

The Usage of Presence Measurements in Research: A Review

Dimitri Hein, Christian Mai and Heinrich Hußmann

Ludwig Maximilian University of Munich

Abstract

The research on presence in recent years resulted in a number of definitions and methods to measure it. Due to the variety of concepts it is difficult for today's and future researchers to compare and interpret presence ratings of individual studies. We conducted a literature review in order to gain insight into the usage of presence measurements in studies using head-mounted displays in the years 2016 and 2017. We show that 93% of the 41 reviewed studies rely on eight presence questionnaires with the Witmer and Singer Presence Questionnaire being utilized the most. The usage of behavioral or physiological measures is very rare. With our work we aim to foster the discussion about guidelines of presence measurements that help practitioners and researchers to evaluate their work in a sustainable manner.

Keywords: virtual reality, presence, head-mounted displays, measurements, literature review

The Usage of Presence Measurements in Research: A Review

Introduction

The introduction of consumer grade VR headsets such as the *HTC Vive*¹ or *Samsung Gear VR*² or *Google Daydream*³ lead to a growing number of users in industrial, commercial, research and private contexts. For example the steam app store⁴ today offers 2.600 commercial VR experiences. Also the research interest shows intensive growth. A search for the exact terms "*head mounted display*" and "*head-mounted display*" in the ACM digital library shows 279 research items for the years 2010 to 2015 and 340 research items for the years 2016 and 2017.

In order to continuously improve VR technology and experiences, a constant evaluation and comparison is needed. A key measurement for VR systems is the degree to which a user is enabled by the system to feel present in the virtual world. The concept of being present in the VR is defined in a number of different ways (e.g. (Lombard & Ditton, 1997; McMahan, 2003; Slater & Wilbur, 1997a; Witmer & Singer, 1998)). We define the feeling of presence as the outcome of a cognitive process, which is different to the term "immersion", describing the system's properties. Both terms will be defined in more detail in the following sections. Important to us is the question of how this measured variable can be quantified. In general, methods can be divided into (1) physiological measurements, (2) behavioral observations, (3) questionnaires and (4) interviews. Each of the categories covers a number of methods to measure the term presence. The work of Baren and IJsselsteijn (van Baren & IJsselsteijn, 2004) summarized the existing methods in the year 2004. Their literature research revealed 29 different questionnaires that address the measure of mainly spatial, but also social presence.

Several challenges arise due to the number of different methods. The comparability between different research methods is difficult as (1) all measurements

¹<https://www.vive.com>

²<http://www.samsung.com/global/galaxy/gear-vr>

³<https://vr.google.com/daydream>

⁴<http://store.steampowered.com>

have a different underlying concept – e.g. measuring a mental construct vs. asking for system properties – and (2) construction of the measurement tool – e.g. different step width in the Likert scales – and (3) the resulting measured variable will be different in absolute numbers. We argue that by using inconsistent measurement methods between studies, the results of studies might only give limited insight into the increase of presence, since researchers might measure different aspects of presence, entirely different qualities of a VR experience or simply cannot properly compare their findings. Some of these problems become visible in the meta-analysis of Cummings and Bailenson (Cummings & Bailenson, 2015) on the influence of immersive attributes towards the feeling of being spatially present in VR. Another problem might arise in the near future. With the growing distribution of head-mounted displays (HMD) in public spaces the need for regulations, or the system's design, might arise. One question might be how immersive, or how strong, the feeling of presence is allowed to be in public spaces, such as trains. They may also be product liability cases concerning psychological concerns due to the long term use of HMDs that we do not understand yet. A single measurement of an effect mostly gives limited statistical evidence. To find the truth, a meta analysis of data from multiple studies would be the optimal method. However the available data needs to be in a form that allows researchers to conduct a meta analysis. As indicated by related work (Cummings & Bailenson, 2015; Rosakranse & Oh, 2014), we argue that this is not the case today.

Therefore we want to gain insight into the usage of measures used in present research, accelerated by the introduction of favorable consumer grade HMDs. The goal is to show how presence is measured in recent research, give an overview of common measurement methods and discuss the findings of the review on the current state of presence research.

After methodically reviewing 41 papers published in 2016 and 2017, the years consumer grade HMDs were introduced, several interesting findings were discovered that indicate problems which were already partly pointed out by other researchers (Cummings & Bailenson, 2015; Rosakranse & Oh, 2014; Sanchez-Vives &

Slater, 2005; Slater, 2004). Those problems include the prevailing reliance on presence questionnaires instead of behavioral measures in addition to the lack of data collection that can give more context to the presence ratings in most studies, as well as the variety of used measurements which reduce comparability. It was found that the Witmer and Singer Presence Questionnaire (PQ) (Witmer & Singer, 1998) was still the most commonly used questionnaire. Although different measurement methods cover a different aspect of the feeling of presence, we could not identify a clear connection between measurement method and the addressed research question.

Related Work

In presence research many different definitions of the terms presence and immersion have been proposed, resulting in likewise diverse measurement methods. For instance, Schuemie et al. (Schuemie, van der Straaten, Krijn, & van der Mast, 2001) summarized these definitions and highlighted the importance to distinguish between the terms presence and immersion.

The quite common interchangeability of those two terms was later pointed out by McMahan (McMahan, 2003) for the context of virtual reality and video games. A further implication made in that work is that the distinction of *presence* and *immersion*, using well-defined criteria, can help game developers to judge the "immersiveness" and degree of presence. This can also be applied to the improvement of perceived presence in VR. McMahan also uses Steuer's (Steuer, 2000) definition of *Telepresence* as a foundation for presence.

In their compendium Baren and IJsselsteijn (van Baren & IJsselsteijn, 2004) summarize approaches for measuring presence, including subjective methods, such as questionnaires and objective approaches like behavior measures. Not only did they summarize the methods, but they also included relevant critique by other researchers such as the failed "reality check" by Usoh et al. (Usoh, Catena, Arman, & Slater, 2000) for the PQ by Witmer and Singer (Witmer & Singer, 1998).

In a more recent review of existing definitions of immersion by Nilsson, Nordahl

and Serafin (Nilsson, Nordahl, & Serafin, 2016) several partly opposing views on immersion are presented and categorized. Their work summarizes and discusses the definitions and differences of the term presence over the last decades. They do not, however, give insights into the dissemination of the measurement methods in the body of research.

Rosakranse and Youn Oh conducted a literature review in order to find out about the development in usage of presence questionnaires between 1998 and 2012 (Rosakranse & Oh, 2014). They show that the popularity of the Witmer and Singer PQ reduced after the introduction of the Igroup Presence Questionnaire (IPQ) (Schubert, Friedmann, & Regenbrecht, 2001) and the ITC-Sense of Presence Inventory (Lessiter, Freeman, Keogh, & Davidoff, 2001), but PQ and Slater-Usoh-Steed Presence questionnaire still were dominant. Unlike our work they focused on venues specialized on the research on presence.

The meta-analysis done by Cummings and Bailenson (Cummings & Bailenson, 2015), with the goal to investigate the impact of technological immersion on presence, also included the review of presence measurements. The criteria for the selection of studies is focused on self-reported presence and, in contrast to our work, older studies from 1995 to 2014 which includes display system as well as CAVE like systems. In their work they point out that the most promising measures, like behavior, cognitive and physiological measures, are too disparate to meet the criteria of a meta-analysis. Similar issues arise in the meta-analysis of social presence and self-presence. For their meta-analysis of spatial presence, with the use of self-reported measurements, they had to limit their procedure to a very specific decision tree. Cummings and Bailenson focus on the implications of their analysis, such as the fact that certain features of technological immersion have a greater impact on presence, rather than reviewing the utilized presence questionnaires and their comparability.

Conceptualizing Presence and Immersion

Research on presence spans for decades and resulted in different measurement approaches (e.g. (Schubert et al., 2001; Slater & Wilbur, 1997b; Witmer & Singer, 1998)). Arguably one of the most common phrases used in those works is more or less a quote of Minsky's description of *telepresence* which includes the sense of "being there" (Minsky, 1980). Even though Minsky describes telepresence in the context of a remotely controlled machine, it can still be applied to presence in a virtual environment (VE) in the context of VR. In this chapter we discuss some of the common interpretations of the presence term in order to give an insight on the variety and complexity of possible definitions that make it difficult to compare findings between research referring to different concepts.

Following Slater and Wilbur's work from 1997, presence can be defined as a person's psychological state which manifests itself in two phenomena: (1) The person evaluates the VE as "place-like" and feels some degree of "being there". (2) The person behaves in the VE in a similar manner as in the everyday reality (given analogous circumstances) (Slater & Wilbur, 1997b, p. 606). It is important to note that those two phenomena differ in terms of how one can measure them, the first one being a subjective evaluation reported by the person and the latter one being objectively observable.

The Components of Presence in VR

A somewhat complementary continuation of the above stated definition is made by Schubert, Friedmann and Regenbrecht (Schubert et al., 2001), as they describe presence as an outcome of the user's construction of a special type of mental model of the virtual space. Additionally an important distinction between interactions with a display and a presence inducing VE is stated. While a user might be focused on a certain task in a VE using a display, he/ she might not pay attention to their position in the virtual space. Presence emerges only when the user's actions, which are represented mentally, are depicted as bodily actions within the virtual space, being functionally related to navigation, manipulation of objects or even interactions with

other agents in the space (Schubert et al., 2001, p. 268). This resembles Slater and Wilbur's hypothesis that presence requires a virtual body (Slater & Wilbur, 1997b, p. 606). In order for presence to emerge, the experiencing person does not only need to have a sense of being located in the VE, but also being able to act in it.

Another crucial, yet not entirely unique, component of presence is directly connected to the person's attention. In other words it should describe the degree to which the person is concentrating on the VE and ignoring the real environment (Schubert et al., 2001, p. 269).

Through a factor analysis study Schubert, Friedmann and Regenbrecht were able to further explore the individual components within their construct of presence (Schubert et al., 2001). By using items from different questionnaires of several researchers in a factor analysis, it was found that by using their measurement method, presence consists of three main components which combine subjective interpretations of the participant's experiences in the VEs:

- *Spatial presence*: often described as the sense of being in the VE or feeling as if one is located in the VE
- *Involvement*: the degree of attention (and awareness) to either real or virtual environment
- *Realness*: addresses the person's judgment of the realness of the VE in comparison to the known actual reality

Not only did Schubert and colleagues' studies confirm that presence is a multidimensional construct, but in particular that the items that were extracted from other questionnaires do, in fact, split into the distinct factor groups presence, immersion and interaction (Schubert et al., 2001). The importance of distinguishing between, for instance, presence and immersion factors is explained in the section on immersion below.

These factors describing presence, especially the spatial presence in VEs, are defined similarly in related work researching on presence, e.g. the cross-media presence questionnaire by Lessiter et al. (Lessiter et al., 2001).

Presence Types

As already stated, presence can be defined as a multidimensional construct which is not only a description of the spatial feeling of being present in the VE. Following the review by Lombard and Ditton (Lombard & Ditton, 1997) where presence is identified using two broad categories, it is useful to distinguish three main types of presence (Riva, Davide, & IJsselsteijn, 2003):

- Physical presence: roughly reflects the component *spatial presence*, the sense of being physically located in the mediated space
- Social presence: describes the feeling of being together, social interaction with a virtual or remotely located communication partner
- Co-Presence: combination of physical and social presence, where a sense of being together in a shared space is created

It is therefore important to also consider social and co-presence in cases where some kind of interaction exists with other autonomous agents in a VE. A study by Sanchez-Vives and Slater showed how Co-Presence in a VE can influence not only task performance but can also result in reactions that resemble everyday life behavior (Sanchez-Vives & Slater, 2005). In our work we mainly focus on the measuring of the "being there" feeling as measured by Schubert and colleagues (Schubert et al., 2001). But we also include studies that measure these other forms of presence additionally.

Immersion as an Objective and Technological Description

In this paper the term *immersion* was purposely avoided in the context of conceptualizing presence so far, due to the varying interpretation of the words. We follow Slater and Wilbur's distinction of presence and immersion (Slater & Wilbur, 1997b). This means that immersion, in contrast to presence, is objective and quantifiable, since it describes what a particular system provides. They also define four

main characteristics of an immersive system: inclusion, extensiveness, surrounding and vividness. Due to the differing usage of the term immersion it is also sometimes called *system immersion*, emphasizing its affiliation to the system characteristics. For simplicity's sake in this paper, it is simply called immersion.

By defining immersion in such a way, presence can be identified as a psychological response to immersion. Slater and Wilbur name several cases where this distinction is helpful (Slater & Wilbur, 1997b, p. 609). By being able to investigate both presence and immersion separately, it is possible to find characteristics of an immersive system that enhances presence. Good examples for attributes of an immersive system are, for instance, the latency and resolution of a VR Head-Mounted display or of other sensory modalities such as spatialized sound.

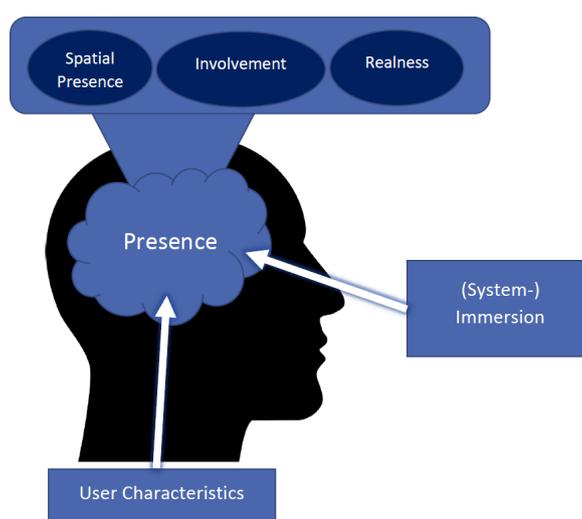


Figure 1. Presence as a psychological response to immersion influenced by the person's characteristics based on propositions by Schubert, Friedmann, Regenbrecht (Schubert, Friedmann, & Regenbrecht, 2001) and Slater (Slater, 1999).

Another reason why Slater recommends the differentiation between the quantifiable technological characteristics of a system and the resulting psychological state of the user can be found in his critique of the presence questionnaire published by Witmer and Singer (Witmer & Singer, 1998). In the response to Witmer and Singer's questionnaire it is argued that the answers to the questions are generally subjective opinions of the participants (Slater, 1999). While Witmer and Singer agree that the

degree of experienced presence depends on both the cognitive processes of the individual and the characteristics of the system, it is pointed out that their questionnaire makes it impossible to separate both influences. Since it is very likely that differences in user characteristics such as traits, experiences or even biological varieties influence presence (Riva et al., 2003), it is important that one is able to measure them independently from the system factors. By considering this necessity it is, for instance, possible to study the influence of system factors on presence with people who share similar traits (Slater, 1999). A simple illustration of a possible framework for presence based on the previously elaborated definitions is shown in figure 1.

Although Schubert, Friedmann and Regenbrecht's results from a factor analysis of presence support Slater's critique, some items from Witmer and Singer's questionnaire were still used in their studies (Schubert et al., 2001). Some researchers also argue that while Witmer and Singer's questionnaire contains questions that might be interpreted by the participants as technology related, it is possible to fix the ambiguity by re-phrasing them carefully (Lessiter et al., 2001).

Measuring Presence

In most VR presence studies presence is measured by having participants carry out some sort of task in a VE (e.g. while wearing a virtual reality head-mounted display) and afterwards answer presence related questions. Even though using a questionnaire is quite common in presence research, one usually can use four different general methods for its measurement:

- Analysis of the user's behavior
- Physiological measures
- Questionnaires
- Interviews

Even though there are more specialized methods to measure presence such as the thinking aloud or task performance⁵ approaches, they can in some sense be allocated to one of the above listed groups.

For the behavioral approach the point of focus is on whether the participants behave in the VE as though they are under similar conditions in the physical environment. In order for this method to be applicable, the VE requires features that trigger bodily responses such as ducking in response to a flying object (Sanchez-Vives & Slater, 2005, p. 335).

Furthermore there is the usage of physiological measures, which is also somewhat connected to the user's behavior, like electrocardiogram recordings which can be used to estimate the users workload (Sanchez-Vives & Slater, 2005, p. 335). The reason why this approach is not as widely used is that the participant's psychological response has to be rather obvious, which is difficult to achieve in mundane situations.

In a study by Slater and Steed (Slater & Steed, 2000) *breaks in presence* (BIPs) are introduced. Instead of trying to improve factors that result in a higher degree of presence, events that cause the participant to break out of the state of presence are evoked. Participants would therefore report BIPs whenever there is a transition from the VE to the real physical world. Examples for the causes of BIP include the participant's collision with real world objects such as the lab equipment or highly unrealistic effects such as rendering trees as pixel maps rather than solid objects (Sanchez-Vives & Slater, 2005, p. 335).

Questionnaires

There are many different proposed questionnaires for measuring presence including "general-purpose" presence questionnaires, social-presence questionnaires and cross-media questionnaires (Greef & Ijsselsteijn, 2001; Lessiter et al., 2001; Slater, Usoh, & Steed, 1994). Using scales from 1 to 7, typical items in many presence questionnaires use questions such as "I had a sense of being there in place X".

⁵Slater and Wilbur argue that there is no evidence for a positive correlation between presence and task performance (Slater & Wilbur, 1997b)

Slater and Sanchez-Vives take issue with the heavy reliance on presence questionnaires in most studies and argue that the corresponding methods can be unstable (Sanchez-Vives & Slater, 2005, p. 335) and prone to methodological circularity (Slater, 2004). Mentioning presence in the items of the questionnaire might bring up the very phenomenon of presence and therefore distort results. Questionnaires that use such questions include, for instance, the Murray et al. questionnaire (Murray, Arnold, & Thornton, 2000) or the questionnaire on presence and realism by Parent (Parent, 1998).

Methodology: Review of Measurement Methods in Recent Research

In order to review the measurement methods utilized in recent presence research a sample of scientific papers was analyzed. Due to the variety of contexts of such studies (e.g. presence in VEs, VR or even cave systems) the selection of papers was based on multiple criteria. Firstly, only papers that were published either in 2016 or 2017 were collected as we want to include research from domains that were attracted to VR by the affordability of consumer grade HMDs. Additionally the paper had to include an actual user study where participants take part in some sort of VR experience using an HMD such as the *Oculus Rift* or mobile VR devices. Other setups such as presence in general VE studies using usual computer monitors or cave systems were not considered. Lastly the study had to utilize some sort of presence rating measurement such as questionnaires, observation of behavior or interviews.

The search for the relevant papers was done using the *ACM Digital Library* and *IEEEExplore Digital Library*. Keywords for the search were comprised of "virtual reality", "presence" and "immersion" and the years 2016 and 2017 were considered as they are expected to include research using consumer grade headsets. The specific timespan that was used for the selection is from January the 1st 2016 till November the 30th 2017.

By using the above mentioned search method, 41 scientific papers were found for the meta-analysis and then analyzed for the criteria *Used Measurement Method*, *Used Questionnaire*, *Reasons for Specific Method*. Even though VR as the context of the user

study was an important criteria, the secondary context (more specific use case) was not a criteria for the paper's inclusion in the meta-analysis. Thus, the selected papers vary in terms of what kind of applications were used in the study in addition to varying use cases. Examples include actual simulations, games and 360° story telling experiences. The most important factor of the analysis was the utilized measurement methods where the search was not limited to specific approaches or methods.

Results

As reported in the past, researchers rely on presence questionnaires (Sanchez-Vives & Slater, 2005, p. 335). The results from our review confirm this trend as 38 out of 41 papers used some kind of questionnaire for the presence measurement shown in figure 2. Three of the studies therefore solely used objective measures or interviews. Six out of 41 papers used both a presence questionnaire and at least one of either an interview or some objective measure. Objective measures were mostly recorded behavior of the participants (via video) and electrodermal activity. Moreover the thinking aloud method was used in the case of a VR game as well as the analysis of the user's behavior. Most papers gave no specific reason explaining why the specific method or questionnaire was used. In three instances researchers referred to the fact that the questionnaire of Witmer and Singer (Witmer & Singer, 1998) is widespread and commonly used.

One of the exceptions is the work by Kaul, Meier and Rohs in which they refer to Nacke and Lindley's experiments (Nacke & Lindley, 2008) that show significant differences in different experiment conditions when measuring psycho-physiological parameters for presence (Kaul, Meier, & Rohs, 2017). Furthermore, they point out that those differences could not be clearly correlated to accompanying questionnaires and state that they decided to use the Witmer and Singer PQ, which they describe as a widely used qualitative method.

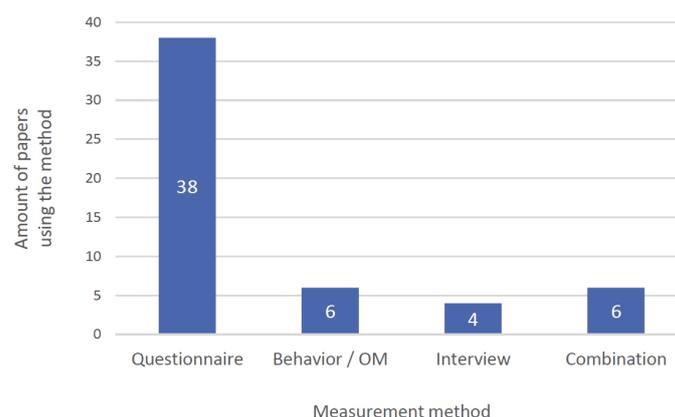


Figure 2. Utilized presence measurement methods in 41 presence studies from the years 2016 and 2017. The *Combination* category includes papers where both a questionnaire and at least one of the other two methods is used. *OM* stands for *Objective measure*.

Results on the Usage of Questionnaires

In most publications a specific questionnaire is used. However there are cases in which the questionnaires are altered in some way or individual items are removed. In 18 studies at least one questionnaire was used in its standard form (without adding or removing items), while in 13 papers it is stated that at least one of the used (published) questionnaires was either adapted in a specific way, had items removed or added. The stated reasons for the alteration varied from the lack of the necessity of sub factors (for the removal of items), missing relevant factors (for the addition of items) or even no stated reason whatsoever. Due to this problem of questionnaire alteration a questionnaire is marked as utilized if it is directly stated that a significant portion of it is used or when significant parts of the corresponding questionnaire are extracted. In figure 3 the distribution of the used questionnaires is shown. The Witmer and Singer PQ (Witmer & Singer, 1998) is utilized in 13 out of 41 studies. The IPQ (Schubert et al., 2001) was used in seven studies and the Slater-Usuh-Steed PQ (Usuh et al., 2000) in four studies. Eight other published questionnaires are used in total besides these three questionnaires. In three cases they were used in addition to at least one of the three most frequent ones. In those cases either two or more questionnaires are used separately or are combined into one questionnaire, while using portions of the individual

questionnaires which the researchers seemed to prioritize. In some works the Witmer and Singer PQ was modified due to missing aspects in the experiment. For instance, in the experiments of vibrotactile feedback in VR by Lee, Bruder and Welch (Lee, Bruder, & Welch, 2017), items that include aspects of 3D navigation were removed from the used questionnaire. Other questionnaires that were utilized were, for instance, the General Presence Question by Barfield et al. (Barfield & Weghorst, 1993) and social presence questionnaires like the Bailensen et al. questionnaire (Bailenson, Rex, Beall, & Loomis, 2001). The *Custom* category includes all papers in which it was either not specified what kind of questionnaire was used or which stated that a custom questionnaire with self-defined item-sets was designed.

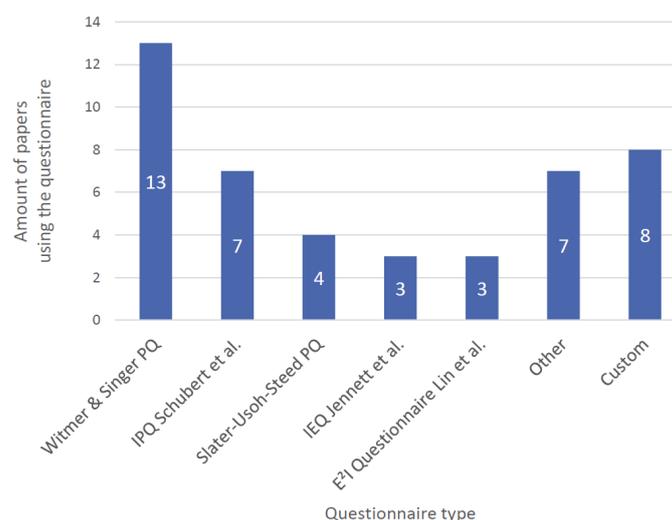


Figure 3. Utilized presence questionnaires in 41 reviewed VR research papers from the years 2016 and 2017. The *Other* category contains published questionnaires that are only used in one or two papers. The *Custom* category includes specially designed questionnaires.

Besides the presence questionnaire, Witmer and Singer published a corresponding immersive tendencies questionnaire (ITQ) (Witmer & Singer, 1998) which is designed to be used in combination with the PQ. The ITQ was used in one of the 13 studies with the Witmer and Singer PQ. Additionally the Witmer and Singer PQ is the most used questionnaire in cases where only a single questionnaire was used for the measurement (seven out of 13 studies). Both the Regenbrecht & Schubert and Slater-Usuh-Steed PQ

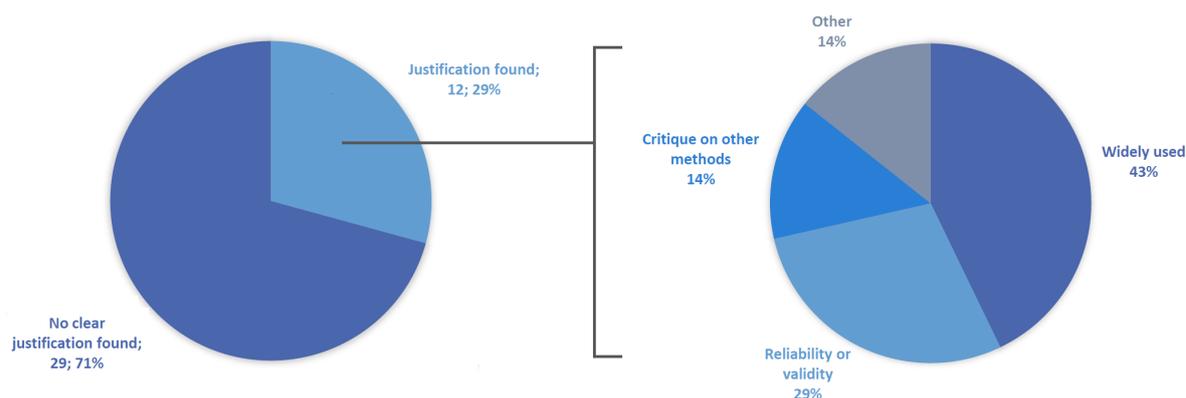


Figure 4. Justifications for the usage of specific measurement methods. In 29 out of 41 papers no clear justification could be found.

were only used in one study without the addition of other questionnaires.

In most papers it was not clear why one of the less popular questionnaires was used. Although one possible reason for their use might be the addition of other aspects such as the inclusion of enjoyment factors in the case of the E²I questionnaire by Lin et al. (Lin, Duh, Parker, Abi-Rached, & Furness, 2002). This was, for instance, used in the airflow simulation experiments by Rietzler et al. (Rietzler et al., 2017). In table 1 the individual questionnaires and the frequency of occurrences is shown, including the less used ones, in order to give a more detailed overview.

Method Selection and Further Insights

29 out of 41 papers do not explain the selection of the used method to measure presence. In figure 4 the reported reasons for selecting a method and their frequency is presented. Not all of the 12 cases are direct justifications per se, but can also be remarks from the authors that can be interpreted as one of the reasons explaining why the specific method was used. For instance, remarks that prominently mention the frequent use of Witmer and Singers PQ in many previous studies are also counted as a reason. The most common reason is referring to the wide usage of the method, followed by stating that the utilized method was shown to be reliable or to some degree proven to be valid. Justifying the use of a method by pointing out flaws or problems of other

Table 1

List of utilized (and already published) presence questionnaires in 41 reviewed VR research papers from the years 2016 and 2017, including their individual usage frequency.

Questionnaire	Usage Frequency
Witmer & Singer PQ	13
Witmer & Singer ITQ	3
Regenbrecht & Schubert IPQ	7
Slater-Usoh-Steed PQ	4
IEQ Jennett et al.	3
E ² I Questionnaire Lin et al.	3
Lessiter et al. ITC-SoPI	2
MEC Spatial Presence Questionnaire	1
Semantic Differential Technique - Social PQ	1
UCL Presence Questionnaire (SMoP)	1
RJPQ Baños	1
Barfield et al. Questionnaire	1

methods such as using behavior measures instead of questionnaires only made up 14% of the cases. Other justifications include cases where a specific method was used based on the attributes of the method, e.g. the encompassing nature (Newbutt et al., 2016) of the ITC-SoPI (Lessiter et al., 2001).

In addition to the presence measurements, eight out of 41 papers measured simulator sickness. Other measured factors include usability, body ownership and workload.

Discussion

About 93% of the papers used a presence questionnaire as the main measurement method or at least as one of multiple utilized methods. This number shows that questionnaires are still the first choice for researchers when it comes to measuring presence in VR applications. Even in cases where interviews were conducted, questions mostly addressed factors that might support the data collected by using questionnaires or give additional context to the questionnaire results. In only one out of four papers,

using the interview method, an interview was conducted without using a questionnaire and even in that case the presence factor resembled typical items from presence questionnaires identified as "Subjective experience of being in one place or environment, even when one is physically situated in another" (Witmer & Singer, 1998). Although a questionnaire is easy to use and gives detailed insights on the presence feeling, Slater and Sanchez-Vives (Sanchez-Vives & Slater, 2005, p. 335) say that the main issue in using questionnaires to measure presence is that prior information, experience and task expectations can affect the participant's presence rating (Freeman, Avons, Pearson, & IJsselsteijn, 1999). Therefore, data collected via questionnaires today might turn out to be of little help for addressing future challenges in the upcoming/next ten years, as users get used to the technology.

Questionnaire Discussion

As questionnaires are the most common method to assess the feeling of presence in virtual reality, the following section discusses this in more detail. *Variety of questionnaires*—As expected, a variety of questionnaires is currently present in research: eight self-made questionnaires were used in addition to 11 different published presence questionnaires listed in table 1. One might argue that the used questionnaires, at least in principal, measure the same aspects of presence. Even though this assumption is already problematic due to intrinsic differences in the creators' conceptualizations of presence, demonstrated by Slater's critique on Witmer and Singers PQ (Slater, 1999), there are also other factors that make comparisons of presence ratings difficult. The most obvious example is that the used questionnaires themselves differ in terms of scope and, inevitably, the level of detail of the different factors which are measured. Additionally, there are often differences in measured factors concerning both the scales used and how they are defined. For instance, when one would want to compare findings of two studies, one based on the IGroup Presence Questionnaire (IPQ) (Schubert et al., 2001) and the other on the ITC-SoPI (Lessiter et al., 2001), it is crucial that their differences are considered. The IPQ is identified by three presence factors, contains 14

items and is based on the idea of presence developing from the construction of a spatial-functional mental model of a VE. The ITC-SoPI, on the other hand, is identified by four presence factors, contains 44 items and was developed with the intention to create a valid cross-media questionnaire. Looking at these variations, it is rather obvious that the results of the corresponding studies can only be properly compared if the differences are taken into account. This problem is outlined in a recent analysis by Nilsson, Nordahl and Sefarin (Nilsson et al., 2016) including the characterizations of presence by Slater and Witmer & Singer, where the authors state that it is important that researchers are mindful when it comes to interpreting and comparing the results of other studies. To what extent this proposal is considered in the current state of presence research would go beyond the scope of this discussion, but it is important to point out nonetheless. Since, in many cases, given questionnaires are also modified or combined differently in the respective study, it might be even harder to come back to that work in a timely distance of 10 years. In terms of possible correlations between used questionnaire and VR context, the Witmer and Singer PQ was used in a wide variety of studies including general presence, simulations, social-presence, games and storytelling. These results therefore do not indicate any correlation between the usage of the Witmer and Singer PQ and a specific use case. The same applies to the IPQ Regenbrecht & Schubert and the Slater-Usoh-Steed PQ. We do not analyze the other questionnaires in that regard as their usage frequency is too low in the years 2016 and 2017.

The Witmer and Singer PQ– The Witmer and Singer PQ is still the most popular questionnaire in use. Also its construction is critically discussed by Slater (Slater, 1999) and also partially by Schubert, Friedmann and Regenbrecht (Schubert et al., 2001). Interestingly it is only used in three cases in combination with the immersive tendencies questionnaire (ITQ) which is designed to give more insights into the results of the Witmer and Singer PQ by measuring the participants' tendencies to immerse themselves. Following the criticism of Slater (Slater, 1999) it is also interesting that the Witmer and Singer PQ and the Slater-Usoh-Steed PQ are both used together in two studies. In these cases they were not combined, but rather used separately in order to

compare their individual results (Salanitri, Lawson, & Waterfield, 2016; Skarbez, Brooks, & Whitton, 2017). The use of both the Witmer and Singer PQ and the Slater-Usuh-Steed PQ are not that surprising, since both were developed by well known researchers in the field of presence research. The use of the E²I Questionnaire by Lin et al. (Lin et al., 2002) seems to be justified as in addition to presence, it also measures enjoyment which was part of the goals of the studies, e.g. in the context of a VR exploration game (Frommel, Sonntag, & Weber, 2017). The same applies to studies with a more specialized use case, such as the specific measurement of spatial presence with the MEC Spatial Presence Questionnaire (Vorderer et al., 2004) or an additional social presence factor with the Social Presence Questionnaire by Short, Williams & Christie (Short, Williams, & Christie, 1976). In the section *Results on the Usage of Questionnaires* the variety of other less common questionnaires in the reviewed studies is shown, but drawing other valid conclusions besides the lack of their utilizations in recent presence research is difficult, due to their infrequent use.

In terms of custom questionnaires that were designed specifically for the individual studies, the items are included in the corresponding papers except for one case. The need for their disclosure is self-evident and enables comparisons to other questionnaires. Some of the questionnaires measure aspects that resemble presence factors such as spatial presence or attention, but are on the other hand too specific to be compared in a simple manner, e.g. the evaluation of collision avoidance effects on discomfort in VEs (Sohre, Mackin, Interrante, & Guy, 2017).

Behavior and Interview Discussion

Since only four papers were found which used an interview to measure presence, it is difficult to draw a valid conclusion concerning the most common practices. All found cases differ in at least one of the important attributes such as scope, measured factors or whether they are combined with other methods such as questionnaires. Therefore, when it comes to comparing the results the same problems occur as in the case of questionnaires.

Looking at the measurement of behavior, the results indicate that such methods are rather unpopular due to the fact that only six out of 41 papers stated their utilization. The specific cases vary from verbal reactions, non-verbal reactions, electrodermal activity, analogous behavior compared to real world and the thinking aloud method. In only two out of the six works solely behavioral or objective measures is used without the utilization of questionnaires and interviews. This could have multiple reasons, most likely being the possibility that researchers see those methods as less reliable, valid or simply as a secondary source of data. Besides the more common behavior measurement methods such as real life analogous behavior, no other ("exotic") proposed methods such as Breaks in Presence (BIP) (Slater & Steed, 2000) were found.

Other Relevant Data

In the section on immersion the idea of presence as a response to immersion influenced by the person's characteristics is elaborated and is reflected in some of the papers' methods. The most prominent example is the measurement of simulator sickness or cyber-sickness which was found in eight papers. Cyber-sickness can play an important role as people who report higher symptoms tend to also report less presence when using a presence questionnaire (Witmer & Singer, 1998). Other factors influencing the presence ratings, like personality traits or previous experience with the used technology were only found in a few cases. Such additional information, however, could help future researches to go back to today's data and estimate the influence of the personality traits on the results from the presence measurements. This could prevent the discussed incomparability of future results or meta analysis to today's data.

In order to get a better insight into how and why researchers use different methods we looked for reasons as well as justifications for using certain methods. These observations from the field might help in a future standardization process, as they might tell us more about usability issues of today's methods. In the majority of the studies no clear reasoning for using a particular method could be found. Five papers stated that their methods were widely used and although identified as a kind of

justification here, arguably it can be seen as a remark without the authors' intention to use it as a justification. Nonetheless there might be the possibility that the high usage frequency of a questionnaire influences the authors' decision to pick the corresponding method. However, some authors stated some justifications concerning special attributes of questionnaires or even criticized the use of presence questionnaires in general. To this point we found a variety of methods and do not see any trend towards a certain type of measurement in the near future's research.

Conclusion

This work contributes a review of the most recent VR presence studies in terms of utilized measurement methods. Combined with the most common definitions and propositions regarding presence in VEs, the results of the review give an overview of the current state of presence measurements. Especially in the case of HMDs, which will likely become a larger user group in the next years, presence measurements might have a big impact in both industrial and research use cases. Therefore it is crucial to also review the current state of research in a broader scope in order to help remove prevailing problems by giving other researchers not only an overview but also point out topics for improvement. Even though there is work that focuses on comparing presence measurement methods and their underlying conceptualizations of presence, actually reviewing presence studies in terms of the used methods is not as common. Several interesting findings were pointed out including the prevailing reliance on presence questionnaires, the diversity of used questionnaires and the unpopularity of behavior measures. With these findings it is possible to describe current trends in presence measurement and more importantly point towards deficiencies that researchers could approach. If presence in VR can be improved through a proper understanding of its measurement, so that one is able to adequately compare and review findings of other studies, it is likely that this will have positive effects on the development of better VR applications for simulations, entertainment and teaching.

Future Work

As we argued in the beginning, a wider circle of professional and non-professional users will come into contact with virtual reality technology in the coming years. In order to improve the comparability of recent and future results, and ensure the applicability of today's results in future systems, we hope to foster discussion on common practices and standardization.

References

- Bailenson, J. N., Rex, C., Beall, A. C., & Loomis, J. M. (2001). Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *10*, 583–598.
- Barfield, W. & Weghorst, S. (1993). The Sense of Presence within Virtual Environments: A Conceptual Framework. In G. Salvendy & M. J. Smith (Eds.), *HCI (2)* (pp. 699–704). Elsevier.
- Cummings, J. & Bailenson, J. (2015). How immersive is enough? a meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272–309. doi:10.1080/15213269.2015.1015740
- Freeman, J., Avons, S. E., Pearson, D. E., & IJsselsteijn, W. A. (1999, February). Effects of Sensory Information and Prior Experience on Direct Subjective Ratings of Presence. *Presence: Teleoperators and Virtual Environments*, *8*(1), 1–13. doi:10.1162/105474699566017
- Frommel, J., Sonntag, S., & Weber, M. (2017). Effects of Controller-based Locomotion on Player Experience in a Virtual Reality Exploration Game. In *Proceedings of the 12th International Conference on the Foundations of Digital Games (30:1–30:6)*. FDG '17. New York, NY, USA: ACM. doi:10.1145/3102071.3102082
- Greef, P. D. & Ijsselsteijn, W. (2001). Social Presence in a Home Tele-Application. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, *4*, 307–316.
- Kaul, O., Meier, K., & Rohs, M. (2017). Increasing Presence in Virtual Reality with a Vibrotactile Grid Around the Head, 289–298.
- Lee, M., Bruder, G., & Welch, G. F. (2017). Exploring the effect of vibrotactile feedback through the floor on social presence in an immersive virtual environment. In *Virtual Reality (VR), 2017 IEEE* (pp. 105–111). IEEE.
- Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: the itc-sense of presence inventory. *Presence: Teleoperators and Virtual Environments*, *10*(3), 282–297. doi:10.1162/105474601300343612

- Lin, J. J. W., Duh, H. B. L., Parker, D. E., Abi-Rached, H., & Furness, T. A. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings IEEE Virtual Reality 2002* (pp. 164–171). doi:10.1109/VR.2002.996519
- Lombard, M. & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, 3.
- Mcmahan, A. (2003). Immersion, engagement, and presence: A method for analyzing 3-D video games. *The Video Game Theory Reader*, 67–86.
- Minsky, M. (1980). Telepresence. *Omni*, 2(9), 45–52.
- Murray, C. D., Arnold, P., & Thornton, B. (2000, April). Presence Accompanying Induced Hearing Loss: Implications for Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 9(2), 137–148. doi:10.1162/105474600566682
- Nacke, L. & Lindley, C. A. (2008). Flow and Immersion in First-person Shooters: Measuring the Player's Gameplay Experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (pp. 81–88). Future Play '08. New York, NY, USA: ACM. doi:10.1145/1496984.1496998
- Newbutt, N., Sung, C., Kuo, H.-J., Leahy, M. J., Lin, C.-C., & Tong, B. (2016, September). Brief Report: A Pilot Study of the Use of a Virtual Reality Headset in Autism Populations. *Journal of Autism and Developmental Disorders*, 46(9), 3166–3176. doi:10.1007/s10803-016-2830-5
- Nilsson, N., Nordahl, R., & Serafin, S. (2016). Immersion Revisited: A Review of Existing Definitions of Immersion and Their Relation to Different Theories of Presence. *Human Technology*, 12, 108–134.
- Parent, A. (1998). A virtual environment task analysis workbook for the creation and evaluation of virtual art exhibits. *National Research Council of Canada-Reports-ERB*.
- Rietzler, M., Plaumann, K., Kränzle, T., Erath, M., Stahl, A., & Rukzio, E. (2017). VaiR: Simulating 3d Airflows in Virtual Reality. In *Proceedings of the 2017 CHI*

- Conference on Human Factors in Computing Systems* (pp. 5669–5677). CHI '17. New York, NY, USA: ACM. doi:10.1145/3025453.3026009
- Riva, G., Davide, F., & IJsselsteijn, W. A. (2003). *Being there: Concepts, effects and measurements of user presence in synthetic environments*. Ios Press.
- Rosakranse, C. & Oh, S. Y. (2014). Measuring presence: the use trends of five canonical presence questionnaires from 1998–2012. In *In proceedings of the 15th international workshop on presence (ispr'14)*.
- Salanitri, D., Lawson, G., & Waterfield, B. (2016). The Relationship Between Presence and Trust in Virtual Reality. In *Proceedings of the European Conference on Cognitive Ergonomics* (16:1–16:4). ECCE '16. New York, NY, USA: ACM. doi:10.1145/2970930.2970947
- Sanchez-Vives, M. & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature reviews. Neuroscience*, 6, 332–9.
- Schubert, T., Friedmann, F., & Regenbrecht, H. (2001, June). The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10(3), 266–281. doi:10.1162/105474601300343603
- Schuemie, M. J., van der Straaten, P., Krijn, M., & van der Mast, C. A. (2001, April). Research on Presence in Virtual Reality: A Survey. *CyberPsychology & Behavior*, 4(2), 183–201. doi:10.1089/109493101300117884
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Wiley.
- Skarbez, R., Brooks, F. P., & Whitton, M. C. (2017). Immersion and coherence in a visual cliff environment. (pp. 397–398). IEEE. doi:10.1109/VR.2017.7892344
- Slater, M. (1999, October). Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 8(5), 560–565. doi:10.1162/105474699566477
- Slater, M. (2004, August). How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 13(4), 484–493. doi:10.1162/1054746041944849

- Slater, M. & Steed, A. (2000). A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments*, 9(5), 413–434. doi:10.1162/105474600566925
- Slater, M., Usoh, M., & Steed, A. (1994, January). Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 3(2), 130–144. doi:10.1162/pres.1994.3.2.130
- Slater, M. & Wilbur, S. (1997a, December). A framework for immersive virtual environments five: speculations on the role of presence in virtual environments. *Presence: Teleoper. Virtual Environ.* 6(6), 603–616. doi:10.1162/pres.1997.6.6.603
- Slater, M. & Wilbur, S. (1997b, December). A Framework for Immersive Virtual Environments Five: Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoper. Virtual Environ.* 6(6), 603–616. doi:10.1162/pres.1997.6.6.603
- Sohre, N., Mackin, C., Interrante, V., & Guy, S. J. (2017, March). Evaluating collision avoidance effects on discomfort in virtual environments. (pp. 1–5). IEEE. doi:10.1109/VHCIE.2017.7935624
- Steuer, J. (2000). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, 42.
- Usoh, M., Catena, E., Arman, S., & Slater, M. (2000, October). Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments*, 9(5), 497–503. doi:10.1162/105474600566989
- van Baren, J. & IJsselsteijn, W. (2004, March). *Measuring Presence: A Guide to Current Measurement Approaches*.
- Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., . . . Jäncke, P. (2004). Mec spatial presence questionnaire (mec-spq): short documentation and instructions for application. In *Report to the european community, project presence: mec (ist-2001-37661)*.
- Witmer, B. G. & Singer, M. J. (1998, June). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225–240. doi:10.1162/105474698565686