Literature Review

Presence in Virtual Reality: Methods of Measurement

Johannes Schirm

25040946

Contents

1   Introduction

Research on elementary aspects of virtual reality is being conducted for several decades now, which has led to significant progress in the field. Nevertheless, some rather important questions are still difficult to answer when designing immersive virtual environments (IVEs). These are mostly related to the way humans perceive IVEs, as the involved technology steadily improves and therefore manages to provide experiences of increasing quality. An example of technical limitations that presented less and less of a problem over time is frame rate: Barfield and Hendrix (1995) thought that, even with only a maximum of 25 frames per second, 'computational resources could be saved using a slower update rate while maintaining a given level of presence', whereas 17 years later, van Waveren (2016, p. 38) explains that '[the delay between sampling tracking coordinates and emitting corresponding photons] should be 20 milliseconds or less for compelling virtual reality experiences.' This is twice as many frames to compute as two decades ago, but with high-performance graphics cards and optimisation techniques, this has become the standard for contemporary implementations of IVEs.

However, human-related problems are often more difficult to solve because they usually depend on how an individual perceives their environment and the specific purpose an IVE is used for. An important factor in this regard is the degree to which a person feels presence in an IVE, in literature commonly referred to as an individual's sense of "being" at the virtual place. While it is possible to collect objective data about computational performance and act upon the results by improving hardware and software, measuring how present an individual is when experiencing an IVE is generally a very difficult task. This literature review aims to give an overview of which objective and subjective methods have already been used in order to scientifically measure presence and how reliable their results are seen as by different authors.

2   Selection Criteria

Most of the progress on presence in IVEs until now has been accomplished by slow degrees in the context of conference papers and journal articles. In a recent literature survey about concepts and measurement of presence by Skarbez, Brooks and Whitton (2017b), the authors reference more than 150 publications and still conclude the following: 'An ideal measurement of presence would be reliable, that is, producing repeatable results, both

within and between subjects; valid, that is, demonstrated to correlate with the subjective feeling of presence; multi-level-sensitive; and objective [...]. These are standard principles of sound testing practices [...]. We would go further and suggest that such an ideal metric should also be measurable contemporaneously, continually, and without modification to the scenario, and should be generalizable across [I]VEs. No measure of presence yet exists that meets all of these criteria.' Firstly, this statement helps to understand that this topic has been worked on for many years now and there is still no entrenched method everybody agrees on, which is why literature mainly consists of many different model proposals or studies that apply these models. Secondly, it becomes clear that only a fraction of the relevant publications can be considered in this context. This section explains both formal inclusion and exclusion criteria and the content-related selection strategy used.

2.1   Inclusion Criteria

The body of relevant literature primarily features studies for which there was a direct or indirect need to measure the presence of participants in an IVE, as well as theoretical publications explicitly focussing on models of presence. Both of these can be valuable resources when compared with one another. Since research in this field is still in the phase of finding novel approaches, it was considered more useful to include a broader range of interesting ideas and opinions rather than a full collection of already established methods. Another reason for this is that the latter are biased towards subjective measurement methods, as these are much more practical.

A publication was seen as suitable for inclusion if it had been peer-reviewed, especially in the context of a conference or a journal with strong relation to the topic. This way, it was guaranteed that the scientific community with experts in this field reviewed it and approved its methods. Usually, one has to be a little bit cautious when including conference papers, as they sometimes present very novel ideas or preliminary conclusions in a smaller format. For this topic, however, a good number of them was included in order to broaden the spectrum of ideas. Interesting examples for this are Deniaud, Honnet, Jeanne and Mestre (2015), Soyka et al. (2015) and Nunez and Blake (2003). Another property that made publications particularly eligible for inclusion was the way they dealt with existing approaches: Even if a study was not primarily concerned with understanding presence, it could provide valuable input just through the justification for the authors' choice of a specific method of presence measurement or through newly invented methods.

2.2 Exclusion Criteria

It should go without saying that only peer-reviewed publications were considered eligible for inclusion. But to define an even more precise requirement, it was expected that each document was accessible through its own Digital Object Identifier (DOI). Some documents on the topic, for instance smaller reports or dissertations, were interesting but never formally published and therefore had no DOI assigned to them. Examples for this are 'Immersion and Presence' by Daniel R. Mestre or 'Physiological Reaction as an Objective Measure of Presence in Virtual Environments' by Michael Meehan. Others had their DOI apparently withdrawn recently or were still in the phase of waiting for a new DOI to be assigned to them. One interesting publication had to be excluded because it is not written in English: 'Presence in Virtual Environments: Objective Metrics vs. Subjective Metrics: A Pilot Study' by Melo, Rocha, Barbosa and Bessa.

In general, the topic seemed to be dominated by high-quality research which contributed to major conferences or journals. Because of this, there were very few cases in which a publication did actually not meet the requirements described above. Furthermore, there are almost no published treatises or books on the topic, as it was always evolving and a much more appropriate subject to experimental and theoretical publications. Although this almost entirely removed the need for further exclusion criteria, there was one content-related condition set: The term 'presence' is not necessarily connected to IVEs, since 'media such as the telephone, radio, television, film and many others offer a lesser degree of presence as well', following Lombard and Ditton (1997). It was therefore verified that included publications were predominantly relating the term 'presence' to IVEs and not to other forms of media.

3 Literature Analysis

The perception of presence in IVEs constantly evolved and a good number of models were thought of, trying to describe and prioritise all components involved. Even so, a few methods of measurement became particularly popular and were much more frequently incorporated into studies than others. One of the most present examples is the questionnaire presented by Witmer and Singer (1998), references to which can be found in many publications on the topic. It is taken by participants after experiencing the IVE and contains questions from the four categories control, sensory, distraction and realism.

In a later publication by Slater, Spanlang and Corominas (2010, p. 92:2), the authors discourage the use of the questionnaire as the only measure by referencing several problems that were found: 'It does not seem to be able to distinguish between an experience in reality and virtual reality [...], the measurements may be unstable [...], it has problems in actually assessing the concept itself [...], and there are methodological problems in analyzing subjective rating data as if it were interval or ratio data [...].' In Slater (2009), they presented an alternative model instead, splitting up the term into 'Place Illusion' (PI) and 'Plausibility Illusion' (Psi): 'We have called PI the "being there" qualia that was referred to as "presence" in the original literature: It is the feeling of being in the place depicted by the virtual environment (even though you know that you are not there). We call the Psi the illusion that what is happening is real (even though you know that it is not real).' Following further publications from Slater and Garau (2007), Sanchez-Vives and Slater (2005) and Slater (2004), the research community slowly moved towards always using both subjective methods and objective methods in order to make their measures of presence more comprehensible.

3.1   Models of Presence

Many different models have been considered for explaining how the concept of presence should best be broken down into components that are easier to assess. Skarbez et al. (2017b, p. 96:18) manage to reference 14 models, without even including 'higher-order concepts' (p. 96:17) like PI and Psi. An in-depth discussion of the models would go beyond the scope of the literature review — this subsection will only summarise the main concepts.

On the most fundamental level, Flach and Holden (1998) thought through different approaches to the 'science' of presence. They asked to what extent it is relative to the individual experiencing an IVE and specifically pointed out a possible explanation with the help of Gibson's well-known studies of the human perception as being solely defined through perceived interactions with the environment (pp. 93-94). While Bystrom, Barfield and Hendrix (1999) already started decomposing the phenomenon into immersion, presence and performance, Lombard and Ditton (1997) listed a considerable number of views on the term itself, primarily social and perceptual aspects.

Alongside the model based on the concept of neuro-linguistic programming by Slater, Usoh and Steed (1994), the questionnaire by Witmer and Singer (1998) then gained popularity within the research community. Even though Nichols, Haldane and Wilson (2000) and Hoffman, Richards, Coda, Richards

and Sharar (2003) took first steps towards objective methods of measurement by considering behaviour, attention and magnetic resonance imaging in their experiment designs, the results were perceived as ambiguous and hard to interpret even by the authors themselves. Both appeared to be much more confident presenting data from questionnaires and verbal ratings.

Only after PI and Psi were introduced by Slater (2009) and further work on objective measurements — particularly utilising methodologies from psychophysics — was accomplished by Slater et al. (2010), it became a standard to combine subjective and objective methods of measurement. From then on, not many publications (none in the subset reviewed here) explicitly tried to develop new models of presence. One exception is an interesting theoretical framework by Szczurowski and Smith (2017), proposing three hypothetical experiments that evaluate 'capacity of the medium for inducing presence', 'factors loading on presence' and also 'brain function theory hypothesis in relation to [IVEs]'. This is one of the most recent publications considered in this literature review and provides a promising starting point for designing a novel presence study in an 'applied research' (p. 17:7) context.

3.2 Subjective Methods

Subjective measurements of presence are taken through verbal judgements in real-time or post-questionnaires on paper. An early approach by Slater et al. (1994, p. 18) was to directly expose participants to their definition of presence and ask them to rate a few items on a Likert scale. They concluded that 'subjects have a certain baseline level of presence in the [I]VE, and the questionnaire may lead to an expectation that the experimenters are looking for answers that are beyond this baseline' (p. 28), which showed first problems with this approach. Barfield and Hendrix (1995, p. 7) expanded this questionnaire by items regarding 'Fidelity of Interaction', which they viewed as implicit factors on presence. They also noted that 'questions relating to interactivity [...], which we thought a priori would form a separate factor, loaded significantly on the virtual presence factor.' Again, there was no clear notion of what factors actually influence presence.

For a study of fear of heights, Regenbrecht, Schubert and Friedmann (1998, p. 240) created their own questionnaire 'from interviews with long-term users of VR applications'. Although Witmer and Singer (1998) already presented their comprehensive questionnaire, Slater, Steed, McCarthy and Maringelli (1998) continued refining their own questions for studying body movement: To them, it mixed the distinct concepts of immersion and presence.

7

From then on, these main approaches always had to be considered by the research community: Studying the effect of different levels of IVE quality on presence, Nunez and Blake (2003, p. 104) acknowledged that 'although several presence scales are available in the literature, the choice of which to use is far from trivial' and considered not only the questionnaire by Witmer and Singer (1998) but also the ratings by Slater, Usoh and Steed (1995). Bouchard, St-Jacques, Robillard and Renaud (2008) studied anxiety and combined the two approaches, as well, arguing that 'single-item measures [...] are less intrusive than questionnaires if one wants to use them during the immersion.' And while McQuiggan, Rowe and Lester (2008) only used the questionnaire by Witmer and Singer (1998) for an evaluation of their learning environment, Phillips, Interrante, Kaeding, Ries and Anderson (2012) only used the questions by Slater et al. (1995) for perceptual studies.

Depending on the context and the preferred model of presence, different methods were utilised in almost all subsequent studies. Even though these were still derivatives of the two main approaches referenced before, modified statistical procedures and question pools lead to a large number of employed subjective measurement methods. Since it would not make sense to randomly pick a few of them for this review, the comprehensive literature survey by Skarbez et al. (2017b, p. 96:25) is referenced at this point instead.

3.3   Objective Methods

From the included publications, Nichols et al. (2000) were the first authors to explicitly make objective measures part of their experiment design. They compared objective measurements like 'physiological monitoring, self-report of symptoms and other experiences, postural assessment and visual, physical and psychomotor performance tests' (p. 475) with a subjective measurement in a first experiment. Although they found the results of the reflex response measure to be promising, they retained more confidence in use of rating scales that were generally easier to apply (p. 438). In a much earlier study by Slater et al. (1995), the authors evaluated the effect of two locomotion techniques on presence both through a questionnaire and the path taken by participants through the IVE. But since measures like this are completely dependent on the specific IVE, they are not considered in this section.

IJsselsteijn, de Ridder, Freeman and Avons (2000) presented a summary of what measures of presence needed to be considered at the time. They did not only consider the existing categories of subjective measures of presence, but also postural responses, physiological measures, dual task measures and

social responses as objective forms measurements (p. 524). Their conclusion is the following: 'It seems reasonable to assume that not one overall presence measure will be developed, but rather an aggregate measure of presence that is comprised of both subjective and objective components, in order to avoid the limitations of either alone' (p. 527). This was understood by many other authors, which is why most of them started to expand their experiment designs by more objective measures.

Furthermore, Zimmons and Panter (2003) measured skin conductance and heart rate for their study about influences of rendering quality on presence and noticed a significant correlation between heart rate, reported presence and task performance. Several years later, Slater et al. (2007) started their work on the 'Presenccia' project in order to 'assess the extent of presence using tools beyond traditional questionnaires, and therefore [...] avoid many of the problems involved with sole reliance on these'. In this context, Slater et al. (2010) published a study on the psychophysics of presence: They now argue that 'presence is identified with its operationalization as a measureable property of the actions of people within IVEs compared with their expected or actually observed behavior within similar real-world settings.' Building upon their most recent model of presence — which consists of PI and Psi — they introduced a response function that can be used to predict how 'an average participant' would respond to a specific IVE in various regards (p. 92:8).

Through an empirical comparison between entrenched questionnaires and physiological measures, Skarbez, Brooks and Whitton (2017a) were able to draw several conclusions, including 'good evidence that the [...] presence questionnaire [by Witmer and Singer (1998)] responds to higher levels of immersion as a main effect' (p. 398). Some more examples of authors who incorporated both subjective and objective measurements into their studies are quickly described hereafter: Malbos, Rapee and Kavakli (2012) showed that their experimental behavioural scale was 'consistent, reliable, and [a] valid instrument to measure presence in threatening virtual situations', as well as 'an easy-to-use complementary test to the subjective and physiological measures of presence.' Although this scale can only be used in particularly threatening IVEs, it could be of use in situations where this is deliberate. Deniaud et al. (2015) took many physiological measurements for studying driving simulation and concluded that 'presence measurement can't be only based on subjective measures' (p. 747). Combining driving performance, heart rate and skin conductance level, they had a diverse set of objective data to compare the usual questionnaire results against.

4 Conclusion

Despite several decades of research and many innovations in this field, there is no complete agreement within the research community about what exactly defines presence and how it should be measured in a reliable way (Skarbez et al., 2017b). In general, subsequent studies should be designed to collect both subjective and objective measurements. For subjective measurements, there is a huge number of questionnaires to choose from. Nevertheless, it might be such a huge number because it is often needed to adjust entrenched instruments to better fit a specific scenario. From this perspective, it might be good to always derive a customised procedure for each scenario. There are not as many objective measures that could possibly be taken, but on the other hand, they usually require a lot of effort to integrate into an experiment design. Although there were several possibilities discussed in section 3.3, the work of Skarbez et al. (2017b, pp. 96:29-96:31) is referenced at this point, as they offer a more general pool of common objective methods.

Even though it is important to consider objective methods to the same degree as subjective methods, one must be careful to not choose methods that are too intrusive and change the IVE or the experiment scenario too much. The less influence an objective method has on the participant, the better. This is why physiological measures such as heart rate and skin conductance are sometimes preferred to methods that would need the IVE to change, for example make it more threatening. Still, it takes more effort to implement these, as the experimenter has to manage all the corresponding hardware and the results need to be mapped and analysed afterwards, as well.

With regards to advancing presence research, Szczurowski and Smith (2017) bring very inspiring ideas into play with their recent paper. Firstly, they propose three hypothetical experiments, which could be adapted to a specific device and a specific IVE. Interpreting the results would be much easier than with a study that had to be designed from scratch, since the authors also provide 'possible experiment results and their interpretation'. Secondly, they call for 'an Open Science model for reporting experiments results and sharing as much data and designs as it's legally possible' (p. 5), an idea which they already supported by allowing others to adapt their experiment templates and therewith accelerate the increase of the critical mass of the research community, something that is urgently needed (p. 5).

References

Barfield, W. & Hendrix, C. (1995, June). The effect of update rate on the sense of presence within virtual environments. *Virtual Reality*, *1*(1), 3-15. doi: 10.1007/BF02009709

Bouchard, S., St-Jacques, J., Robillard, G. & Renaud, P. (2008, August). Anxiety increases the feeling of presence in virtual reality. *Presence*, *17*(4), 376-391. doi: 10.1162/pres.17.4.376

Bystrom, K.-E., Barfield, W. & Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *8*(2), 241-244. doi: 10.1162/105474699566107

Deniaud, C., Honnet, V., Jeanne, B. & Mestre, D. (2015). An investigation into physiological responses in driving simulators: An objective measurement of presence. In *Science and information conference (sai)* (p. 739-748). doi: 10.1109/SAI.2015.7237225

Flach, J. M. & Holden, J. G. (1998). The reality of experience: Gibson's way. *Presence: Teleoperators and Virtual Environments*, *7*(1), 90-95. doi: 10.1162/105474698565550

Hoffman, H. G., Richards, T., Coda, B., Richards, A. & Sharar, S. R. (2003). The illusion of presence in immersive virtual reality during an fmri brain scan. *CyberPsychology and Behavior*, *6*(2), 127-131. doi: 10.1089/109493103321640310

IJsselsteijn, W. A., de Ridder, H., Freeman, J. & Avons, S. E. (2000). Presence: concept, determinants, and measurement. In *Spie electronic imaging* (Vol. 3959, p. 520-529). doi: 10.1117/12.387188

Lombard, M. & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*(2), JM321. doi: 10.1111/j.1083-6101.1997.tb00072.x

Malbos, E., Rapee, R. M. & Kavakli, M. (2012, August). Behavioral presence test in threatening virtual environments. *Presence*, *21*(3), 268-280. doi: 10.1162/PRES_a_00112

McQuiggan, S. W., Rowe, J. P. & Lester, J. C. (2008). The effects of empathetic virtual characters on presence in narrative-centered learning environments. In *Proceedings of the sigchi conference on human factors in computing systems* (p. 1511-1520). New York, NY, USA: ACM. doi: 10.1145/1357054.1357291

Nichols, S., Haldane, C. & Wilson, J. R. (2000). Measurement of presence and its consequences in virtual environments. *International Journal of Human-Computer Studies*, *52*(3), 471. doi: 10.1006/ijhc.1999.0343

Nunez, D. & Blake, E. (2003). Conceptual priming as a determinant of presence in virtual environments. In *Proceedings of the 2nd international conference on computer graphics, virtual reality, visualisation and interaction in africa* (p. 101-108). New York, NY, USA: ACM. doi: 10.1145/602330.602350

Phillips, L., Interrante, V., Kaeding, M., Ries, B. & Anderson, L. (2012, May). Correlations between physiological response, gait, personality, and presence in immersive virtual environments. *Presence*, *21*(2), 119-141. doi: 10.1162/PRES_a_00100

Regenbrecht, H. T., Schubert, T. W. & Friedmann, F. (1998). Measuring the sense of presence and its relations to fear of heights in virtual environments. *International Journal of Human–Computer Interaction*, *10*(3), 233-249. doi: 10.1207/s15327590ijhc1003_2

Sanchez-Vives, M. V. & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews.Neuroscience*, *6*(4), 332-9. doi: 10.1038/nrn1651

Skarbez, R., Brooks, F. P. & Whitton, M. C. (2017a). Immersion and coherence in a visual cliff environment. In *Ieee virtual reality* (p. 397-398). doi: 10.1109/VR.2017.7892344

Skarbez, R., Brooks, F. P. & Whitton, M. C. (2017b, November). A survey of presence and related concepts. *ACM Computing Surveys*, *50*(6), 1-39. doi: 10.1145/3134301

Slater, M. (2004). 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. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1535), 3549-3557. doi: 10.1098/rstb.2009.0138

Slater, M., Frisoli, A., Tecchia, F., Guger, C., Lotto, B., Steed, A., . . . Bernardet, U. (2007, July). Understanding and realizing presence in the presenccia project. *IEEE Computer Graphics and Applications*, *27*(4), 90-93. doi: 10.1109/MCG.2007.93

Slater, M. & Garau, M. (2007). The use of questionnaire data in presence studies: Do not seriously likert. *Presence: Teleoperators and Virtual Environments*, *16*(4), 447-456. doi: 10.1162/pres.16.4.447

Slater, M., Spanlang, B. & Corominas, D. (2010). Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. In *Acm siggraph 2010* (Vol. 29, p. Article 92). New York, NY, USA: ACM. doi: 10.1145/1833349.1778829

Slater, M., Steed, A., McCarthy, J. & Maringelli, F. (1998). The influence

of body movement on subjective presence in virtual environments. *Human Factors: The Journal of Human Factors and Ergonomics Society*, *40*(3), 469-477. doi: 10.1518/001872098779591368

Slater, M., Usoh, M. & Steed, A. (1994). 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., Usoh, M. & Steed, A. (1995). Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction*, *2*(3), 201-219. doi: 10.1145/210079.210084

Soyka, F., Kokkinara, E., Leyrer, M., Buelthoff, H., Slater, M. & Mohler, B. (2015). Turbulent motions cannot shake vr. In *Ieee virtual reality* (p. 33-40). IEEE. doi: 10.1109/VR.2015.7223321

Szczurowski, K. & Smith, M. (2017). Measuring presence: Hypothetical quantitative framework. In *23rd international conference on virtual system and multimedia* (p. 1-8). doi: 10.1109/VSMM.2017.8346261

van Waveren, J. M. P. (2016). The asynchronous time warp for virtual reality on consumer hardware. In *Proceedings of the 22nd acm conference on virtual reality software and technology* (p. 37-46). New York, NY, USA: ACM. doi: 10.1145/2993369.2993375

Witmer, B. G. & Singer, M. J. (1998, June). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225-240. doi: 10.1162/105474698565686

Zimmons, P. & Panter, A. (2003). The influence of rendering quality on presence and task performance in a virtual environment. In *Ieee virtual reality* (p. 293-294). IEEE. doi: 10.1109/VR.2003.1191170