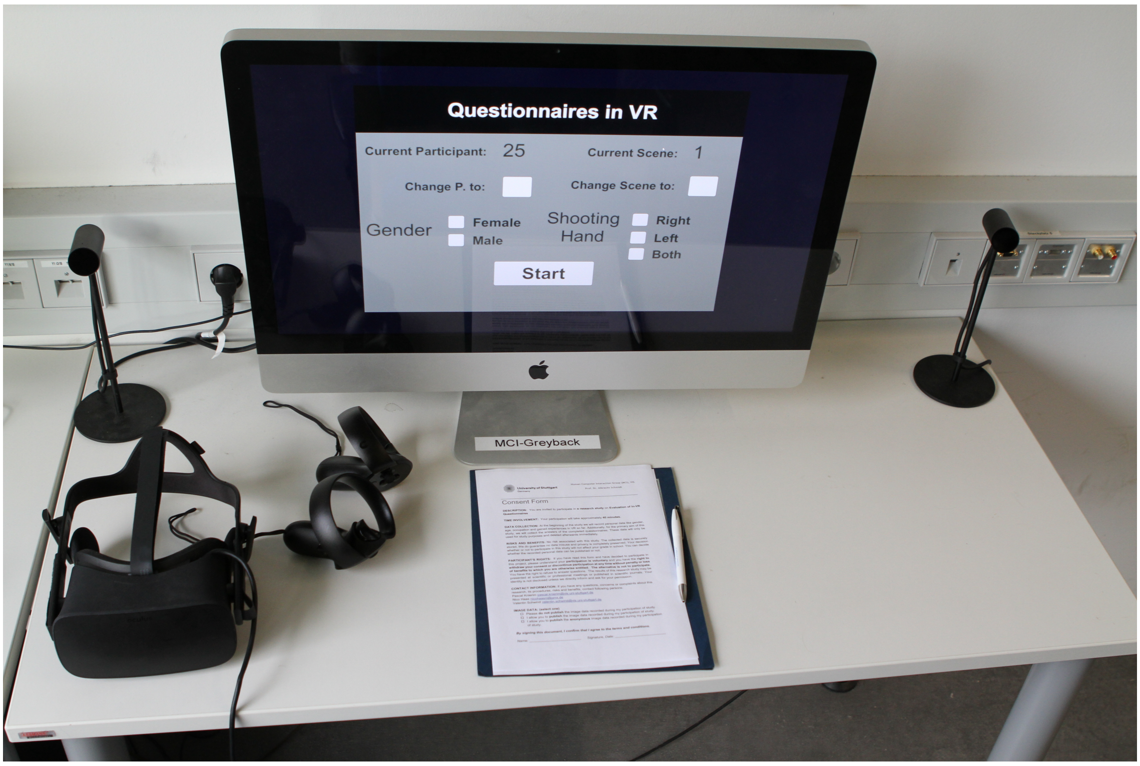


Figure 4.1: Setup with iMac (starting scene on screen) and Oculus Hardware

investigated. All relevant questionnaires mentioned in the previous chapter are listed in the **??** facilitating the overview. We left out research where the items of the questionnaires were never reported.

Two factors were essential in choosing the right questionnaires. First of all, the questionnaire must fit in our setup. Because presence is a multidimensional and ambiguous construct, different questionnaires address different factors with varying environments. Secondly, the questionnaire should be proven to be effective.

Gerhard et al. (3) extended previous work (8) to adjust on a multi-user setup. Since in our study only one participant at a time will be there, questionnaire 3 does not meet the first condition. The IPQ was not necessarily applied directly after a VR experience. Most of the participants completed the questionnaire on the Web and the questions were referring to the last made experience in VR. Due to this general applicability the IPQ (4) fulfill the setup condition. Also with 600 citations and the reported validity it is part of the shortlist. Because the validity was never reported for their questionnaires the work of Murray et al. (6), Baños and colleagues (10), Nowak et al. (13), Schroeder et al. (14) and the Memory Characteristic Questionnaire (MCQ) (19) were ruled out. Due to the social presence and the accompanying SVEs the work of following authors do also not meet our condition: Basdogan et al. (16), Gunawera and Zittle (17, “Globaled Questionnaire”) as well as Biocca et al. (18, “Networked Minds Questionnaire”). Bailenson and colleagues addressed social presence, too. But they used an agent instead of other real persons so the work could

	Questionnaire	Year	Citations	# Items	Comment
1	Barfield	1996 (93)	266 (311)	2 (3)	$\surd$
2	Dinh	1999	293	13+1	$\surd$
3	Gerhard	2001	48	23	Shared VE (SVE)
4	**IPQ**	**2001**	**600**	**600**	$\surd$
5	Kim&Boccia	1997	581	8	$\surd$
6	Murray	2000	36	10	V not reported
7	Nichols	2000	131	9	$\surd$
8	**PQ**	**1998**	**2933**	**32**	$\surd$
9	Presence&Realism	1998	6	8	Never used
10	Reality J&P	2000	123	77 / 18	V not reported
11	**SUS**	**1996/99**	**712 / 467**	**6** (3)	$\surd$
12	SVUP	2001	47, 9	17	$\surd$
13	Nowak	2003	469	18	V not reported
14	Schroeder	2001	146	10	V not reported
15	Bailenson	2001	296	5	SP
16	Basdogan	2000	393	8	SP, SVE
17	Globaled	1997	1648	10	SP, SVE
18	Networked Minds	2001 / 2002	237 / 189	40 (38)	SP, SVE
19	MCQ	1995 / 1996	26	21	V not reported

Table 4.2: Overview and Comparion of (accessible) questionnaires (V = Validity, SP = Social Presence)

fit in our experiment. Since it is the the last questionnaire addressing social presence and it is difficult to combine social presence and physical presence in one experiment, we decided against this questionnaire. The questionnaire of presence and realism (9) from Parent was only listed in her work but never used.

The remaining questionnaires (1, 2, 4, , 7, 8, 11, 12) fulfill both conditions. Because the IPQ (6), PQ (8) and the SUS (11) have proved to be successful, and are by far the most used and cited ones (see column “Citations” in table **??**) we decided to use them for our purpose.

4.3 Stimuli

For applying presence questionnaires we need obviously a VE where the participants can gather experience whose quality will be then evaluated by these questionnaires. For this purpose we decided to implement 2 different worlds: an abstract one and a more immersive world. In both VEs the participants task is to destroy drones with an abstract weapon or a more realistic pump gun (more detailed discussion section **??**). For the abstract scene

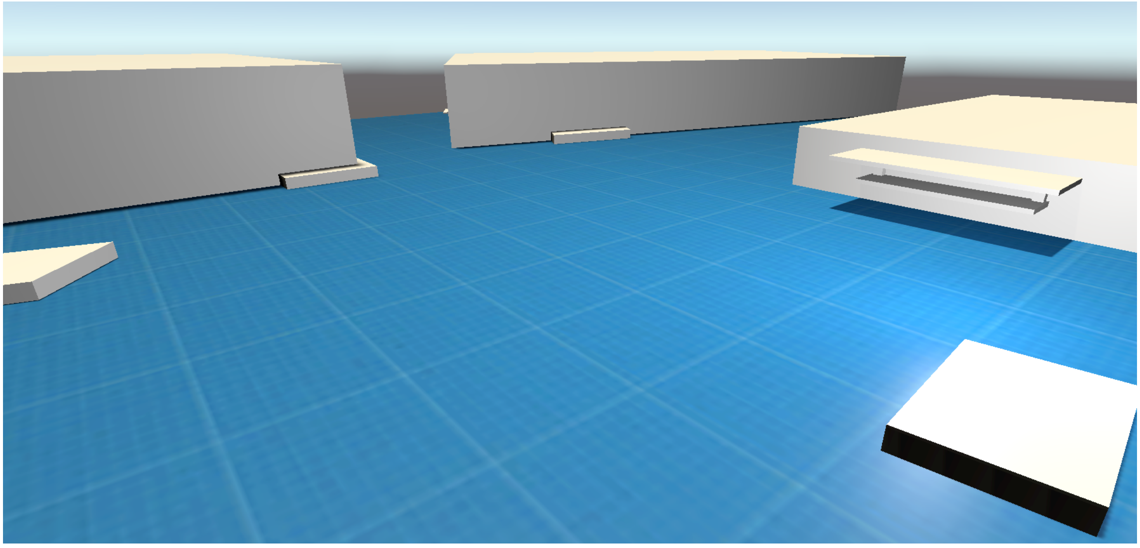


Figure 4.2: The abstract scene

(Which is the VE in the Unity3d editor) we have reached the decision that we make us of the unity project already implemented by us in our previous work where the idea of this thesis stems from [SKT+17]. In figure **??** you can see the structure of the abstract world: It mainly consists of simple default boxes from the Unity editor, just re-scaled to create an abstract environment. Besides, we changed the floor texture to a blue one so the participants can see more clearly the also simple made drones which are needed for the participants task. A drone is simply a greater white cube for the body and 4 smaller cubes which represent four drone propellers. Furthermore, the default skybox in unity is used for the background.

The immersive scene is mainly constructed by the detailed "Adam Demo Scene" made from unity which is available for free download in Unitys own *Asset Store* as a complete project. We only had to disable the so-called "PostProcessing script", because with it undesirable graphic errors in the form of black and white pixels in the main camera area occurred (when being in VR). In contrast to the abstract drones we purchased and imported the "Drone Controller(PC/Gear VR/Joystick/Mobile) + 3D Drone Models" asset [1].

The drones from the Asset Store are made for flying in an environment controlled by the user himself with VR-Controller, PC Input devices or even with a mobile phone. However, in my setup the drones are not used by the participants but are the targets for the shooting task. So we had to manipulate the original script to adjust the behavior to our need in self-flying drones. Apart from change the original code slightly we also wrote a new script where you can set flying direction, time when the drone will start flying and an *auto destroy time* so even when the participants do not hit the drone with their gunshots all drones will be finally destroyed after this time. The *auto destroy time* is particularly important for the implementation of the different waves which will be discussed in the section **??**. Both the

[1] https://www.assetstore.unity3d.com/en/content/61327

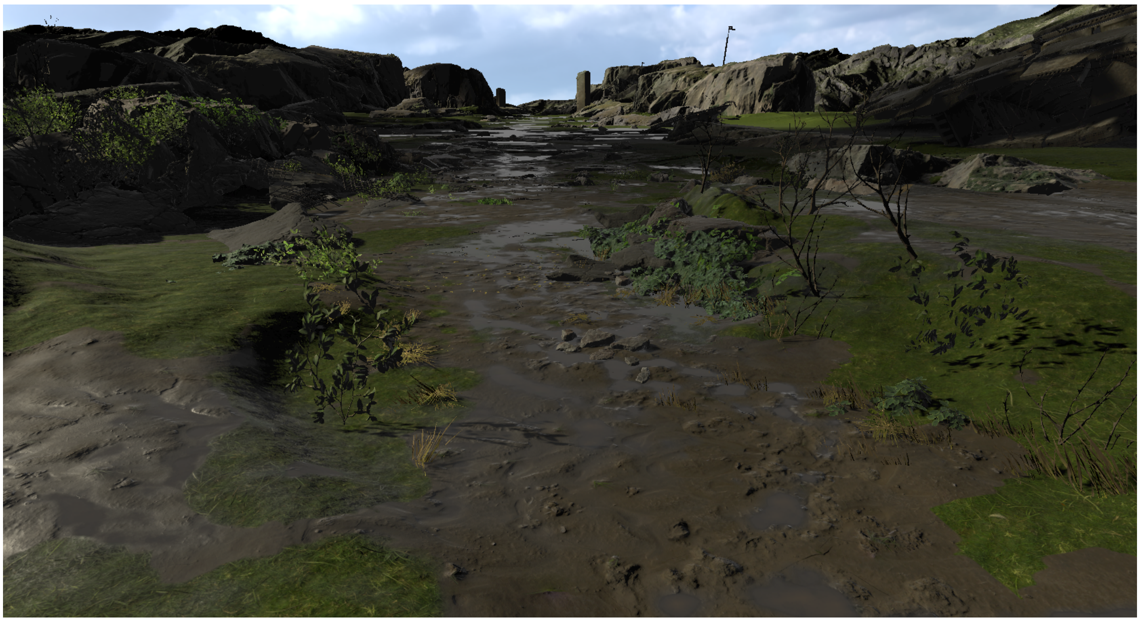
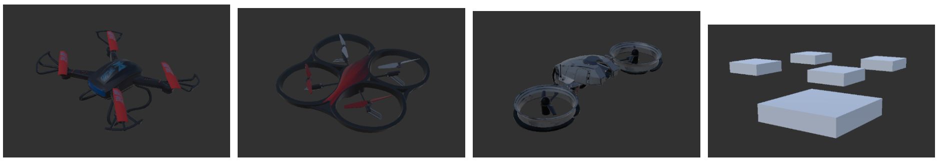
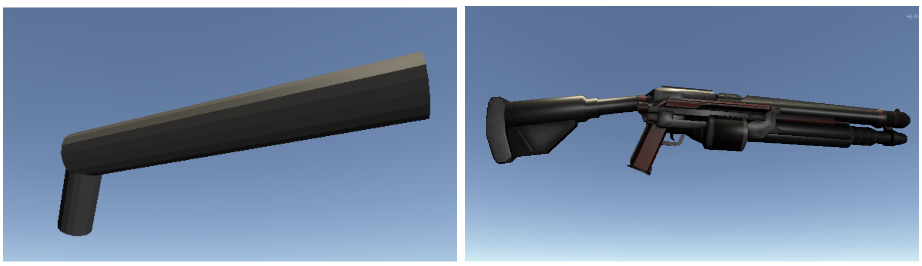


Figure 4.3: The immersive scene



(a) Buzz Drone **(b)** Red Drone **(c)** Spliter Drone **(d)** Abstract Drone

Figure 4.4: All drone models in comparison: (a)-(c) immersive drones and abstract drone (d)



(a) Abstract Weapon **(b)** Immersive Weapon

Figure 4.5: Both weapon models

original *movement script* and the new created *drone handler script* are used in the immersive scene and in the abstract one as well.

Furthermore, since the authors intention was not to create destructible drones or disappearing drones in general some kind of destruction animation and drone disappearing had to be implemented. To achieve this we imported the free asset *Explosion System* from the Unity Asset Store. This asset already includes physical explosions with animation, fire, smoke, sparks and sound in form of a Prefab (prefabricated game object with components like scripts and animations attached to it). With a little workaround we achieved our aim to destroy a drone having this explosion animation and make the drone disappear, simultaneously: First, a little script guaranteed that for each frame the position of the explosion object (which is by default disabled of course and therefore not visible in the virtual world) was aligned with the drones position. Second, when the drone will be hit by a shot the drone game object was destroyed with the *destroy(gameobject)* function so the drone just disappear and the explosion object simultaneously enabled. The consequence is that from the participants point of view, it seems the explosion completely destroyed the drone.

Thanks to the detailed *Adam Demo Scene* environment and the nice looking drones attached with a realistic sound and the explosion Prefab we created the basis for an immersive world. In the abstract scene we removed the explosion sound, smoke and spark achieving a less immersive world. In both scenes the main camera object was replaced by the camera prefab from the Asset *Oculus Utilities for Unity* which is - as expected - optimal for using the Oculus Rift and its specific configurable options. Additionally, the *Local Avatar* Prefab from another Oculus asset was attached to provide the full functionality of the Oculus Touch Controllers. With this Prefab you simply can simply add a weapon 3d model on each hand and just need to adjust its size and direction. We imported the free available *Shotgun* asset [2]. Since we chose a shotgun we also downloaded an appropriate pump gun sound from www.soundbible.com.

The abstract weapon was made of 2 simple re-sized pipes from the *City Props* low poly package which is now no longer available. A white noise formed the appropriate shooting sound.

In figure **??** you can see both weapons.

With these configurations and adjustments we achieved our goal delivering both, an immersive and an abstract world for the VR experience as mentioned in the opening of this chapter.

Without going into much detail: For shooting we implemented no assistance like a laser or cross-hair known from many shooting games even though it would be no problem at all to implement these as there already implemented functionality like in the *VRTK - Virtual Reality Tooklkit* asset [3]. The hit boxes were big enough and while shooting the direction

[2] https://www.assetstore.unity3d.com/en/content/26685
[3] https://www.assetstore.unity3d.com/en/content/64131

of the weapon barrel were checked if they hit one of these drone hit boxes the explosion mechanismus described above was performed.

Tasks

Before the actual task starts the participant is given the opportunity to look around and get an impression from the more detailed or abstract environment in the immersive world or abstract world, respectively.

Apart from the given environments and different kinds of drones the first three drone waves are exactly the same in both scenes. In the first wave all drones are located directly in front of the participant, varying in the height and are spread horizontally in the participants view. Except a small number of drones which only turn on its own axis all drones do not move. After testing with other members of the lab we decided to double the number of drones and increased the auto destruction time in the first wave in the immersive scene giving the participants more time to get used to the shooting mechanism. In the abstract scene, however, the testing showed that there were no problems with getting used to shoot: people learned very fast how to aim correctly.

In the second wave all drones are situated a little bit farther and higher, but in the way the participants sill can locate the drones very easy without using the head uncomfortable. One by one the drones are moving to he bottom and stay still right above the ground. Participants can only shoot at drones already started giving them a small challenge because they have to wait until a drone moves and then anticipate correctly and of course, hit the drone. But even if they would miss, they still have the chance and time to shoot the drones while they are just staying above the ground and waiting for getting destroyed. The third wave is a mix of the previous drones plus drones which are flying through the environment but always appear near the participant so that he or she can see and shoot them, eventually.

The immersive scene has in comparison to the abstract world a fourth wave, called “Crazy Wave”. 6 groups of two drones are flying one by one towards the participant but the next group only starts when the previous is destroyed by the participant or smashed against the wall of the Adam environment behind the participant. The 2 drones are randomly mixed and starting from different distance so the participant has the chance to shoot them both successively. The groups were set to have an extra speed in comparison to the group before. This wave also provides some kind of challenge and also can lead to immersive situations when the drone get destroyed right next to the participant showing the explosion within close proximity.

Note: The first idea of the drone appearances was to spread them all over the place and let the user anticipate where they or their sound come from. The problem was that there is no surround sound of the drones. You can only identify that drones are there but you can not exactly localize the position of them. Testing showed that these conditions are leading to more effort and frustration to catch the drones. Although, in principle, the implementation would be practical on Oculus side. While the project was build Oculus released an update which includes the possibility to set up 3 and more Oculus Sensors. Thus, make it possible to provide a fully 3D experience because the tracking of the Oculus Touch Controllers do not get lost while looking and acting on the other side of the default 2 sensors (in case of 2 sensors the body is in the way and would disturb the signal of the infrared sensor).

5 User Study

In this chapter we will firstly introduce the applied measurements and describe the procedure of our experiment. Afterwards we will show the characteristics of our participants. In the last two chapters “results” and “discussion” we will first present the data of our evaluation, both quantitative and quantitative results, and then will discuss these results with respect to our hypothesis and new emerged observations.

5.1 Measurements

The changing variable is the kind of questionnaire the participant will be completing. Depending on the condition the questionnaire is the In-VR-Q or the common type of questionnaire - in the real world on the real iMac. For getting a meaningful evaluation the aim was to build both questionnaire “worlds” in the way they look similar as possible, i.e. we tried recreating the real lab in the In-VR-Q scene. In both variants the questionnaire itself looks exactly the same - whether it is on the real iMac or on the virtual iMac. As you can see in figure **??** the questionnaire is constructed as follows: On the screen there is only one question at a time, whose font size was chosen big enough so the participants can read the question very clearly even if they sat one meter before the table (due to the Oculus own *guardian system*). The same applies to the font size of the anchors from the fixed seven-point Likert scale, which is placed directly under the question, of course. On the same level there are 2 arrows on each side to give the participant a visual hint to choose his answer by going right or left with the controller. At the beginning we planned using a Xbox 360 Controller for our project which is the reason for the icons looking in Xbox manner. Despite the change to the Oculus Touch Controller we let the icons as they were because first, the buttons and labels are very similar and second the D-Pad icon only shows the direction to the next or previous question in general. To go to the next you have to confirm your chosen answer with “A, B, X or Y” and then turn one thumb stick down (both controller can be used). This instruction you can see below the Likert scale. The idea for having a confirmation mechanism is giving people more time to rethink their answers and not rush through the questionnaire. Besides, with this mechanism we avoid skipping questions accidentally since the controller thumb stick is very sensitive (this failure often occurred during testing). The questionnaire was displayed on an Apple iMac11,3 (Mid-2010). We chose this computer because we already possessed this computer in both forms, in real and as Prefab for the Unity scene which perfectly suits our needs aiming for an identical mapping of the real lab.

Outside-the-VR-Questionnaire (Outside-the-VR-Q)

Basically, this is how the scene with the common questionnaire looks like: As full screen application on the real iMac in the real lab. As i mentioned both questionnaire worlds should be as similar as possible. In theory there were x options to display the questionnaire on the real iMac. The first try was to use the iMac as second screen. Although researching shows that is possible with later iMac versions like ours this option seems only work when the operating system installed is *macOS*. However, macOS was fully removed and replaced with Windows. So this option does not work. Secondly, we thought about using the iMac separately with its own application only consisting of the Outside-the-VR-Q and not the entire project. It would be a bit cumbersome but should work in general. But because we decided using the Oculus Touch Controllers we relied on the Oculus requirements. However, the iMac specifications are not good enough: For example the iMac lacks in providing USB 3.0 and HDMI connections for the sensors and the HMD as well. Since we only needed the iMac as screen the solution was simple and brilliant as well: We installed on both computers Steam and used the included "In-Home-Streaming" function which let you stream a game from one computer to another computer when both are connected to the same home network. Eventually, we used *Age of Empires 2 HD Edition* on Steam that first launches a setting pop-up before the actual game starts. This starting screen you can simply minimize and use Windows on both machines, simultaneously. Figure **??** shows this function in action.

In-VR-Questionnaire (In-VR-Q)

The aim was building the In-VR-Q scene so it matches as much as possible the real world lab. We used the Oculus own Prefabs *OVRCameraRig* and *Local Avatar* like we did in the task scenes. In addition, we enabled the option to see not only the hand but the controller itself as well. You can see how a participant fill the questionnaire in **??** Unfortunately, we could not replace the abstract hands with more realistic ones like we did in our previous work with the LEAP Motion Setup [SKT+17]. Apart from that a room was created with similar size and inserted free available 3D objects found in the internet. These objects were computers, laptops, TVs, tables, chairs and the Prefab from the iMac was placed on the table as in the real world. For the floor, the walls and the ceiling we took photos from the real lab and inserted them as a texture in each object, respectively.

The virtual reality lab you can see in **??**

NASA Task Load Index (NASA TLX)

Additionally, we make use of the NASA TLX which is a popular technique for measuring subjective mental workload (see Appendix A). This assessment tool contains 6 items, accompanied by 100 point rating scales subdivided in 5 point steps (21 gradations). The questionnaire includes following subscales: Mental demand, physical demand, temporal

Figure 5.1: Lab Scene in the VR

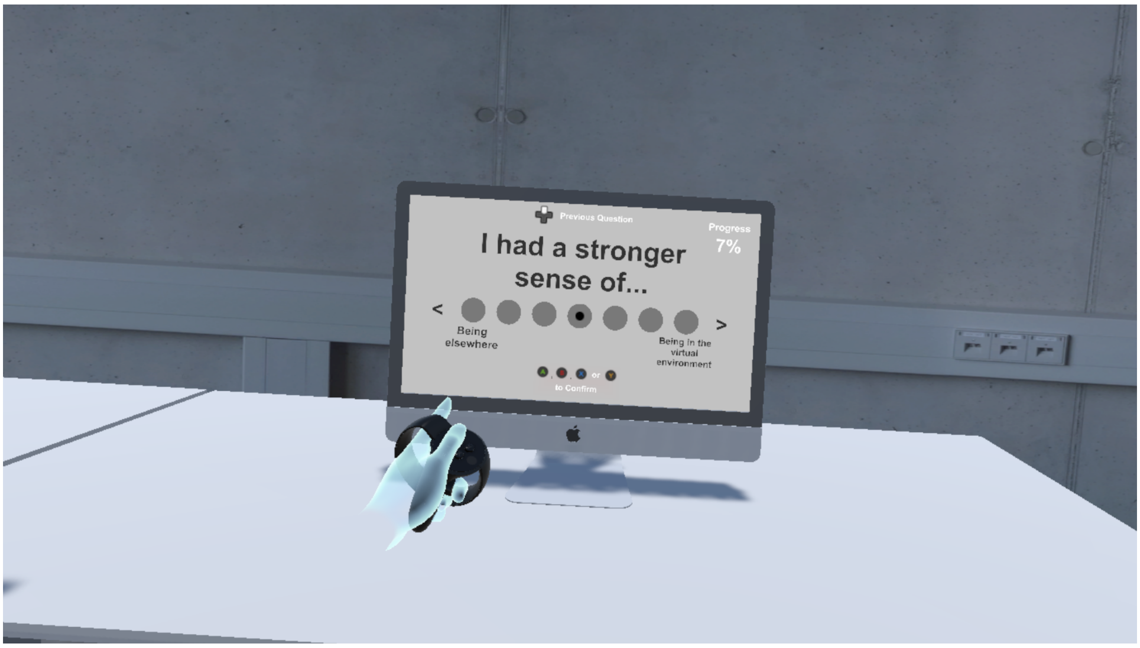


Figure 5.2: Participant fill questionnaire outside the VR

Figure 5.3: Participant fill questionnaire in the VR

demand, performance, effort and frustration. Furthermore, for better understanding and rating of the "questions" a description for each subscale is given.

We apply the NASA TLX to gain further insights into possible discrepancies between filling the questionnaire in VR and outside the VR. When completing the questionnaires is more exhausting in one of both, this would indicate using the questionnaire that is less exhausting to fill out.

5.2 Procedure

As described in chapter **??** the discussion of the study design resulted in a number of 12 participants to pass every condition one time. So we needed multiple of 12 participants and decided us for 36 participants for getting reasonable study results. One way recruiting participants was via the mailing list of theHuman-computer interaction (HCI) department. Either the participation counted as exam preparation for the human-computer interaction lecture or they got a compensation of five Euro. In addition, the email includes a Google Form link which was used for scheduling the participants study dates. On this site they should enter their email address for getting confirmations of their participation and gender so we could plan ahead the distribution on the right conditions and prefer the current gender minority in case of multiple sign ups for one date (which was the female part as expected). Further subjects we got by simply asking fellow students and in two cases, participants even asked room mates of their flat-sharing communities which gave us 4 more participants.

The study took place in the lab 0.023A, Pfaffenwaldring 5a, Stuttgart and was signposted at the entry. As the participant arrived, the examiner welcomed him and asked him to sit down before the table with the iMac on it. After a little bit small talk and questions concerning VR experience, information about VR in general, the Oculus devices and a short introduction of the study was given. Of course it was taken care of not giving to much information about the study goals which could lead to negative impacts on the results.

Afterwards, the declaration of consent was given to the participant which should take his or her time to read it carefully and eventually sign it. This document is mainly about providing demographic information and data collected completing the questionnaires which, of course, will not passed to third parties. These data was only used in terms of evaluation. Besides, the participant could check one of three boxes if we are allowed to publish image data or not and if so whether it should be anonymous (See also Appendix **??**).

The Participant filled in the demographic data form which included age, gender, Glasses yes/no, occupation and VR experience gained so far (See Appendix **??**).

Once the formalities were completed, the examiner went through the entire study-flow together with the participant: First, he described that the project consists of the 2 worlds/scenes followed in both cases by a questionnaire. Depending on the condition the exact sequence and kind of questionnaire was explained. Then, the examiner told him or her that the task is to shoot at drones which occur different waves. After he explained how to shoot and that the first drones appears with the first shot made, the examiner simply told the participant something like this: "Just try to shoot so many drones as you can. In the first round you have enough time to accustom to the shooting mechanism. I do not measure any performance at all, it is all about having a little bit fun in the VR and gain experience". We did not want the participant to get frustrated because this would reduce his or her overall experience. Depending on the control group the examiner told the participant that either the HMD will be taken of to fill the questionnaire or the scene will transition to the In-VR-Q scene and so on.

Finally, the examiner gave a introduction in the usage and operations principle of the Oculus devices like describing how the position tracking of the head and the hands works. Also, the participant was once again told how to shoot with the Oculus Touch Controller and how to navigate through the questionnaires: Turning the stick left and right to choose answers, confirm the chosen answer with the controller buttons "A", "B", "X" or "Y" and go to next or previous question by turning the stick down or up, respectively.

Before the actual study started, the last preparation was to put on and adjust the HMD correctly. To facilitate this, Oculus provides in his Oculus Home environment (only visible with HMD) a menu setting "Adjust your lenses" which gives instruction how to do so: First moving headset up and down while head stays still and 2nd with slider which is spotted on the bottom of the Oculus Rift Consumer Version and moves the lenses left and right. This procedure is especially for the control group in the In-VR-Q of importance because they have to read the questions inside the VR. All but one of the 13 participants wearing glasses had no problems putting on the HMD at all. The only one having problems with his glasses was in the other control group and thus, he could use his glasses again when completing the questionnaires outside the VR.

Lastly, the participant was told if any kind of questions arose he or she should just feel free to ask them (even though knowing that this could reduce the feeling of presence).

After all preparations were done, the examiner start the application “User Study” and in the starting scene, he entered the respective settings (gender and condition).

After the participants have finished the experiment we asked them to fill one last questionnaire: The NASA TLX. In this thesis the task which should be rated was completing the questionnaires in the procedure before (In-VR-Q and Outside-the-VR-Q).

Then the participant was given the five Euro or the examiner signed the printed form for the HCI lecture that he or she had participated in this study. To get some qualitative results the examiner also asked 2 personal questions after the participants was given the change to speak out his opinions, critics or whatever comments he or she had about the whole study and in particular, completing the questionnaires. The personal questions were like “Did you feel different completing the questionnaires” or “Where you more focused than completing a common questionnaire” which arose after two out of the first five participants from the In-VR-Q control group mentioned it.

5.3 Participants

A total of 36 people participated in the experiment. The ID of the participants were adjusted so each condition had 2 male subjects and 1 female subject. This means that the ID numbers do not correspond to the chronological times the respective participant conduct the study. Our sample was drawn from students of our university. The youngest participant was 19 years old and the oldest one 32 years. The mean age was 23.81 years with an standard deviation of 2.81. 20 participants had at least experience in VR like participation in previous studies or wearing a HMD at bigger events like Gamescom, the rest had no experience at all. About one third of the subjects were wearing glasses. As compensation the students either received an expense allowance of five euro or the participation counted as exam preparation for the human-computer interaction lecture. A detailed listing of all participants is shown in

5.4 Results

On average, participants spent 24.3 minutes (SD = 6.7) in the study. The average questionnaire completing time was 19 minutes (SD = 6.4). Nobody took a break or had to abort the study.

5.4.1 Quantitative Results

Independent from the kind of questionnaire, the immersive scene delivered higher presence scores than the abstract world. In the SUS the average answer for an item for abstract scene and immersive scene are 3.65 (SD = 1.11)) and 4.50 (SD = 0.93), respectively. The mean scores in case of IPQ were 4.16 (SD = 1.01)/4.77 (SD = 0.73) and the PQ resulted in 4.65 (SD = 0.59)/4.98 (SD = 0.51). Hence, with 0.82 difference the SUS showed the biggest distinction estimating the VR experience. The IPQ results show a difference of 0.61 whereas the presence scores of the PQ only differ with 0.31 “points” on an average.

In respect to the kind of questionnaire, the resulting presence scores for the three applied presence questionnaires SUS, IPQ and PQ in form of average ratings are shown in figure **??**. The results of the group of participants that completed questionnaires only in VR are represented with the red bars (“VR in”). Cyan bars are representing the counter part, the answers of the participants which filled the questionnaire on the real iMac. The black range-lines on the top of each bar illustrate the respective standard deviations.

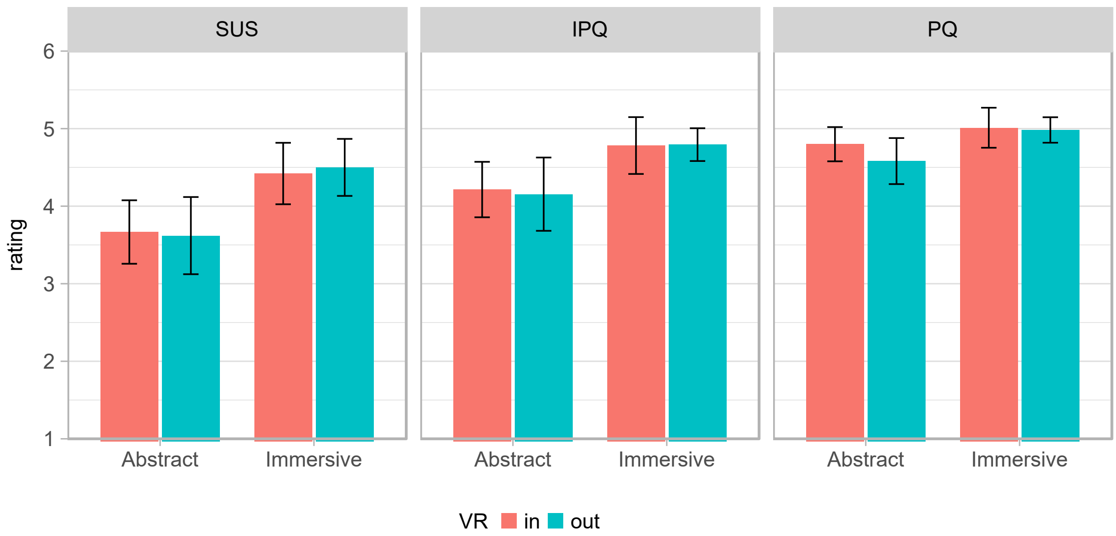


Figure 5.4: Presence Scores

Our fist approach for evaluating was computing the the differences between the presence scores of the In-VR-Q and Outside-the-VR-Q groups for which we will use the term “Presence Difference Score (PDS)”.

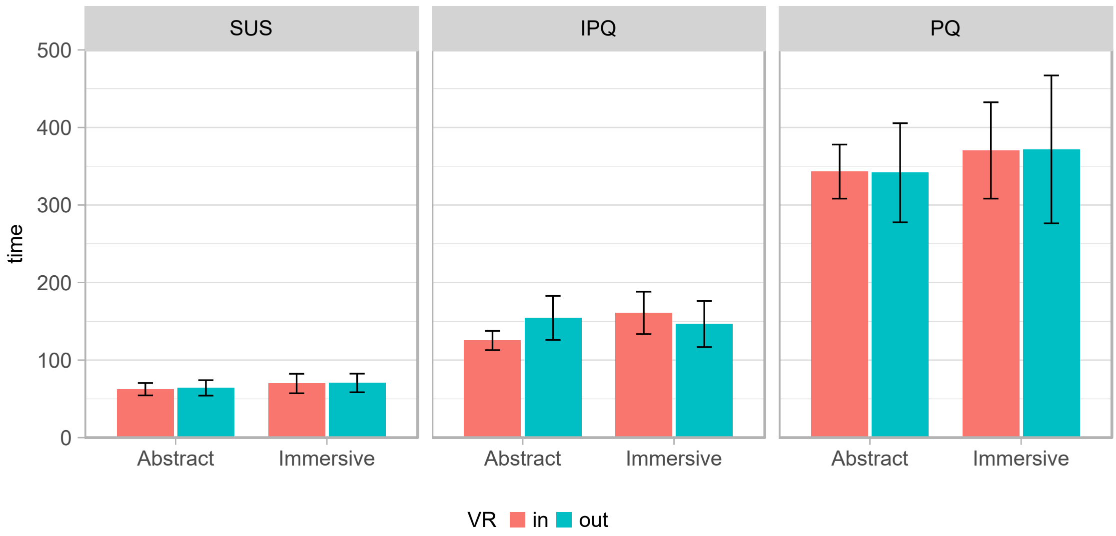


Figure 5.5: Completing Times

Be $Q \in \{SUS, IPQ, PQ\}$ the presence questionnaire and $TS \in \{Immersive, Abstract\}$ the respective task scene. The PDS is computed using the following formula:

$$PDS(Q, TS) = PS_{InVR}(Q, TS) - PS_{NotInVR}(Q, TS) \tag{5.1}$$

Where $PS_{InVR}(Q, TS)$ is the presence score of the questionnaire Q applied to scene TS in the In-VR-Q group and $PS_{NotInVR}(Q, TS)$ is the respective presence score for the Outside-the-VR-Q participants.

Thus, a positive value indicates a higher presence score in the case of the integrated questionnaire in comparison to the questionnaire outside the VR.

We get the following PDSs for the SUS scenes...

$$PDS(SUS, A) = 3.67 - 3.62 = 0.05$$
$$PDS(SUS, I) = 4.42 - 4.50 = -0.08$$

and for the IPQ scenes...

$$PDS(IPQ, A) = 4.21 - 4.15 = 0.06$$
$$PDS(IPQ, I) = 4.78 - 4.79 = -0.01$$

and finally for the PQ scenes we get:

$$PDS(PQ, A) = 4.76 - 4.60 = 0.16$$
$$PDS(PQ, I) = 4.99 - 5.01 = -0.02$$

As second approach a three-way analysis of variance was conducted on the influence of three independent variables (task scene, questionnaire appearance, questionnaire type) on presence in terms of presence score evaluated by completing the presence questionnaires. The effects included two levels were task scene (abstract, immersive) and questionnaire appearance (in VR, not in VR). Questionnaire type consisted of three levels (SUS, IPQ and PQ). All effects were statistically above the .05 significance level except for the task scene and questionnaire type factors. The main effect for task scene yielded an F ratio of $F(1, 204) = 23.4446$, $p < 0.0001$, indicating a significant difference between abstract ($M = 4.15$, $SD = 1.02$) and immersive ($M = 4.73$, $SD = 0.78$). The main effect for questionnaire type yielded an F ratio of $F(2, 204) = 14.2222$, $p < 0.0001$, indicating a significant difference between SUS ($M = 4.05$, $SD = 1.10$), IPQ ($M = 4.47$, $SD = 0.93$) and PQ ($M = 4.81$, $SD = 0.58$). However, if a questionnaire was completed in VR ($M = 4.48$, $SD = 0.94$) or not ($M = 4.45$, $SD = 0.96$) did not have an significant effect on presence, $F(1, 204) = 0.0109$, $p = 0.9169$ » 0.05. The interaction effect between task scene and questionnaire appearance yielded an F ratio of $F(1, 204) = 0.3162$, $p = 0.5745$ » 0.05 which also indicates no significant effect on our measured construct, presence. Likewise the interaction effect between task scene and questionnaire kind($F(2, 204) = 1.4020$, $p = 0.2485$) and between the factors questionnaire appearance and questionnaire type ($F(2, 204) = 0.0773$, $p = 0.9257$) was not significant, too. And the interaction effect between all 3 factors on presence was also not significant, $F(2, 204) = 0.0150$, $p = 0.9851$.

Likewise, we conducted a Factorial analyses of variance (ANOVA) with the same conditions on “Completing Time”.

Now the three-way analysis of variance was conducted on the influence of the same three independent variables (task scene, questionnaire appearance and questionnaire kind) on the time (in seconds) needed to complete one questionnaire. There was only a significant effect of questionnaire type, $F(2, 204) = 151.3253$, p « 0.0001, indicating a significant difference between SUS ($M = 66.44$, $SD = 26.27$), IPQ ($M = 146.95$, $SD = 62.84$) and PQ ($M = 358.92$, $SD = 162.86$). The effects of the remaining factors were all not statistically significant at the .05 significance level. If a questionnaire was filled in VR ($M = 570$, $SD = 185.26$) or not ($M = 574.42$, $SD = 261.17$) had no significant influence, $F(1, 204) = 0.0094$, $p = 0.9230$. Neither had the task scene which main effect yielded an F ratio of $F(1, 204) = 1.2035$, $p = 0.2739 > 0.05$, indicating that the effect of being in the immersive scene (M = , SD =) or in the abstract scene (M = , SD =) before the questionnaire scene was not significant. Additionally, the interaction effect of all possible combination was also not significant: The questionnaire type neither has an significant effect in combination with task scene: $F(2, 204) = 0.2068$, $p > 0.05$, nor between it and questionnaire appearance: $F(2, 204) = 0.0566$, $p > 0.05$. The interaction effect between task scene and questionnaire appearance yielded an F ratio of $F(1, 204) = 0.1828$, $p > 0.5$, which also indicates no interaction effect on time. Last, the interaction effect between task scene and questionnaire appearance and questionnaire type was not significant, too: $F(2, 204) = 0.2734$, $p = 0.7611$.

From the analysis of the filled out NASA TLX forms, it emerged that in the case of being outside the VR the subscale “Mental Demand”, “Physical Demand” and “Frustration” have a slightly higher amount of participants who voted with 14 or more “points”. “Temporal Demand” and “Performance” show the opposite behavior.

The total scores of the fist subscale "Mental Demand" were the highest one among all subscales but do not show any significant differences in the questionnaire appearance: Participants in VR rated with 158 points in total, the other 18 students come to 159 points. The standard deviations values are 4.64 (M = 8.78) and 6.03 (M = 8.83), respectively.

Interestingly, in terms of *Physical Demand* people in VR rated with 2.67 points in mean (SD = 1.70) whereas the mean value yielded for the Outside-the-VR-Q is 4.22 (SD = 5.05) although people have the same control devices (Oculus Touch Controllers). This may indicates that reading on the real iMac screen was more exhausting than reading on its 3D object counterpart in the VE. But overall both means were below 5 and so are rated almost "Very Low" (left anchor).

The subscale *Temporal Demand* showed that people felt less "hurried or rushed" in the real world (M = 5.94, SD = 3.82) than in the VR scene (M = 7.28, SD = 4.72).

Evaluating the subscale *Performance* also showed higher values for the group in VR (Mean = 5.11, SD = 4.62) in comparison to the students outside the VR (M = 3.44, SD = 2.89). But with the higher mean of 5.11 both are near the "Perfect" anchor on average.

The means for *Effort* in VR and not in VR are 6.33 (SD = 4.89) and 4.94 (SD = 5.08), also indicating no significant differences and also no big influence on task load in general.

Surprisingly, the subscale *Frustration* with an overall average value of 9.51 indicates a relative high impression of feeling"insecure, discouraged, irritated, stressed, and annoyed" regarding the fact that they "only" have to fill a questionnaire. Maybe the formulation of the questions and the length of the whole questionnaire (52 items) played an important role..

The figure **??** shows the computed "raw TLX" Score (raw means that subscales were treated equally without weighting due participant comparisons of pairwise subscales).

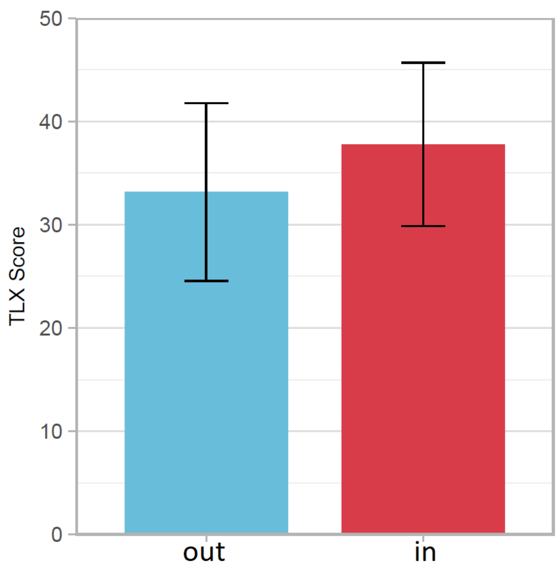


Figure 5.6: TLX Scores

The right red bar shows the NASA TLX score of the participants completed the questionnaire in VR. The standard deviation you can take from the black range line on top of the score. The same counts for the Outside-the-VR-Q counterpart represented by the left cyan bar. As you can see there is a difference of about 5 points in favor of the Outside-the-VR-Q group (higher score means higher task load).

Additionally, to evaluate if the participants in VR felt significant less exhausted completing the questionnaires we conducted a Single Factor ANOVA on the influence of the independent variable *questionnaire appearance* on the NASA TLX score. As before, the questionnaire appearance consisted of two levels (in vr, not in vr). Yet, the main effect for questionnaire appearance yielded an F ratio of $F(1,34) = 0.472$, $p = 0.4967$, indicating a non-significant difference between being in VR ($M = 36.44$, $SD = 15.59$) or not ($M = 34.61$, $SD = 23.06$) to fill the questionnaire.

In addition to these approaches we also investigated the resulting data in terms of correlations. For this purpose we deployed a correlation matrix with all items and the questionnaires themselves to determine whether an item is related to another item or the whole questionnaire. The corresponding matrix with p-values indicates if obtained results are statistically significant ($p < 0.1$) or not.

The highest correlating item of the SUS was item 2 with a factor of 0.74 and $p < 0.001$.

Item 6 of the IPQ with respect to the entire questionnaire showed with 0.8 point and p « 0.001 a extremely high correlation between the result of item 6 and the result of the IPQ itself.

In case of the PQ item one has the highest correlation coefficient with 0.67. Item 10 and Item 27 showed a coefficient of 0.64. All three items have a p-value smaller than 0.001.

5.5 Discussion

The first approach using PDSs did not yield satisfying numbers. The differences from presence scores between the In-VR-Q and the Outside-the-VR-Q groups are marginal. This already indicated that the influence of the questionnaire appearance is not significantly. We tried getting deeper insight using the three-way factorial analysis of variance. We only got significantly effects on presence from the factors task scene and questionnaire. That the choice of the task scene has a big influence is not surprising. We built the respective scenes in a way that one is more immersive than the other hoping it results in higher presence scores, like it did. This indicates confirmation for the assumption that the presence may be a “direct function of immersion” like Schubert et al. stated in their work [SFR01]. The fact that the different questionnaires resulted in difference overall presence scores is simply due to the different amount of items in them. The same applies to the analysis of variance of the time: Of course, participants need more time for filling out the PQ consisting of 32 items than completing the SUS or IPQ with 6 or 14 items, respectively. Unfortunately in both cases, the factor questionnaire appearance had no significant influence on presence and neither did interactions with the other factors. Also, the analysis of the NASA TLX showed no significant effects of In-VR-Qs on task load. So these quantitative measures lead to the result that In-VR-Qs have no significant influence on presence.

Nevertheless, “there was no significant effect” does not mean that there are not positive conclusions to draw.

First, let us closely take another look on the observation made that there is no significant effect: That there is no *significant effect on presence* also means that there is **no significant negative effect**! The first conclusion is that you can insert breaks in VR applications without significant loosing presence. When it does not matter if you are filling out a questionnaire in or outside the VR in respect to the perceived presence than you could do a break in that time, too. This may be helpful for people suffering from *simulator sickness*: If the breaks do not have a significant effect, that means that someone spending the same amount of time in an application feel probably not more present than if he would do with a certain amount of breaks. Participants spent about 19 minutes (SD = 6.4) filling out the questionnaire. We assume that breaks lasting this amount of time would then also not significantly disturb the experience in the VR. Secondly, in combination with the NASA TLX results that indicates that participants were not significantly more exhausted (perceived workload) than the other group of participant this also means for designer of VR applications with surveys that they can choose if they want implement an In-VR-Q or Outside-the-VR-Q. A potential application area could be games in the testing phase (like “early access” games or games in the beta-phase): The player can give immediate feedback without leaving the VR. So he or she can fill out the questionnaire and directly move on with gaming.

Besides, the evaluation of the correlation matrix revealed that item 6 of the IPQ highly correlates with the respective questionnaire. This result indicates that using this item alone could be used as questionnaire on its own and would provide almost the same results. Interestingly, question 6 was: “I felt present in the virtual space”. Lessiter et al. proposed that an “ideal presence questionnaire [...] First, an understanding of presence should not be assumed by directly asking respondents how present they feel”. However it is not clear how many of the participants have ever heard of the term presence.

Furthermore, we collected qualitative information for gaining further insights and possibly discovering aspects whose are not arrived by the quantification used before. Achieving additional data we were applying the “think aloud protocol” and asked personal questions after the entire experiment. Yet, for minimizing the negative impact on presence, the think aloud protocol was mainly done in the questionnaire scenes. The experimenter asked the participant to concentrate on the tasks and environment when being in the drone environments but of course, if question arose or something else they had to say they were free to talk. The resulted assumptions and conclusion based on these investigations are presented in the following.

TLX Analyses - Factor Focus

As already seen in the section before, the quantitative results show that there was no significant influence when filling out the questionnaire inside the VE or outside. In addition, the evaluation of the NASA TLX indicates that the choice of questionnaire appearance even has no influence on the work load which basically means: first, the presence results are the same and second, completing the questionnaire itself has no significant difference either. But in return, we made the following observation: The examiner had a feeling participants in the Outside-the-VR-Q-group got easier distracted by the (real) environment and asked more questions than the other participants from the other group. Even if you ask the participants to concentrate on the study or specifically on the questionnaire in that case there always are times in the lab when probably more people are in the room conducting their own studies or working in groups on their own projects and eventually, there will be ever moments were they can get distracted. That does not mean they take their time for a small talk but rather that first, through the noisy environment it is harder to concentrate and secondly, the participants may get the feeling being watched, precisely their answer choices. This may be just pure speculation and the examiner was not looking more deep into it cause he wanted the participants not bring in an uncomfortable situation. But qualitative feedback indicated the opposite behavior. After the experiment at the end of the talk the examiner asked participant with ID 7 the obligatory question if he or she has any further questions or comments on the study, especially on filling out the questionnaire inside the VR. The answer was: "At the very beginning I need getting used to it, especially with one big question about the entire screen. But after the first two questions it felt just like a normal questionnaire and to be honest, being in my own world I was more *focused* than I was the last times I filled out questionnaires." After this response I added the question to In-VR-Q group members with respect to a potential boost in terms of focus. And surprisingly, especially after evaluating the quantitative results, six participants confirmed, two participant denied the question and the rest could not say if there was a difference (not to forget the three participants before which are also in the In-VR-Q group). When we take the NASA TLX and look at the subscale *performance* with description "How successful were you in accomplishing what you were asked to do?". They just have to choose their own answers and move on, nothing difficult. Which you also can take from the NASA TLX results. But maybe we get difference answers when we added to the description "and how well could you concentrate/focused?" Besides it is difficult to measure if they answers "correctly" meaning that their answers comply with their real opinions about that. More distraction could lead to not filling the questionnaires as they would when they were totally alone in a room and sure nobody is observing them.

Back to reality: Is the time being in VR relevant?

Another "effect" was observed when the students finally put off their HMDs. In case of the Outside-the-VR-Q group there were no reactions at all. But most of the other participants staying the entire study in the VR reacted with short statements : "Oh, now I'm back in the real world, what a pity" (P13) or "Already finished? Can we start again?" (P17) and in almost every case combined with sighing. Some participants only sighed which also can be seen as a sign of exhausting. These behavior differences arise the question: Is the time being in VR an important factor, too? Even though the applied ANOVAs did not show any significant effects on presence these reactions may indicate people were more captivated by the VE when being in the VR the whole experiment. If this is the case, our short task scene experiences with mean times of 102 and 220 seconds for the abstract and immersive scene, respectively, could have been a limiting factor.

Lab Environment (In-VR-Q)

Based on the feedback the mapping of the real lab environment was highly successfully. Only two female participants expressed some critics but with respect to the illustration of the Oculus Touch Controller accompanying hands. In contrast to the Leap Motion device which tracks almost the full hand (bones) the touch sensors of the controller can only recognize if a finger is on the button or not and thus, the representation of the thumb in the virtual world is either something like a "thumb up" gesture or it is laying on controller surface. The two reactions were a continual response like "These hands are so weird. [...] This is still absolutely weird!" (Participant 29) and "That (behavior of the thumb) is really disgusting" (P25). But other than that, in reply to my question if they meant the environment they both negated: "No, the lab looks really nice. I like the windows" (P25). From the 18 participants in the In-VR-Q group, all but three provided positive feedback. Two others did not say anything about the lab and the other one just stated that the distances were not considered perfectly (apparently someone, who is often in this lab). 4 people even mentioned they were "missing" me because they talked with (to) me but could not see me: "(Participant enters Q-Scene and looks around) Cool, but there is no Nico (examiner) here." (P9) or "Ahh, every time we talk I look left and can't find you." (P3). The latter impressions may indicate presenting the experimenter with an avatar in the VE could increase the overall investigated presence, eventually, in a "transformation" manner: If there is an effect of the proposed In-VR-Q on presence (in this study we found no evidence but maybe there will be future work who can find support for our hypothese with other settings or improvements) there could be function F: Presence(Q in VR) $\rightarrow$ Presence(task scene) (similar to the immersion -> presence function of Schubert et al. [SFR01]) which describes the individual function that a user's perceived sense of being in the In-VR-Q environment directly effect the overall presence. That does not mean that it is an essential factor of feeling presence in the actual VE. It illustrates that the perceived presence could be also influenced by the perceived presence of the questionnaire scene: Presence(ActualScene) = Presence(Actual Scene) +- influence of Q-scene.

6 Conclusion and Future Work

With the rise of VR technology during the last decade commercially available products become cheaper and widely available. It is no longer just a research “thing” but games and console manufacturers now invest lot of money in new products and features with respect to VR. But not only the gaming industry is concerned, in almost all imaginable areas VR is now a part of such as in the area of education, medicine, military, sports ,broadcast and so on.

In this thesis, we investigated whether integrated questionnaires (in-VR) have a significant effects on presence. After creating two task scenes with different level of immersion, two scenes for completing the questionnaires in VR or on the Computer in the real world, testing and conducting the user study we came to the conclusion that In-VR-Questionnaires (In-VR-Qs) can have an effect on the experience in VR based on qualitative feedback, even though quantitative results did not support our hypothesis. Besides, examining the collected data we found that with these data it is possible to build an own questionnaire evaluating presence and to validate with our apparatus.

So even if we did not found evidence for supporting our main hypothesis, we believe that it is worth implementing In-VR-Qs in general. It is not difficult anymore to build a questionnaire scene or integrate a questionnaire in a existing scene, even for non-professional. Although no *significant* effects were found, there are almost no negative effects either. Disregarding simulator sickness, nothing speaks against the introduction of In-VR-Qs in both research and gaming areas or whatever field of interest you like to apply integrated questionnaires.

Limitations and Future Work

The discussion of the quantitative and qualitative results revealed some limitations of our project and potential further research which can be done in the future.

Eliminating Disruptive Factors and Adding more Time in VR

To get reasonable data it was an important issue to map the real lab environment as good as possible so both groups (in VR, not in VR) have the same conditions except being in the virtual world or in the real lab. Nonetheless, there are several other factors that can influence the outcomes of the experiment. Especially the questionnaire items addressing the real environment are of importance in this case, like item 7 from the IPQ: "How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)". During the study there were times when the real lab environment was empty and was nearly full as it could be. Also the shooting mechanism with the weapon of the immersive scene seemed to be different for participant. Some adapted very fast and others had almost problems the whole iteration in this scene. Maybe we should choose the task for a more general audience or we choose participants with higher skills in respect to the task (if this is possible). Of course are participants of both In-VR-Q and Outside-the-VR-Q affected and we had with 36 participants an appropriate distribution. But nevertheless, these factors should be considered and reduced as good as possible in the future when doing such a project. Like discussed in the qualitative feedback section in the previous chapter the short time being in the task scenes could have a main influence on our results. Maybe significant difference only emerge in terms of longer task experiences. There might be a minimum time starting from this the difference get significant enough to have an influence on the presence. Further investigation has to be done regarding times being active in VR.

Is Answer equal to Answer?

The observation described in the previous chapter showed that participants of the In-VR-Q group had the feeling being more focused than usual. In contrast, the other participants showed signs of distraction. The question arose whether this may influence the answers as well. Since participants were split into In-VR-Q and Outside-the-VR-Q (between subjects) we have no clue how the same participant may be answered the same questions in the other condition in respect to the questionnaire appearance group. When participants feel more Measuring this issue seems not being easy. A within subject design with this settings may be virtually not doable. First, the participant have to spend the double amount of time for the study. Furthermore with the previous made thoughts about the task time this would be even more. But secondly, the within subject design itself could be a problem. The so-called "carryover effect" would influence obtained data.

One way would be like mentioned above eliminating all disturbing factors and provide a calm environment where not more than one additional person, the examiner, is with the participant.

Potential Influence of Experience in VR applications

In the discussion we made the assumption that breaks in VR applications may also have no significant influences on the VR experience. If this is the case, further research can be done to find out how much time the breaks can take without have a significant effect on presence. Also in the field of gaming design there is potential research work for implementing integrated questionnaires.

New Presence Questionnaire

We used the three questionnaires SUS, IPQ and PQ with the aim getting a complete analysis on the subject measure of presence since we found that presence is a complex construct and is defined differently among researchers. The evaluating of the data revealed that is possible to create a new presence questionnaire combining items of the used questionnaires.

A User Study Documents

- Consent Form
- Demographic Information Sheet
- NASA-TLX
- The other 3 Q's maybe to much?!

A.1 Consent Form

University of Stuttgart
Germany

Human Computer Interaction Group (MCI), VIS
Prof. Dr. Albrecht Schmidt

Consent Form

DESCRIPTION: You are invited to participate in **a research study** on **Evaluation of In-VR Questionnaires**

TIME INVOLVEMENT: Your participation will take approximately **40 minutes.**

DATA COLLECTION: At the beginning of the study we will record personal data like gender, age, occupation and gained experiences in VR so far. Additionally, for the primary aim of this study, we will collect the answers of the completed questionnaires. These data will only be used for study purposes and deleted afterwards immediately.

RISKS AND BENEFITS: No risk associated with this study. The collected data is securely stored. We do guarantee no data misuse and privacy is completely preserved. Your decision whether or not to participate in this study will not affect your grade in school. You can decide whether the recorded personal data can be published or not.

PARTICIPANT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your **participation is voluntary** and you have the **right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled**. **The alternative is not to participate.** You have the right to refuse to answer questions. The results of this research study may be presented at scientific or professional meetings or published in scientific journals. Your identity is not disclosed unless we directly inform and ask for your permission.

CONTACT INFORMATION: If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, contact following persons:
Pascal Knierim pascal.knierim@vis.uni-stuttgart.de
Nico Haas nicohaasni@gmx.de
Valentin Schwind valentin.schwind@vis.uni-stuttgart.de

IMAGE DATA: (select one)

- ☐ Please **do not publish** the image data recorded during my participation of study.
- ☐ I allow you to **publish** the image data recorded during my participation of study.
- ☐ I allow you to **publish** the **anonymous** image data recorded during my participation of study.

By signing this document, I confirm that I agree to the terms and conditions.

Name: ________________________ *Signature, Date:* ________________________

Figure A.1: Consent Form

A.2 Demographic Information Form

Demographic Information Form

***Participant number*:**

Instructions: Please provide a response for each of the following questions:

1. What is your age? __________

2. What is you sex?

Female ○ Male ○

3. Wearing glasses?

Yes ○ No ○

4. Occupation: ____

5. Do you have any experiences with VR:

No ○ Yes ○ ,____

Figure A.2: Demographic Information Form

A.3 NASA Task Load Index

NASA Task Load Index

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

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure A.3: NASA Task Load Index

B Demographic Data of Participants

See next page.

ID	Age	Sex	Occupation	Glasses	VR-Exp.
1	24	male	Student, Business Informatics	no	no
2	23	male	Student, Computer Science	yes	yes
3	24	male	Student, Computer Science	yes	no
4	22	male	Student,	yes	no
5	21	male	Student, Software Engineering	no	yes
6	21	male	Student, Media Informatics	no	yes
7	21	male	Student, Computer Science	yes	yes
8	29	male	Student, Human Factors and Ergonomics	no	no
9	26	male	Student, Automotive and Motor Engineering	yes	no
10	26	male	Student, Computer Science	no	yes
11	23	male	Student, Media Informatics	no	yes
12	22	male	Student, Mechatronics	no	no
13	28	male	Student, Business Informatics	no	no
14	24	male	Student, Computer Science	no	yes
15	23	male	Student, Computer Science	no	yes
16	29	male	Student, Information Technology	yes	no
17	22	male	Student, Machine Engineering	yes	no
18	23	male	Student, Computer Science	no	yes
19	27	male	Student, Computer Science	no	yes
20	32	male	Student, Software Engineering	yes	no
21	23	male	Student, Software Engineering	No	Yes
22	26	male	Student, Computer Science	yes	yes
23	25	male	Student, Computer Science	no	yes
24	23	male	Student, Computer Science	yes	yes
25	20	female	Student, Social Science	no	yes
26	19	female	Student, Scoial Science	no	no
27	23	female	Student, Media Informatics	yes	yes
28	21	female	Student, Computer Science	yes	no
29	23	female	Student, Computer Science	yes	no
30	23	female	Student, Computer Science	no	yes
31	23	female	Student, Software Engineering	no	no
32	22	female	Student, Engineering	no	yes
33	22	female	Student, Computer Science	no	yes
34	27	female	PhD Student, Computer Science ?	no	yes
35	26	female	PhD Student, Mathematics	no	no
36	22	female	Student, Biomedical Engineering	no	no

Table B.1: Demographic data of participants

Bibliography

[62] *Sensorama simulator*. US Patent 3,050,870. Aug. 1962.

[BBBL01] J. N. Bailenson, J. Blascovich, A. C. Beall, J. M. Loomis. “Equilibrium theory revisited: Mutual gaze and personal space in virtual environments.” In: *Presence: Teleoperators and virtual environments* 10.6 (2001), pp. 583–598.

[BHG01] F. Biocca, C. Harms, J. Gregg. “The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity.” In: *4th annual international workshop on presence, Philadelphia, PA*. 2001, pp. 1–9.

[BHSS00] C. Basdogan, C.-H. Ho, M. A. Srinivasan, M. Slater. “An experimental study on the role of touch in shared virtual environments.” In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 7.4 (2000), pp. 443–460.

[BW93] W. Barfield, S. Weghorst. “The sense of presence within virtual environments: A conceptual framework.” In: *Advances in Human Factors Ergonomics* 19 (1993), pp. 699–699.

[DCJ+07] J. Difede, J. Cukor, N. Jayasinghe, I. Patt, S. Jedel, L. Spielman, C. Giosan, H. G. Hoffman. “Virtual reality exposure therapy for the treatment of posttraumatic stress disorder following September 11, 2001.” In: *Journal of Clinical Psychiatry* 68.11 (2007), p. 1639.

[DWH+99] H. Q. Dinh, N. Walker, L. F. Hodges, C. Song, A. Kobayashi. “Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments.” In: *Virtual Reality, 1999. Proceedings., IEEE*. IEEE. 1999, pp. 222–228.

[GMH01] M. Gerhard, D. J. Moore, D. J. Hobbs. “Continuous presence in collaborative virtual environments: Towards a hybrid avatar-agent model for user representation.” In: *International Workshop on Intelligent Virtual Agents*. Springer. 2001, pp. 137–155.

[GW95] R. H. Gilkey, J. M. Weisenberger. “The sense of presence for the suddenly deafened adult: Implications for virtual environments.” In: *Presence: Teleoperators & Virtual Environments* 4.4 (1995), pp. 357–363.

[GZ97] C. N. Gunawardena, F. J. Zittle. “Social presence as a predictor of satisfaction within a computer-mediated conferencing environment.” In: *American journal of distance education* 11.3 (1997), pp. 8–26.

[HB96] C. Hendrix, W. Barfield. “Presence within virtual environments as a function of visual display parameters.” In: *Presence: Teleoperators & Virtual Environments* 5.3 (1996), pp. 274–289.

[Hee92] C. Heeter. “Being there: The subjective experience of presence.” In: *Presence: Teleoperators & Virtual Environments* 1.2 (1992), pp. 262–271.

[ID98] W. IJsselsteijn, H. De Ridder. “Measuring temporal variations in presence.” In: *Presence in Shared Virtual Environments Workshop, University College, London*. 1998, pp. 10–11.

[IJs04] W. IJsselstein. “Presence in depth.” PhD thesis. PhD thesis, Technische Universiteit Eindhoven, 2004.

[IRH+98] W. IJsselsteijn, H. de Ridder, R. Hamberg, D. Bouwhuis, J. Freeman. “Perceived depth and the feeling of presence in 3DTV.” In: *Displays* 18.4 (1998), pp. 207–214.

[KB97] T. Kim, F. Biocca. “Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion.” In: *Journal of Computer-Mediated Communication* 3.2 (1997), pp. 0–0.

[LD97] M. Lombard, T. Ditton. “At the heart of it all: The concept of presence.” In: *Journal of Computer-Mediated Communication* 3.2 (1997), pp. 0–0.

[LVK01] P. Larsson, D. Västfjäll, M. Kleiner. “The actor-observer effect in virtual reality presentations.” In: *CyberPsychology & Behavior* 4.2 (2001), pp. 239–246.

[MAT00] C. D. Murray, P. Arnold, B. Thornton. “Presence accompanying induced hearing loss: Implications for immersive virtual environments.” In: *Presence: Teleoperators and Virtual Environments* 9.2 (2000), pp. 137–148.

[Min80] M. Minsky. “Telepresence.” In: (1980).

[MIWB01] M. Meehan, B. Insko, M. Whitton, F. P. Brooks. “Physiological measures of presence in virtual environments.” In: *Proceedings of 4th International Workshop on Presence*. 2001, pp. 21–23.

[NHW00] S. Nichols, C. Haldane, J. R. Wilson. “Measurement of presence and its consequences in virtual environments.” In: *International Journal of Human-Computer Studies* 52.3 (2000), pp. 471–491.

[Par98] A. Parent. “A virtual environment task analysis workbook for the creation and evaluation of virtual art exhibits.” In: *NATIONAL RESEARCH COUNCIL OF CANADA-REPORTS-ERB* (1998).

[PR08] T. D. Parsons, A. A. Rizzo. “Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis.” In: *Journal of behavior therapy and experimental psychiatry* 39.3 (2008), pp. 250–261.

[RHS+00] B. O. Rothbaum, L. Hodges, S. Smith, J. H. Lee, L. Price. “A controlled study of virtual reality exposure therapy for the fear of flying.” In: *Journal of consulting and Clinical Psychology* 68.6 (2000), pp. 1020–1026.

[SFR01] T. Schubert, F. Friedmann, H. Regenbrecht. “The experience of presence: Factor analytic insights.” In: *Presence: Teleoperators and virtual environments* 10.3 (2001), pp. 266–281.

[She92] T. B. Sheridan. “Musings on telepresence and virtual presence.” In: *Presence: Teleoperators & Virtual Environments* 1.1 (1992), pp. 120–126.

[SKT+17] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, N. Henze. “These are not my hands!: Effect of Gender on the Perception of Avatar Hands in Virtual Reality.” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM. 2017, pp. 1577–1582.

[Sla99] M. Slater. “Measuring presence: A response to the Witmer and Singer presence questionnaire.” In: *Presence: Teleoperators and Virtual Environments* 8.5 (1999), pp. 560–565.

[SUS94] M. Slater, M. Usoh, A. Steed. “Depth of presence in virtual environments.” In: *Presence: Teleoperators & Virtual Environments* 3.2 (1994), pp. 130–144.

[SUS95] M. Slater, M. Usoh, A. Steed. “Taking steps: the influence of a walking technique on presence in virtual reality.” In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 2.3 (1995), pp. 201–219.

[Sut68] I. E. Sutherland. “A head-mounted three dimensional display.” In: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. ACM. 1968, pp. 757–764.

[SW97] M. Slater, S. Wilbur. “A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments.” In: *Presence: Teleoperators and virtual environments* 6.6 (1997), pp. 603–616.

[TMTC03] P. Turner, I. McGregor, S. Turner, F. Carroll. “Evaluating soundscapes as a means of creating a sense of place.” In: Georgia Institute of Technology. 2003.

[TS12] A. Toet, M. G. van Schaik. “Effects of signals of disorder on fear of crime in real and virtual environments.” In: *Journal of Environmental Psychology* 32.3 (2012), pp. 260–276.

[TTC+03] S. Turner, P. Turner, F. Carroll, S. O’Neill, D. Benyon, R. McCall, M. Smyth. “Re-creating the Botanics: towards a sense of place in virtual environments.” In: *The 3rd UK Environmental Psychology Conference*. 2003.

[UAW+99] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, F. P. Brooks Jr. “Walking> walking-in-place> flying, in virtual environments.” In: *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co. 1999, pp. 359–364.

[WS98] B. G. Witmer, M. J. Singer. “Measuring presence in virtual environments: A presence questionnaire.” In: *Presence: Teleoperators and virtual environments* 7.3 (1998), pp. 225–240.

All links were last followed on August 13, 2017.

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

place, date, signature