A Separate Reality: An Update on Place Illusion and Plausibility in Virtual Reality

Mel Slater[1,2]*, Domna Banakou[1,2], Alejandro Beacco[1], Jaime Gallego[1], Francisco Macia-Varela[1] and Ramon Oliva[1]

[1]Event Lab, Faculty of Psychology, Universitat de Barcelona, Barcelona, Spain, [2]Institute of Neurosciences of the University of Barcelona, Barcelona, Spain

We review the concept of presence in virtual reality, normally thought of as the sense of "being there" in the virtual world. We argued in a 2009 paper that presence consists of two orthogonal illusions that we refer to as Place Illusion (PI, the illusion of being in the place depicted by the VR) and Plausibility (Psi, the illusion that the virtual situations and events are really happening). Both are with the proviso that the participant in the virtual reality knows for sure that these are illusions. Presence (PI and Psi) together with the illusion of ownership over the virtual body that self-represents the participant, are the three key illusions of virtual reality. Copresence, togetherness with others in the virtual world, can be a consequence in the context of interaction between remotely located participants in the same shared virtual environments, or between participants and virtual humans. We then review several different methods of measuring presence: questionnaires, physiological and behavioural measures, breaks in presence, and a psychophysics method based on transitions between different system configurations. Presence is not the only way to assess the responses of people to virtual reality experiences, and we present methods that rely solely on participant preferences, including the use of sentiment analysis that allows participants to express their experience in their own words rather than be required to adopt the terminology and concepts of researchers. We discuss several open questions and controversies that exist in this field, providing an update to the 2009 paper, in particular with respect to models of Plausibility. We argue that Plausibility is the most interesting and complex illusion to understand and is worthy of significant more research. Regarding measurement we conclude that the ideal method would be a combination of a psychophysical method and qualitative methods including sentiment analysis.

Keywords: virtual reality, presence, place illusion, plausibility, body ownership, questionnaires, measurement

INTRODUCTION

Some physicists argue that there are an infinite number of parallel universes and even that these universes interact with one another at the quantum level (A very interesting discussion of the multiverse can be found in (Deutsch, 2011)). It is very hard for people not trained in mathematics or physics to have any idea what this might mean. However, with Virtual Reality (VR) we do have an example of how a parallel Universe (of sorts) can occupy the same physical space as our Universe in space and time. You are in your living room, and you don a head-mounted display. You perceive an

alternate computer-generated world through a wide field-of-view head-mounted display (HMD), in stereo. As you move your head the images that you see and the sounds you hear update in predictable ways, enabled through 6 degrees of freedom head tracking. If you see a close object, you can move your head to see what is behind it. If you bend down, you might see under it. In this case the world that you perceive is a concert that happened in 1983. You look around and see that you are in a theatre. There are thousands of other cheering people there. You see several people run onto a stage and when they are in position some guitar music starts to play. From a loudspeaker booming out across the theatre you hear "Good evening ladies and gentlemen, would you please welcome to the stage—Dire Straits!"

Where are you? Of course, you are in your living room. But your sensory system only perceives the reality of the theatre. The band starts to play "Sultans of Swing" and the audience around you dance in rhythm with the music. You find that you are dancing too. You look around at the people close to you. One of them appears to be staring at you. Another quickly turns away when you look at him. What am I doing here? a woman alone at this concert, with some quite creepy men around me. Maybe one of them will come over and start to try to engage with me. This is a nightmare.

This actually happened. Research that started from computer vision techniques to extract 3D geometry, texture and animation from video footage of a Dire Straits concert, and that used crowd simulation to create the audience transforming the whole into a VR scenario, was experienced as a "nightmare" by some participants in a pilot study (Beacco et al., 2021). Although unexpected and apparently negative, this was a highly informative result. In reality there was no theatre, no band playing and no audience. Nevertheless, the objective reality of lit pixels on two small screens and the sound from loudspeakers on the HMD were experienced as a live scene where many events were happening—including, for some women, frightening ones such as audience members "ogling" and "pretending not to be staring at" them, arousing fears of possible sexual harassment. In fact, none of these negative events at all were programmed into the scenario—there were no virtual audience members staring at participants (except a momentary glance by chance)—yet such illusory events were perceived by about half of the 20 participants in the pilot study.

How is it possible for people to have such experiences? We argue that this occurs through a number of illusions that are generated through VR experiences. People act out of these illusions as if they were real. In the rest of this paper we discuss those illusions, how they have been measured, and follow-on with some open questions and controversies.

THE ILLUSIONS OF VR

VR and other immersive media such as augmented reality (AR, considered later) can generate at least three unique *illusions* that are not possible with other media. Here we mean illusions in the sense that people have perceptions that arise from digital sources that are totally different from what is actually being perceived.

The perceptions are real perceptions, and people may act on them as if they were real (Chalmers, 2017, 2022; Slater and Sanchez-Vives, 2022). For example, a participant in VR perceives a 0.5 m³ cube stationary in space which reflects light that suggests that it is metallic, and which can be walked around and looked at from various locations, including ducking underneath it. There is an illusion of seeing a heavy cube somehow floating in the air. But what is actually happening? The participant is looking at two small screens modulated by an optical system, which displays illuminated pixels in a variety of colours. One shows a left eye view and the other a right eye view computed from the perspective projections of computer graphics. The brain integrates all this information into the perception of a cube. What is perceived is nothing whatsoever like the source of the perception. While standing in front of the cube it suddenly falls to the ground. Most participants would jump backwards–an automatic response caused by fear of the cube landing on their toes. This "falling" is nothing more than a sequence of colour changes of some of the illuminated pixels. Although one could make the same argument about perception in reality (we do not see "reality"—atoms, quarks etc. but an interpretation constructed by our sensory apparatus and brain processing) the source is quite different. For example, the entire physical apparatus responsible for perception in VR (computers, head mounted displays etc) can also be decomposed into atoms, quarks etc.—but the physical atoms cannot be decomposed into pixels on a display. So the VR apparatus and system is at a different level to physical reality, and the fact is that in spite of perceiving and reacting to a cube, there is no cube there.

The first two illusions we consider are part of the concept of "presence" originally considered as the sense of "being there" in the virtual world (Held and Durlach, 1992; Sheridan, 1992, 1996; Sanchez-Vives and Slater, 2005). Presence has been decomposed into two dimensions—"Place Illusion" (PI) and Plausibility (Psi). PI is the illusion of being in the place depicted in the VR ("being there") in spite of the sure knowledge that this is not the case. "Plausibility" (Psi), is the illusion that events in the VR are actually occurring, that what is perceived is happening (a cube is falling towards your toes), again in spite of the knowledge that the events are digitally generated and nothing that is apparently happening is actually happening in reality (Slater, 2009). Events in VR correspond to changes in the illumination of pixels, changes in generated sound, and possibly other digitally generated sensory information such as haptics or olfactory.

The basis of PI is that perception in VR is through natural sensorimotor contingencies (O'Regan and Noë, 2001)—that is, participants perceive through using their body following much the same rules as in physical reality—turning the head, bending down, reaching out, looking around objects. A head turn, for example, results in updates to the displayed images so that participants would see visual changes and hear auditory changes corresponding to what would occur in reality. This is based on 6 degrees of freedom head-tracking, and possibly eye tracking. This integrates both interaction and display—since natural sensorimotor contingencies demand that what is displayed conforms with body movements. For example, if the participant looks closely at an object and sees pixels, then this breaks PI.

Psi relies on 1) the virtual environment responding to actions of the participant (for example, the participant looks towards a virtual human character that then responds by looking back); 2) events in the environment contingently referring to the participant (for example a virtual human character smiles towards the participant without any prior cause). 3) Where the virtual reality is a simulation of something that could occur in reality, depicting a situation in which the participant has expertise, then virtual events should meet expectations. This may be highly specific to the individual. In the virtual concert scenario discussed above, some participants noticed that the drummer was sometimes not moving in sync with the beat, or even that the light reflected from the guitar strings of the performers did not match the sounds that they were making. Also the concert was supposed to be in the 1980s but no one in the audience was smoking. This last example required that the participant concerned knew that smoking in theatres was still possible in London, United Kingdom, in the 1980s. Other participants who did not know this would not find the lack of smoking to be a failure of expectations. In a study of medical doctors who were in consultation with virtual patients, many complained that they could not read the screen of the computer on their (virtual) desk that they would typically be doing in reality (Pan et al., 2016). Skarbez et al. (2018a) introduced the term "coherence" to describe some of the factors that go into producing Psi, and we will return to this concept later.

PI and Psi are conceptually distinct and can be considered as orthogonal axes that make up an overall response that might be labelled as "presence." Empirically they may be correlated, although they have been found to be different in, for example, (Slater et al., 2010a; Hofer et al., 2020) but with unclear results in (Brubach et al., 2022). Psi is beginning to be extensively studied e.g. (Bergström et al., 2017; Skarbez et al., 2017; GalvanDebarba et al., 2020). The important point of this is that when both PI and Psi operate people tend to respond realistically to situations and events in the VR, even though they know for sure that these are illusions and not reality.

Psi is the more complex and interesting aspect, since a seemingly unimportant feature of the environment that does not fit expectations can result in its loss. For example, in early versions of a study of bystander responses to violence, where the bystander was in a bar talking to one person who was attacked by another over an issue to do with soccer (Rovira et al., 2009; Slater et al., 2013), we were told by some participants that the scenario was not consistent with what would be expected in reality. Why? because a bar with that type of decoration would never be patronised by soccer fans. On the other hand in an environment that was used for cognitive behavioural therapy for people with fear of heights, participants looked into an atrium formed by several tall buildings (Freeman et al., 2019). A whale was "swimming" between the buildings, yet people simply accepted this without comment. In a 3D chess scenario, when participants touched a chess piece it would fly to its next location (Slater et al., 1996). No participant found this odd, and when asked about it, one said that: In this world that is the way things are.

The third illusion is "body ownership." When participants wear a wide field-of-view head-tracked stereo head-mounted display and they look down towards themselves, they will see a life-sized virtual body substituting their own, from their first person perspective (1PP) (if this has been programmed). Utilising real-time body tracking, as the person moves, their virtual body can be programmed to move synchronously and in correspondence with their own movements (referred to as visuomotor synchrony). If something is seen to touch their virtual body, then it can be arranged that a corresponding tactile stimulation is applied synchronously to their own real body (visuotactile synchrony). We refer to this as *embodiment*, which involves multisensory integration of the 1PP view of the body, and visuomotor or visuotactile synchrony, which will typically lead to the illusion that the virtual body is their own—even though they know for sure that it is not. Based on the original finding of the Rubber Hand Illusion (Botvinick and Cohen, 1998) ownership over a virtual body has been demonstrated multiple times—for example, (Petkova and Ehrsson, 2008; Slater et al., 2008; Slater et al., 2010b; Banakou and Slater, 2014). Moreover, changing the type of body can lead to physiological, behavioural, attitudinal, and cognitive changes in the participant. For example, multiple replications have shown that embodying White people in a Black virtual body will lead to sustained reduction in their implicit racial bias (Peck et al., 2013; Maister et al., 2015; Banakou et al., 2016; Hasler et al., 2017; Banakou et al., 2020). However, if the affective situation in which this occurs is a negative one, then implicit bias tends to increase rather than decrease (Banakou et al., 2020).

There is a fourth illusion that can be thought of as a corollary of these three, referred to as copresence (Nowak and Biocca, 2003). When in a virtual environment simultaneously with other remotely located people copresence refers to the extent to which the participant has the illusion of being there with the others, or virtual togetherness (Durlach and Slater, 2000). Participants should be represented in some form with virtual bodies, otherwise it would not be possible to know that they are there or where they are. Sensorimotor contingencies lead to the possibility of the participant having the illusion to be in the same space as the others. For example, each person has a representation that in principle the participant can walk around, look behind, hear the voice from different locations, reach out and touch, and so on. Exactly the same requirements for PI incorporate this aspect of copresence—i.e., the illusion of sharing the same space. Then Psi is needed for participants to be able to take the interaction events as really occurring. Hence the characters should respond when interacted with, for example, change gaze direction, and characters should be able to initiate interaction with the participants (Garau et al., 2005). The characters should move appropriately depending on context, so that the expectation "this is how a human should behave" is met. The requirement to satisfy expectations is the most difficult because it depends very much on context. For example, if the characters are represented as high quality human models, then they should behave according to this, but if they are represented, say, as clearly cartoon characters, then expectations about their behaviour might be different.

Copresence leads to results that would be expected under similar circumstances in reality. A good example of this is

proxemics, where people maintain different distances from one another depending on their relationship at the time: intimate, personal, social and public (Hall, 1973). Bailenson et al. (2003) found that proxemics predictions from reality also operate in VR. Llobera et al. (2010) found that skin conductance varied with distance of virtual human characters from the experimental participant in accordance with proxemics theory. Kastanis and Slater (2012) found that a Reinforcement Learning agent, represented in VR as a human character, learned that if it moved close to experimental participants they would back away (since the character entered their personal or intimate space causing discomfort) and then it used that to manipulate participants to a target spot in the virtual environment. A recent overview and review can be found in (Williamson et al., 2021).

There is a question as to whether the issue of copresence only refers to the situation of remote participants meeting in the same environment, or whether it can also apply to sharing the space with virtual characters controlled by an algorithm. In fact, there is a range of possibilities between these two (full control of the avatar by an embodied person, or control of an avatar by a computer program). Suppose that the avatar is only partially controlled by a remote person—for example, the person controls the location and gaze direction, but not the bodily gestures because there is no motion capture. Or suppose that the remote person only controls the speech, which is then mapped to appropriate lip-sync on the virtual human character, but the facial expressions and body gestures are wholly determined by an algorithm. In any of these situations the issue of copresence (do participants feel together with such representations?) is the same. Therefore, it can be argued that copresence first, is not a separate phenomenon from PI, Psi, or body ownership, if these are operating correctly and applied to the representations and actions of virtual characters, then there will be copresence. Second, copresence can occur whether the other characters are wholly representative of remote people or wholly determined by a computer program, or any variant in between these two extremes.

Recent reviews on the concept of presence can be found in (Parola et al., 2016; Skarbez et al., 2018b; a; Felton and Jackson, 2022) with a meta study in (Cummings and Bailenson, 2016). Grassini and Laumann (2020) focus on the variety of measures that have been used, and Souza et al. (2021) present a meta-study of over 1,200 papers on methods of measurement of presence. It is noteworthy that 86% of these papers used subjective measures and 12% used both subjective and objective. In the next section we discuss different approaches to measurement.

THE PROBLEM OF MEASUREMENT

Questionnaires

Since presence refers to subjective illusions (PI, Psi) the obvious way to elicit the sense of presence has been through questionnaires. Schwind et al. (2019) identified 15 questionnaires in their review. A widely used one is by Witmer and Singer (1998), which focuses on asking opinions of participants about a number of factors that have been thought to promote presence, but it has no questions about the sensation of "being there" itself (Slater, 1999). An early questionnaire that came to be known as SUS, though never formally validated, appeared, for example, in (Slater et al., 1994), and concentrated on the participant's sense of "being there" in the virtual world, the times when the virtual world became the participant's reality so that they forgot about the real world, and the extent to which participants evaluated the virtual world as "somewhere that you saw" or "somewhere that you visited." A fully validated questionnaire based on factor analysis was developed by Lessiter et al. (2001), which was intended also as cross-media (not only for immersive virtual environments). Their factor analysis pointed to three major components. The first is the "sense of physical space" which is closely related to the general view of presence (in particular PI) as "being there." The second factor "engagement" refers to how much participants are involved and interested in whatever is happening in the virtual world. The third factor is "ecological validity," which is concerned with "believability and realism" and overall consistency between the different sensory streams. This is most closely related to the concept of Psi, but is not the same. The fourth factor is "negative effects" such as discomfort, simulator sickness, and so on. Another comprehensive factor analysis study carried out by (Schubert et al., 2001) led to a similar three component model: spatial presence ("being there"), involvement, and realness. This has come to be known as the Igroup presence questionnaire.

There are fewer examples of questionnaires specifically aimed at copresence. A series of three studies where groups of three remotely located people carried out tasks together in VR, developed one of the first questionnaires. These included questions such: "There was a sense of being with the other people," "The computer interface seemed to vanish and there was direct working with the other people," "Rate how closely your sense of being together with others in a real-world setting resembles your sense of being with them in the virtual room" (Tromp et al., 1998; Steed et al., 1999; Slater et al., 2000). These types of questions were derived simply from the idea of copresence. Drawing on the extensive collaborative virtual environments literature Garau et al. (2001) reported an experiment on the impact of different eye gaze models on quality of communication between two remote individuals represented by avatars via a video tunnel. The quality of communication was assessed by a questionnaire eliciting four aspects: the extent of *face-to-face* communication that was perceived (e.g., "I could readily tell when my partner was listening to me" and five other questions), the degree of *involvement* ("I found it easy to keep track of the conversation," "I felt completely absorbed in the conversation"), *partner evaluation* (e.g., "My partner was friendly," and four other questions), and *copresence* ("I had a real sense of personal contact with my conversation partner," "I was very aware of my conversation partner"). These same questions were used in a later study where participants met via immersive VR (Garau et al., 2003). Bailenson et al. (2005) introduced a measure of copresence using three questions ('Even when the "other" was present, I still felt alone in the virtual room,' "I felt like there was someone else in the room with me," 'I felt like

the "other" was aware of my presence in the room.' Bulu (2012) carried out a study using the (2D) Second Life as the shared virtual environment, and assessed copresence as the sense of being part of the group and each participant's assessment of the copresence of the others. A psychometric scale for copresence was developed by Poeschl and Doering (2015) in the context of a fear of public speaking scenario and in the German language. As can be seen while there is a lot of variability in the questions that researchers have used for copresence, there is the common aspect of simply the sense of being with other people.

There are several problems with questionnaires if used alone. One is that they are typically administered after an experience rather than while the experience is taking place. Schwind et al. (2019) showed that questionnaires can be administered during the actual experience so potentially overcoming this disadvantage. However, contrary results were found by (Graf and Schwind, 2020). Moreover, a problem with this is that it forces participants to take a meta-view of their experience during the very time that it is required to just experience it. A second problem with the use of questionnaires alone is that the questionnaire itself may bring about the very feelings that it is supposed to measure. For example, Slater (2004) carried out a study about an entirely invented concept called "the colourfulness of your day." Participants were asked to think about their previous day and answer various questions including how "colourful" it was. Correlations were found between this and various other factors in the questionnaire, just as correlations are found with the sensation of "being there." The problem of course is that before being introduced to this idea of a day being "colourful" participants almost certainly did not ever think in those terms. Similarly, it is possible that the feeling of "being there" never occurs to participants in a VR experience, but that this is introduced to them solely through the questionnaire. In other words, the researcher's conceptual framework is imposed on the participants.

Behavioural and Physiological Surrogates for Presence

In order to obtain a more objective approach to the measurement of presence (PI), behavioural or physiological surrogates have been used. Going back to our opening example, some participants clearly had a strong response to being at the virtual Dire Straits concert, feeling that they were alone, that people were staring at them, and so on. This indicates a high level of both PI (they are at the place of the concert) and Psi (that this was really happening) and also copresence (with the virtual audience around them)—otherwise, becoming stressed about their situation would have made no sense. Meehan et al. (2002) formalised the idea of using physiological measures of stress as a surrogate for presence to measure these types of response. People were placed in different conditions standing by a precipice and their heart rates were measured, the argument being that increased heart rate would indicate arousal, which is what occurred. Spanlang et al. (2007) carried out an experiment where a fire broke out in a virtual bar, and at first other (virtual) people in the bar ignored it—which resulted in physiological responses such as heart rate, heart rate

variability and skin conductance responses indicating significant arousal. Martens et al. (2019) had one group of participants go up a building in an interior elevator and another group in an external elevator, and found differences in a range of physiological responses, with greater arousal amongst those in the external elevator. Ríos and Pelechano (2020) placed participants in a train station, where suddenly virtual travellers started running in a particular direction, as if there were an emergency. Participants tended to follow the escaping virtual characters. Putze et al. (2020) investigated the relationship between physiological responses and questionnaire administration during and after the VR experience, and the results suggested less disruption to presence for questionnaires administered during the experience, and the result of Graf and Schwind (2020) is also relevant to this.

As well as physiological measures of arousal, brain activation has also been measured with EEG, for example (Baumgartner et al., 2006; Kober et al., 2012; Kober and Neuper, 2012; Clemente et al., 2014; Petukhov et al., 2020). One study suggested that placing participants in a stressful situation in VR can lead to extreme stress as indicated by brain activation responses (Fadeev et al., 2020).

Ochs et al. (2022) introduced a new approach for presence and social presence. During the course of social interaction, involving a doctor delivering bad news to a patient, various multimodal measures were recorded, mainly concerned with speech (e.g., length of sentences). Place presence and copresence were measured using the questionnaires in (Schubert et al., 2001) and (Bailenson et al., 2005) respectively. Then machine learning methods were used to successfully predict the levels of reported presence in each case. Association between subjective reports and other non-subjective measures (physiological, behavioural) is one of the best ways to demonstrate that there is indeed an underlying phenomena, which can be measured consistently in independent ways.

A sure sign that participants in VR exhibit correlates of presence such as emotional, behavioural and physiological changes is that one of the main areas where VR has been used in the past 30 years is in clinical psychology. In these applications VR is typically used as an adjunct to cognitive behavioural therapy, where participants are gradually exposed to the anxiety causing situation. Using VR this exposure can take place in the office of the clinician rather than trying to arrange real situations for exposure or giving the client "homework" that they can report on during the next session. This can only work because presence (both with respect to PI and Psi) operate—so that participants, even though they know that the situation is entirely virtual—nevertheless experience similar anxiety as they would in physical reality. Without this level of anxiety, therapeutic techniques would not work. For reviews see (Rizzo and Kim, 2005; Freeman et al., 2017).

The problem with using behavioural and physiological surrogates and EEG for presence is that this technique can only be used in limited circumstances—in particular for environments that cause measurable arousal (whether with negative or positive affect). In circumstances where there are no predicted arousal effects then this method cannot be used, or only where there are specific triggers that are there solely in order

to elicit physiological or brain responses. So, although behavioural and physiological measures, and especially in combination with questionnaires go some way towards a reliable and objective measure, they cannot provide a general answer.

Breaks in Presence

The third main method for assessing presence is based on the idea of "breaks in presence" (BIP). This has its basis in our approach to understanding what presence is and its relationship to sensorimotor contingencies. The argument is that the normal state is that participants will have the sensation of PI in a virtual environment which has at least 6 degrees of freedom head-tracking (whether a Cave or wide field-of-view stereo HMD). However, every so often there may be a glitch, such as a sudden change in frame rate, or an event that breaks consistency or expectation, or even something simple such as bumping into a (real) wall. If we think of PI as essentially "always on" but with occasional failures, then the number of failures and their occurrences through time would provide an interesting measure of presence. This was first put forward in (Slater and Steed, 2000). In this framework we can think of the responses to a post-experience questionnaire as the integral over time of these continual periods of "on" instead of "off" so that the number of BIPs should correlate inversely with post-experience presence questionnaire scores. The problem with BIPs, however, is to know when they occur. In the original paper a BIP was signified by the participant, the argument being that if a BIP had already occurred then reporting that fact could not in itself diminish presence, because it had already been diminished. Another method has been to try to find physiological correlates for BIPs. Slater et al. (2003) studied how breaks in presence were signified by patterns in physiological responses, specifically ECG and skin conductance. Rey et al. (2011) found a relationship between blood flow velocity responses and breaks in presence forced through deliberately caused glitches in the VR. Liebold et al. (2017) viewed BIPs as orienting responses towards the virtual or real world and differentiated between different types of BIPs on this theoretical basis. They found associations between various types of BIP and physiological responses.

Breaks in presence have the advantage over questionnaires that they are intrinsically based on what is experienced during, rather than evaluations after, the VR exposure. Whether a questionnaire is posed during or after an experience it inevitably forces the participant to assess the experience in terms of phrases that are imposed from the outside. A BIP can be just a sudden failure of whatever it is that the participant takes as their presence in their separate reality induced by the VR, and from the point of view of the participant can be considered as conceptually neutral: it broke, we do not really need to know what they mean by that, it is just an observed fact. The problem is in the observation. Do we allow participants to self-report BIPs? perhaps they might forget to report an occurrence in the excitement of the moment, or maybe they feel impelled to report *something* because this is the expectation imposed on them. If we do not allow participants to self-report then some other means needs to be introduced—for example, physiological responses that indicate a BIP. In this case there would therefore be a double layer of assessment—first, whether the physiological measures really indicate a BIP with all the problems of false negatives or positives, and second, instead of directly assessing whether a BIP has occurred it is now an indirect measure. It is more cumbersome, with the high probability that an additional layer of technology involves an additional layer of error. Although relatively neutral compared to a questionnaire, and while the BIP idea is conceptually interesting, it still has potential problems (Skarbez et al., 2020).

Configuration Transitions

Why are measures of presence (whether PI or Psi) useful at all? There are at least two reasons. First, VR uniquely delivers PI (in the sense meant in this paper), and if a system or application does not at least result in that, then there seems to be no point to using VR. Second, generally application designers need some criteria against which to understand trade-offs between different factors that could be included in their designs. Understanding how different trade-offs may influence PI and Psi therefore requires some way of measuring these. To reiterate our view: there is immersion, i.e., the objective capabilities of the system. Presence is a subjective response. The goal is to understand how varying these different immersive capabilities influence presence.

The situation has some parallels with colour perception. There is physics, light being emitted or reflected at different wavelengths into the eye. This is the objective aspect. Then there is the human perceptual response where the wavelength distributions are interpreted as colours. However, the situation is not so straightforward. There is not a 1-1 mapping between wavelength distributions and perception of colour. Different wavelength distributions can result in the same colour perception (metamers) and different people may interpret the same wavelength distribution quite differently. The human visual system builds colour perception from three types of cones in the retina (so-called "red," "green," and "blue"), yet the wavelength distributions are infinite. Colours are quantified, however, through colour matching experiments. For example, a target colour is projected onto a screen and an experimental subject has control over three projectors emitting red, green or blue light, combined onto a different patch. The subject can manipulate the intensities of the three projectors, so mixing red, green and blue in different proportions. The subject's task is to find the intensities of the red, green and blue that result in a colour that perceptually matches the target. By repeating this over a number of people, each colour of interest can be associated with a particular combination of proportions of red, green and blue intensity. Hence there are no questionnaires to elicit a colour (e.g., "How red is this on a scale of 1–7?") but only the colour matching process. This does not require any knowledge of how the person "really sees" the colour only that they match it to the target colour.

In (Slater et al., 2010a) we applied a method analogous to this for the assessment of PI or Psi. Corresponding to wavelength distributions there were factors that could be set at different levels. In this first experiment the factors were: 1) the type of illumination model used in the scenario (Gouraud shading, static view-independent global illumination, dynamic global

illumination including real-time shadows and reflections) 2) the field of view (small or large); 3) display type (a simulated power wall or a head-mounted display) 4) the self-representation of the participant (without a virtual body, with a static virtual body, with a virtual body that moved with the participant's own movements based on real-time motion capture). Participants first experienced the scenario with all these factors set at the highest level (dynamic global illumination, large field of view; head-mounted display, full body representation with real-time movements). They were asked to pay attention either to their sense of being there (PI) or that the situation was really occurring (Psi). Then they had a training session that introduced them to the four factors and how they could change levels within each factor (always in the direction of lowest to highest). After this training period they were placed back in the scenario but now with all the factors set at a low level. At specific times they could choose to increase the level of one of the factors. They were asked to continue to do this until they felt that they had achieved the same sensation of PI or Psi as in their first exposure. The changes that they could make were under cost constraints.

We refer to any set of factor levels as a *configuration*. In this experiment there were therefore $3 \times 2 \times 2 \times 3 = 36$ configurations. Each time that a participant chooses to change the level of a factor there is a transition from one configuration to another. Taking the entire set of transitions, it is possible to estimate a transition probability matrix P with entries p_{ij}, the probability of transitioning to configuration j given that the current configuration is i, $\sum_j p_{ij} = 1$. Now there are several ways that P can be used. From Markov Chain theory it is simple to calculate the long run equilibrium probabilities of the system from any starting point. These specify for each configuration the probability that in the long run this would be the one into which the system settles down (i.e., no more transitions would occur). Second, we can trace the order of transitions—i.e., which are those that are made first to move towards the matching state, which are made second, and so on, and hence we can obtain information about the factor levels that most contribute towards the desired state of PI or Psi. Third we can compute probabilities for the individual factors: for example, taking all the final configurations reached, to compute the probability that a particular factor level would be contained in the final configuration. There are many other possibilities too. In addition this method can result in equivalence classes—i.e., sets of configurations that result in similar probabilities of PI or Psi. Such equivalence classes were argued for long ago by Ellis (1991), and are useful for engineers since they can use these to trade between different configurations, for example, based on a cost function.

What is important to understand is that participants are never asked their opinions or asked to give rating scales. The method is premised on observable events only—the fact of the transitions. The method is entirely based on the participants' decisions that a particular configuration results in a match of their sensation of PI or Psi, i.e., the only important thing are the matches that people make, and not the meanings that might be attributed to these. In the experimental study of this method participants who had been asked to focus on matching their state of PI transitioned between the configurations differently than those who had been asked to match Psi, and the final resulting configurations were different.

This method has been applied several times. Azevedo et al. (2014) combined the method with EEG measures of engagement. Bergström et al. (2017) used this method to evaluate Psi for four different factors influencing the responses to a virtual string quartet performance. Skarbez et al. (2017) evaluated how a number of factors determining the behaviour of virtual characters influenced Psi. Gao et al. (2018) assessed the believability of a VR rock climbing scenario. Fribourg et al. (2020) turned the method towards evaluation of body ownership rather than PI or Psi. GalvanDebarba et al. (2020) studied Psi in relation to body animation features. Clearly the same method could be used for the evaluation of different subjective responses, not only PI of Psi, which we turn to next.

Beyond Presence

Murcia-López et al. (2020) used the configurations transitions method to evaluate the impact of several factors controlling an avatar giving a TED-style talk to participants. There were two main differences from the previous application of the method. First, participants could choose to change a factor level in any direction, so that there was no assumption about levels being ordered, that one factor level was somehow superior to another. The second main change was that the goal was not to improve presence but simply to maximise *preference*. In other words, participants could decide to make a transition from configuration A to B only because they preferred B to A, irrespective of the reason. A consequence of these changes is that the method did not require participants to first experience the optimal configuration, since there was no *a priori* assumption about which configuration would be the best. The method produced coherent results across participants. This fits the idea mentioned above of not imposing criteria of researchers, such as presence, on participants.

This approach has been taken still further by Llobera et al. (2021). Instead of participants themselves choosing how to change the configuration, i.e., choosing which factor to change to which level, an AI agent does this. In this setup a Reinforcement Learning (RL) agent occasionally proposes to the participant a change of level of one of the factors in the current configuration. The participant can reject the proposed change, leaving the current configuration as the one in force, or can accept the change. Over time the RL agent learns which changes are likely to be accepted or rejected, i.e., it forms a policy consisting of probabilities of acceptance of proposed configuration changes given the current configuration. The process stops when the participant has reached a stable state and continues to reject further changes. The only criterion for participants to accept or reject a change is what they prefer. So over time the RL agent finds an optimum (though possibly only a local optimum rather than global) with respect to which configuration is most preferred. In the experiment reported by Llobera et al. (2021) this method was used for four binary factors. The RL was applied to each individual separately rather than cumulatively across all individuals. The optimal solutions happened to conform to what we would expect from previous

research on presence, though to emphasise, the issue of presence (PI or Psi) was never mentioned to participants, only that they should choose according to their preferences.

Sentiment Analysis

A problem with the configuration transitions method is that although it is based either on the fact of a "match" between a target state and an actual state with respect to presence, or simply based on preference, it does not obtain any deeper insight into the reasons for people's choices. In the opening paragraph of this paper, we described the rock concert scenario. This was the first time that we had created such a scenario, one quite unusual for VR. We did not want to impose our notions of presence on participants in the experimental study, so we simply asked them to write a short essay about their responses to it (Beacco et al., 2021). We then applied sentiment analysis (Liu, 2012; Bakshi et al., 2016) to the resulting essays. Sentiment analysis is a machine learning technique that will score text for positive or negative sentiment based on a pre-trained word-to-vector data base of millions of pre-classified words (Mikolov et al., 2017). Hence each essay carried a sentiment score, and the set of scores fell into distinct clusters (from a very low sentiment cluster to a very high sentiment one). The contents of essays of each cluster were then analysed to determine common themes within that cluster. For example, a common theme was "realism," where sentiment scores were higher the more that participants mentioned something about the environment that seemed to be very real. Another theme was "disturbing," where people found the crowd around them to be disturbing in some way, for example, one woman mentioning that she thought two people (of course virtual characters) were acting as predators towards her, by staring at her and then looking away whenever she looked back. Another concerned "failure of expectations" for example, the drummer visually not in sync with sounds of the drumbeats, or sounds of clapping but no one around actually visually clapping. In this way we were able to obtain very deep insight into participant responses to the concert and discover reactions that we would never have discovered through a presence questionnaire, BIPs or configuration transitions used alone. We found that the concert was highly plausible, but for some participants, not at all in the way that we expected. For some (not all) of the participants the concert was a "nightmare." Being a nightmare of course demonstrated a high level of plausibility—since if people had not had the illusion that the events were really happening, then there could be no reason to become disturbed. Similar results were found in (Slater et al., 2022). It should be noted that as with any method the quality of the results depends on the quality of the input. Participants might be reluctant to write even a short essay after experiencing a VR scenario, or what they write might be too short for analysis. Recently we have started recording what participants say in a post experiment interview, rather than requiring them to write.

OPEN QUESTIONS

PI and Sensorimotor Contingencies

PI is the illusion of being in a place. In physical reality we are always in a place. In VR we may have the illusion of being in a different place, with the corollary that we know this is not true (and this knowing it is not true is itself part of the feeling). In reality we might be completely uninterested in what is going on in the physical space in which we are located, but this will not destroy the sense of being there. Our minds might wander, and we start thinking about other things, but this does not change our perception of where we are. Now suppose we are in a VR that renders this same boring place, and similarly we become uninterested and not engaged in what is happening. This actually is not incompatible with the illusion of "being there" because it leads to the same responses as if we were really there. Hence questionnaires that include categories such as "engagement" or "involvement" are not assessing PI, even though in some applications the degree of engagement may be an important, but separate issue. Similarly a relationship between presence and task performance has been posited, with the argument being that greater presence enhances task performance, a debate that goes back a long way (Welch, 1999). Now in physical reality we might be trying to draw money from a bank teller machine (ATM) but not be successful because of the machine's poor user interface. Carrying out a similar operation in VR might also lead to failure for the same reason. But here the failure would not be incompatible with presence—our poor task performance in reality should map to a poor task performance in virtual reality. The user interface is the issue, not PI.

In our view PI is intimately bound up with sensorimotor contingencies for perception. If we carry out bodily actions for perception (turning the head, moving the eyes, turning around, bending down, stretching up, looking around, looking over, looking under, turning our head to hear a sound better, touching, pushing, smelling) and the multisensory displays deliver integrated sensory outputs that correspond to those that would occur with similar actions in reality, then the simplest hypothesis for the brain to adopt is that what we see, hear, feel, . . . signifies where we are.

In our whole lives whenever we turn our head the visual images that we see change in a predictable way. This produces a massive probability for our sense of place: the probability that I am in the place that I see, hear and can touch. This is why, for example, if we are looking at a virtual environment displayed on a screen in a desktop system in front of us, we cannot experience PI in the sense meant in this paper. We only have to turn our head away from the display and we see a different reality, that of the real location in which we actually are carrying out this activity. In (Slater, 2009) we argued that this leads to a partial order over systems. In an immersive virtual reality, for example, realised through a high quality HMD, it is straightforward to simulate the activity of looking at an external screen that portrays a virtual environment. We referred to this as the HMD system being more "immersive" than the desktop system, since the first can be used to simulate experiences of the second, but not vice versa. Here immersion refers to objective capabilities of the system and not the subjective responses. Therefore, it was argued that the qualia associated with PI in the more immersive system cannot possibly be the same as in the less immersive system. In the desktop system one might attain some sensation of "being there" through employing additional attentional resources and imagination,

but in the HMD system it is just a fact—a head turn does not shift your perception out to the real world but is confined within the virtual separate reality. The qualia associated with the HMD and desktop system are at different levels, they are not comparable since they do not refer to the same underlying situation.

In this view there is no "book problem" (Biocca, 2002; 2003) regarding presence. When reading a book we might imagine ourselves into the story line, imagine being in the place depicted in the book. But this is not PI, the perceptual illusion of being in a place, as evidenced by the fact that you can carry out different acts of perception and the sensory stream remains consistent with being in that place. In VR it is simple to program the appearance of a book and allow people to read it while in VR. While they are reading the book they are in a virtual place, at any moment they can glance away from the book and they are still in that virtual place. In the book they might read about a fire breaking out in a train station with all the passengers running away, but they will not run away from that. In VR they might run (Ríos and Pelechano, 2020). The qualia associated with reading a book compared with sensorily being located in a virtual space are quite different. Of course one may choose to call the qualia associated with reading a book "presence," but this is not what we mean.

If PI is so bound up with sensorimotor contingencies (SCs) for perception, why separate these two concepts? This is because SCs can fail, or not be consistent with one another, or delivered with too great latency, and for many other reasons. They are not "perfect" SCs corresponding exactly to real perception. Moreover, people may behave quite differently to one another. If person A stands and simply looks around an environment but B actively explores it looking close up at objects, trying to touch things, moving rapidly, then B might experience quite a different level of PI than A. This also shows that any measurements of PI need to implicitly take account of possible individual differences. Another way to think about this is to consider the level of PI as a random variable even if, as could be argued, that its possible values are restricted to 0 (no PI) and 1 (PI). In this view the expected value of PI (between 0 and 1) would be a function of the sensorimotor contingencies. However, any actually observed value would vary around its expected value, and one of the contributors to departure from the mean would be individual differences. In this approach individual differences could also be modelled as contributing to the variation.

Place Illusion and Plausibility

The study of presence has been a major part of VR research since the early 1990s. We have argued for the view that presence consists of two orthogonal dimensions, Place Illusion and Plausibility. These are logically separable concepts in that it is possible for PI to occur independently of Psi and vice versa. For example, you can have a strong sense of being in a place, but the virtual characters there may have no plausibility—they do not respond to your actions, nor initiate any actions towards you (Garau et al., 2004). On the other hand, you can be interacting with a virtual character on a non-immersive device (e.g., a normal screen) not have any illusion of being in the place, but totally accepting that this is really happening—the person seems to be real, responds to you and initiates interactions with you (and this

happens every day in teleconference video calls). It is really happening but you are not there in the space where the other person is depicted. Of course, despite this logical separability between PI and Psi we might find them to be empirically related, but largely because of the great difficulty (or even impossibility) of having measures that are so precise that they can perfectly distinguish between the two. This is especially the case with questionnaires.

The concept of Plausibility arose out of the fire in the bar experiment discussed above (Spanlang et al., 2007). This was carried out in a Cave, had the best quality rendering that we had ever used, and yet the questionnaire scores on presence were one of the lowest we had seen. It was realised that this was because of failure of one component—repetitive motions of an avatar. Participants only had one way to signify this failure—by giving low scores on a presence questionnaire. So here the measurement of PI was confounded by Psi.

For future research Psi is by far the more important and interesting issue. As mentioned above our view is that in a typical VR with at least head-tracked stereo wide field-of-view HMD that PI will be the default sensation, since most of these systems offer a level of sensorimotor contingencies that are good enough to fool the brain into the illusion of being in the place depicted by the VR to our senses. The major research question remains to be able to understand in greater detail exactly which sensorimotor contingencies are critical, and under which conditions. For example, suppose that the HMD does not have 6 degrees of freedom tracking but only 3 degrees of freedom, so that there are no parallax effects. If participants do not move their heads, or only rotate them without translation, then the 3 degrees of freedom may be good enough. However, for example, in a simulation of a game such as table tennis where players are moving from side to side and back and forth very quickly all the time, the loss of the 3 degrees of freedom would severely diminish the experience (and probably result in simulator sickness). The fundamental research question for PI is therefore the map from sensorimotor contingencies to Place Illusion, irrespective of how this might be measured. A second question is how to model the impact of individual differences.

Although PI might be the default state, low Plausibility might also be default. PI is just part of what VR is (even though it can fail). However, Psi requires deliberate design: how can a scenario be designed and implemented so that participants buy into it? This does not necessarily involve photorealism, nor realism in the sense that the scenario might not be anything that can happen in real life.

Although we have said that PI and Psi can be thought of as conceptual frameworks imposed on participants by researchers, there are some applications where they are nevertheless crucial. For example, consider a therapeutic application for an anxiety disorder such as fear of crowds, one that employs an exposure therapy-based approach, where participants are gradually exposed to the situation that causes anxiety. So first the participant might be in a VR with one other (virtual) person, then two, then four, then ten, and so on, until over time the participant learns to control their anxiety while being in a large crowd. PI is needed because these events have to take place

somewhere, and if the participant does not have the illusion to actually be there, then there is little chance of anxiety being provoked. Second, the crowd should behave like a crowd, for example, it should part as the participant tries to move through it, and virtual crowd members should occasionally look towards the participant or acknowledge them in some way such as a wave or a smile or say "Excuse me." Hence Psi is also critical to the success of this application. This illustrates the point above—that although PI might occur by default, Psi has to be deliberately designed. A corollary of this is that if PI fails (a BIP) it is likely to re-form again, since the sensorimotor conditions that gave rise to it are still there, even if they had temporarily failed in some way. However, if Psi fails it does not re-form. For example, once participants realise that a virtual human character is unaware of their presence, they just lose interest and move on (Garau et al., 2004). Although all this follows from the concepts of PI and Psi there is little empirical evidence available to date. In our view the major part of the research interest should be a focus on Psi—under what conditions do people take events and situations in the VR as actually happening, and when does this fail?

Coherence

Skarbez et al. (2020) introduced the notion of coherence, which is the extent to which a virtual environment "behaves in a reasonable or predictable way." More fully: coherence is defined ... as the set of reasonable circumstances that can be demonstrated by the scenario without introducing unreasonable circumstances, where a reasonable circumstance is a state of affairs in a virtual scenario that is self-evident given prior knowledge" (Skarbez et al., 2018b). In this view coherence stands in the same relation to Psi as immersion (the actual objective affordances of the system) stand to PI. However, this is not straightforward. While "immersion" is something objective and describable independently of any effect that it might cause (e.g., the field-of-view is x, it supports stereo vision with adjustable interpupillary distance, the resolution is y, the colour resolution is z, the tracking capabilities are w, and so on) behaviour in a "reasonable or predictable way" is not something that is objectively describable—it includes evaluations by people. One person may feel that a scenario is reasonable and events predictable, but another may feel that this is not the case. What is self-evident to one may not be to another. We have proposed that the three factors that contribute to plausibility are: 1) the reactivity of the environment to participant actions, 2) contingent references by elements of the environment to the participant, and 3) credibility of expectation—i.e., the environment is constructed based on evidence of what is supposed to happen in real life where this is relevant, so that the application is supposed to be a simulation of events that occur in reality. These are a question of the extent to which these elements are *supported by the hardware and have been programmed*. The *satisfaction of expectations* or the *realisation of environment being coherent* in the sense meant in (Skarbez et al., 2018b; Skarbez et al., 2020) cannot be programmed, since these are the responses of the participant. What is possible though is to carry out research, for example,

using co-design methods in order to attempt to meet these requirements. Another way to think about this is that Psi may fail if the environment never responds to the participant, or if nothing in the environment ever addresses the participant personally, or if there is something that fails to meet expectations. All of these are a function of the hardware and programming and the design of the application. To the extent that coherence might also be regarded in the same way—i.e., there is an attempt to satisfy coherence but it fails—then coherence can be thought of as a way of referring to requirement (3).

This discussion emphasises our distinction between "immersion" and "presence." The "immersion" is fully under the control of the implementation. Whether a virtual character smiles back at you when you smile towards it, will have to have been programmed. The extent to which a virtual environment can meet expectations depends on prior research amongst potential participants about the critical elements that a scenario must support, and then actually programming these. Whether this investigation has been carried out or not, and then the required elements programmed into the application is a matter of fact. Then "presence" (PI and Psi) refers to how people respond to the "immersion." As mentioned above this is not deterministic, since people have different prior experience, personalities, knowledge, and so on. We can think PI and Psi as conditional probabilities, and consider: conditional on a particular immersive configuration what are the probabilities of PI or Psi occurring? Given different immersive configurations these probabilities may change. Thus immersion sets the ground for PI and Psi. Similarly we use the terms "embodiment" to refer to the multisensory factors that provide evidence about the body—for example, it is seen as life-sized from the perspective of the eyes of that body, it moves synchronously with the person's real movements, when something touches it the person feels this on their real body, and so on. Embodiment configurations, which are completely determined by the hardware and programming, may give rise to the illusion of body ownership. Again, this is not deterministic but provides the basis. For example, ballet dancers, always acutely aware of the exact disposition of their bodies might have reduced body ownership in VR because they are more likely to notice small discrepancies between the position and movement of virtual limbs compared to the true positions and movement.

Latoschik and Wienrich (2021) introduced a new model for Plausibility. However, the term "Plausibility" in this model does not have the same meaning as considered here, and it is confusing that the same term was used to describe something quite different. Latoschik and Wienrich (2021) write that ... in contrast to the discussed presence models, we don't assume an illusion of plausibility but define plausibility as a state or condition during an XR experience that subjectively results from the evaluation of any information processed by the sensory, perceptual, and cognitive layers" (p5-6). In their model each of *cognition* ("social-cognitive processes" and "higher order cues"), *perception* ("proximal perception experiences" and "proximal perception cues") and *sensation* ("genetics or life-long habitude perceptions" and "habitual sensory cues") contribute to a level of coherence (as defined by Skarbez et al.). Then Plausibility is a weighted function of these

TABLE 1 | Summary of types of measurement and their properties.

Method	Type of data	Applied	Analysis	Advantages/disadvantages
Questionnaire	Subjective responses on an ordinal scale	Typically after the experience, though it is possible to include questionnaires as part of the experience, with the danger of forcing participants to a meta level (thinking about rather than experiencing)	Frequency tables, median, interquartile range, ordered logistic or probit regression	Very simple, can be applied to any concept. Universal—can be applied to any scenario. Surface results only, lack of intersubjective equivalence in meaning of scores, imposes researcher concepts on participants.
Behavioural and Physiological	Observed behaviours, e.g., based on tracking, eye movements, heart rate, skin conductance, respiration, etc	During the experience	Means and standard deviations, ANOVA, regression, time series	The observed values are based on what participants do, or on their automatic responses and not on their opinions. The measurement process is more difficult, especially with physiological measures. Problems such as drift in skin conductance responses and artefacts—e.g. in ECG data—have to be overcome. Interpretation of results is open to question. The VR scenario requires events or situations that deliberately spark the behaviour or physiological responses which may not be appropriate. It is therefore not universal.
Breaks	Discrete times when breaks occurred, count of the number of breaks	During the experience	Rate of breaks per unit time, times between breaks, survival analysis, mean, standard deviation, Poisson regression or equivalent	Based on observational data rather than subjective. A break occurring is an observable event. It is universal—can be applied to any scenario. It is difficult to ascertain when breaks occur. The subjective element clearly applies when participants are asked to self-report breaks.
Configuration transitions	Transitions between configurations and their timing, and when matches to a specific quale (e.g., Place Illusion) are reported (depending on the particular method)	During the experience	Transition probability matrix, Markov Chain analysis, conditional probabilities	Based on observed data—the actual transitions, and matches (if used in the method). The method can be combined with Reinforcement Learning. Applicable when the goal is to find the optimal configuration for a number of factors. Very large data sets may be required for transition probability matrix estimation when the number of factors is above 4 and the number of levels of factors are above 3.
Sentiment analysis	Freely written or spoken verbal reports, resulting in quantitative sentiment scores	After the experience. Talking aloud during the experience is possible but with the danger of forcing participants to a meta level	Mean, standard deviation, regression, ANOVA, cluster analysis for clusters of sentiments, keyword analysis to find commonalities amongst clusters	The analysis gets beneath the surface of questionnaire or behavioural and physiological responses to allow participants to freely reflect on their responses to the VR experience, yet nevertheless resulting in quantitative measures. The method only applies to sentiment, although researchers are able to create their own dictionaries that concentrate on other aspects of an experience (e.g., presence). Requires participants to write essays or speak about their experiences at enough length to make the analysis feasible.

three aspects of coherence. Plausibility in turn contributes to several different qualia: presence, social presence, copresence, placeness, body ownership—Figure 2 in Latoschik and Wienrich (2021). The authors state that "In our opinion, the proposed model possesses predictive and explanatory power of modern XR experiences."

Brubach et al. (2022) tested this theory with two experiments where participants interacted with some cubes. In experiment 1 ($n = 40$) there were two binary factors: manipulation of the perceptual layer (within group): cubes followed expected physical laws or they floated. The second factor was a cognitive manipulation (between groups): participants were told that they would be weightless, or nothing was said about this. Hence the conditions (floating cubes, weightlessness) and (cubes follow physical laws, no mention of weightlessness) were compatible with coherence, and the other two possibly with incoherence. A questionnaire was used to assess plausibility, which included questions such as "I am used to objects behaving this way," "I had a prior expectation of how the objects would behave," "the behaviour of the objects made sense," **Table 1** in (Brubach et al., 2022). An assessment of presence was based on the Igroup questionnaire (Schubert et al., 2001). The results showed that each individual item on their Plausibility scale had higher scores for the normal cube behaviour on the perceptual factor. However, there were no significant results for the cognitive factor (i.e., the explanation about weightlessness had no effect), and no interaction effects. Presence as measured by the Igroup questionnaire showed a significantly higher mean for "realism" for the cognitive factor "weightlessness" level, whereas "involvement" had a higher mean for the perceptual "floating cubes" level. Participants were asked to choose which perceptual level (the within group factor) was more plausible, and 39/40 chose the normal cube behaviour.

The second experiment ($n = 71$) was on the same lines, but with a much richer story line for the cognitive factor. Similar results were found for the Plausibility questions, but no significant differences for the Igroup presence questionnaire (except in overall scores). The authors concluded that both experiments together showed that "The manipulation of object behavior on the perceptual layer leads to a break in plausibility." However, the hypothesis "Breaks in plausibility lead to break in presence" could only be partially accepted (based on experiment 1), and no conclusion could be reached on whether "Perceived coherence between the cognitive and perceptual manipulation has an effect on presence." Overall the authors argued that the two experiments provide empirical evidence in favour of their model described above.

There are several points to be made about this model and experiment. First there is the unfortunate confusion of terminology, as mentioned. A definition of "plausibility" is given that has little relationship with the original one proposed in (Slater, 2009). However, considering the operationalisation of this via the questionnaire about plausibility, we can see that in this model plausibility refers to evaluations of object behaviour—how real this seems and how much it conforms with expectations. This is not the original meaning of Psi, which is not an evaluation of conformity to

realism but refers to the illusion that *events are actually happening* (even though cognitively the participant knows that nothing like those events are really happening). Psi is like an exclamation "This is really happening!" The cube falls towards your toes and you jump out of the way, an avatar smiles at you and you smile back—these are automatic reactions that follow from the possibly implicit exclamation "This is really happening!" At some level the brain does not know about virtual reality. It responds and acts based on its sensory surroundings. The point is that *there might be a falling cube!*, so the safest action to take is to get out of the way. In our whole lives when we have seen something falling towards our toes, there really was something falling towards our toes, so the safe thing to do is to act on that.

The second problem is that if the idea of plausibility as exhibited in the questionnaire is followed, then applications that depict fantasy worlds, or events that cannot happen in reality, can never be plausible. People adopt different models of reality in VR though, they can fly and be a superhero (Rosenberg et al., 2013), they can see their own virtual body separate itself from their location and carry out some actions (Gorisse et al., 2021) and have an out-of-body experience (Bourdin et al., 2016), or even a simulated near-death experience (Barberia et al., 2018). If these applications were judged on the basis of the plausibility questionnaire adopted in (Brubach et al., 2022), then none of them would be plausible, since events occurred that clearly broke physical laws. Yet each of these examples were effective in leading to changes in participant attitudes and behaviours.

The third issue is the complexity of the model, with three layers, several elements in each layer, and five resulting qualia. Moreover, the idea that "Plausibility emerges from a function of weighted congruence relations (activations)" does not offer a way forward—how could this ever be tested, or rejected? The approach we have proposed is much simpler. On one side is everything relevant that goes into the making of an application that can be objectively described—the hardware (e.g., computational, display and tracking), the programming (which determines the events that occur, the affordances and the valid actions that participants can carry out). On the other side are the responses—PI, Psi and body ownership. This relationship can even be expressed as conditional probabilities: $P(Psi \mid an\ immersive\ configuration)$, and this has been done, for example in (Slater et al., 2010a; Bergström et al., 2017; Skarbez et al., 2017). This model could also include psychological assessments of participants to account for individual differences.

However, there is a different question that concerns the mechanisms that might explain how a particular immersive configuration becomes translated into PI, Psi or body ownership. The approach of Latoschik and Wienrich (2021) could be regarded as a step on the road in that direction.

With respect to Psi, there remains a huge amount of research to do in order to investigate what the illusion that events are "really happening" means and can be assessed, how this might vary from application to application, what types of contingent reference to the participant are important, how faithfully should the environment conform to expectations when it is a simulation

of events that could happen in real life. The issue of how to maximise Plausibility is a massive research area, that probably would benefit from input from theatre and film studies.

Measurement

As we have seen multiple methods of measurement have been proposed and used. These are summarised in **Table 1**. Questionnaires have the advantage that they are relatively simple, and may help answer specific questions, but they are not neutral. If we want to find out what happens during an experience from the point of view of the participants, questionnaires are not ideal—they impose a conceptual framework which might have nothing to do with the actual experiences of participants during the exposure. Physiological and behavioural measures including EEG as surrogates for presence are useful and are more objective measures, but they are not universal. They require environments where there are events that are specifically included to induce automatic responses, or where these responses are well-understood and predictable and would only occur if there were presence. BIPs provide another neutral way to measure presence, and although interesting theoretically, they suffer from the problem of knowing when they occur. The configuration transitions method provides a psychophysical approach to quantifying presence, but this method does not easily give in-depth information about the underlying reasons for transitions. Using sentiment analysis does provide quantitative and qualitative information about responses to a scenario. However, questionnaires are going to be continued to be used because of their simplicity and universality (applicable to any concept and any scenario). Good evidence that the questionnaire responses are meaningful is when their results correlate with a completely different behavioural or physiological measure. For example, the experiment with medical trainees and doctors carried out by Pan et al. (2016) examined the extent to which they would prescribe antibiotics to (virtual) patients, even when the indications suggested a virus, only because the patients vociferously demanded the antibiotics. It was found that whereas almost all the trainees inappropriately prescribed antibiotics, the doctors were less likely to prescribe the greater their reported levels of PI and Psi elicited from questionnaires. In other words the more that they experienced the situation as really happening, the more appropriate their behaviour. The introduction by Meehan et al. (2002) of physiological measures for a stressful virtual environment also included a questionnaire for presence, and the results supported each other. Recently, Archer et al. (2022) used a questionnaire and physiological measures to assess presence in response to odours in the virtual environment, and found correlated results from the two measures. For further discussion of the relationship between questionnaires and physiological measures see (Grassini and Laumann, 2020). We propose that ideal measurements should involve this type of triangulation between several completely different approaches, such as a combination of configuration transitions and qualitative approaches, including sentiment

analysis, and where questionnaires are used then further backup with behavioural or physiological measures where these make sense in the context.

PI and Psi in Augmented Reality

How does this discussion relate to Augmented Reality (AR) where the participant sees the physical surrounding into which are projected virtual entities? PI becomes inverted. VR aims to place the participant into the virtual world. In AR the problem is to incorporate virtual objects into the real world. Although the problem is inverted the solution is the same: the extent to which the AR system supports sensorimotor contingencies for perception of virtual entities. To bring a virtual object into the physical world participants must be able to perceive it by using their bodies as they would objects that are really there. Hence, they must be able to look around it, see it from different orientations, ideally reach out and touch it. It is more complex than in VR since in VR everything is under program control. But in AR virtual objects should reflect light from the real world as well as virtual light, should influence illumination of the real world, should cast shadows on real objects as well as virtual objects, should visually obscure objects that are behind them from the perspective of the participant, and real objects should obscure virtual objects that are behind them, and so on. Although this is much more technically demanding than in VR, the fundamental issue is the same. Sensorimotor contingencies for perception will bring virtual objects into the real world just as in VR SCs place the participant inside the virtual world. Some of the complexities involved in this have recently been discussed by Regenbrecht and Schubert (2021) who have provided a questionnaire for presence in AR and discussed the reasoning behind it. With respect to Psi, however, there is no fundament conceptual difference from VR. Plausibility in AR is based on the same principles and causes as in VR—virtual entities should respond to actions of participants, entities should be able to initiate personal interactions, and—most difficult of all—there should be consistency between real and virtual objects and coherence as discussed earlier.

CONCLUSION

In this paper we have reviewed the concepts of Place Illusion and Plausibility and argued that if these are achieved, then we would have created a separate reality, one parallel with the real world in which the whole experience is taking place. We have also pointed out that PI and Psi are concepts essentially imposed on participants, and while critical in some applications, it is also useful to have methods that rely on the actual experiences of participants rather than what we think those experiences should be. We have reviewed a number of methods for measurement, and found problems with all of them, and argued that a combination of a psychophysical method such as configuration transitions with qualitative methods including sentiment analysis would be the way

forward, but in any case triangulation through multiple approaches would be ideal. Some of these issues have been studied now for about three decades. With the entry of VR (and to a lesser extent AR) into the consumer market, there is the opportunity for conducting very large studies with participants from all walks of life (not just the typical student subjects), with experiences in their own homes (Steed et al., 2016; Mottelson et al., 2021). It also becomes even more pressing to address the vast amount of confusion there is in the literature. Our philosophy is to keep it simple: be there, experience what is going on as really happening, and therefore respond realistically. Additionally, allow participants to express themselves leading to researchers discovering new ways of thinking about the effects of VR on people.

REFERENCES

Archer, N. S., Bluff, A., Eddy, A., Nikhil, C. K., Hazell, N., Frank, D., et al. (2022). Odour Enhances the Sense of Presence in a Virtual Reality Environment. *Plos one* 17, e0265039. doi:10.1371/journal.pone.0265039

Azevedo, A. S., Jorge, J., and Campos, P. (2014). Combining EEG Data with Place and Plausibility Responses as an Approach to Measuring Presence in Outdoor Virtual Environments. *PRESENCE Teleoperators Virtual Environ.* 23, 354–368. doi:10.1162/PRES_a_00205

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal Distance in Immersive Virtual Environments. *Pers. Soc. Psychol. Bull.* 29, 819–833. doi:10.1177/0146167203029007002

Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., and Blascovich, J. (2005). The Independent and Interactive Effects of Embodied-Agent Appearance and Behavior on Self-Report, Cognitive, and Behavioral Markers of Copresence in Immersive Virtual Environments. *Presence Teleoperators Virtual Environ.* 14, 379–393. doi:10.1162/105474605774785235

Bakshi, R. K., Kaur, N., Kaur, R., and Kaur, G. (2016). Opinion Mining and Sentiment Analysis. In 3rd International Conference on Computing for Sustainable Global Development. IEEE, 452–455.

Banakou, D., Beacco, A., Neyret, S., Blasco-Oliver, M., Seinfeld, S., and Slater, M. (2020). Virtual Body Ownership and its Consequences for Implicit Racial Bias Are Dependent on Social Context. *R. Soc. Open Sci.* 7, 201848. doi:10.1098/rsos.201848

Banakou, D., Hanumanthu, P. D., and Slater, M. (2016). Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias. *Front. Hum. Neurosci.* 10, 601. doi:10.3389/fnhum.2016.00601

Banakou, D., and Slater, M. (2014). Body Ownership Causes Illusory Self-Attribution of Speaking and Influences Subsequent Real Speaking. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17678–17683. doi:10.1073/pnas.1414936111

Barberia, I., Oliva, R., Bourdin, P., and Slater, M. (2018). Virtual Mortality and Near-Death Experience after a Prolonged Exposure in a Shared Virtual Reality May Lead to Positive Life-Attitude Changes. *PloS one* 13, e0203358. doi:10.1371/journal.pone.0203358

Baumgartner, T., Valko, L., Esslen, M., and Jäncke, L. (2006). Neural Correlate of Spatial Presence in an Arousing and Noninteractive Virtual Reality: an EEG and Psychophysiology Study. *CyberPsychology Behav.* 9, 30–45. doi:10.1089/cpb.2006.9.30

Beacco, A., Oliva, R., Cabreira, C., Gallego, J., and Slater, M. (2021). "Disturbance and Plausibility in a Virtual Rock Concert: A Pilot Study," in Disturbance and Plausibility in a Virtual Rock Concert: A Pilot Study. 2021 IEEE Virtual Reality and 3D User Interfaces (New York City: VR), 538–545. doi:10.1109/VR50410.2021.00078

Bergstrom, I., Azevedo, S., Papiotis, P., Saldanha, N., and Slater, M. (2017). The Plausibility of a String Quartet Performance in Virtual Reality. *IEEE Trans. Vis. Comput. Graph.* 23, 1352–1359. doi:10.1109/TVCG.2017.2657138

Biocca, F. (2003). "Can We Resolve the Book, the Physical Reality, and the Dream State Problems? from the Two-Pole to a Three-Pole Model of Shifts in Presence," in EU Future and Emerging Technologies, Presence Initiative Meeting, Venice, May 5-7, 2003 (Citeseer), 13–69. https://www.researchgate.net/profile/Frank-Biocca/publication/228788322_Can_we_resolve_the_book_the_physical_reality_and_the_dream_state_problems_From_the_two-pole_to_a_three-pole_model_of_shifts_in_presence/links/548610730cf289302e2c04cd/Can-we-resolve-the-book-the-physical-reality-and-the-dream-state-problems-From-the-two-pole-to-a-three-pole-model-of-shifts-in-presence.pdf.

Biocca, F. (2002). The Evolution of Interactive Media. *Narrat. impact. Soc. cognitive Found.*, 97–130.

Botvinick, M., and Cohen, J. (1998). Rubber Hands 'feel' Touch that Eyes See. *Nature* 391, 756. doi:10.1038/35784

Bourdin, P., Barberia, I., Oliva, R., and Slater, M. (2016). A Virtual Out-Of-Body Experience Reduces Fear of Death. *PLOS ONE* 12 (1), e0169343. doi:10.1371/journal.pone.0169343

Brubach, L., Westermeier, F., Wienrich, C., and Latoschik, M. E. (2022). Breaking Plausibility without Breaking Presence-Evidence for the Multi-Layer Nature of Plausibility. *IEEE Trans. Vis. Comput. Graph.* doi:10.1109/tvcg.2022.3150496

Bulu, S. T. (2012). Place Presence, Social Presence, Co-presence, and Satisfaction in Virtual Worlds. *Comput. Educ.* 58, 154–161. doi:10.1016/j.compedu.2011.08.024

Chalmers, D. J. (2022). *Reality+: Virtual Worlds and the Problems of Philosophy.* New York: W. W. Norton.

Chalmers, D. J. (2017). The Virtual and the Real. *Disputatio* 9, 309–352. doi:10.1515/disp-2017-0009

Clemente, M., Rodríguez, A., Rey, B., and Alcañiz, M. (2014). Assessment of the Influence of Navigation Control and Screen Size on the Sense of Presence in Virtual Reality Using EEG. *Expert Syst. Appl.* 41, 1584–1592. doi:10.1016/j.eswa.2013.08.055

Cummings, J. J., and Bailenson, J. N. (2016). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychol.* 19, 272–309. doi:10.1080/15213269.2015.1015740

Deutsch, D. (2011). *The Fabric of Reality.* Penguin UK.

Durlach, N., and Slater, M. (2000). Presence in Shared Virtual Environments and Virtual Togetherness. *Presence Teleoperators Virtual Environ.* 9, 214–217. doi:10.1162/105474600566736

Ellis, S. R. (1991). Nature and Origins of Virtual Environments: a Bibliographical Essay. *Comput. Syst. Eng.* 2, 321–347. doi:10.1016/0956-0521(91)90001-l

Fadeev, K. A., Smirnov, A. S., Zhigalova, O. P., Bazhina, P. S., Tumialis, A. V., and Golokhvast, K. S. (20202020). Too Real to Be Virtual: Autonomic and EEG Responses to Extreme Stress Scenarios in Virtual Reality. *Behav. Neurol.* doi:10.1155/2020/5758038

Felton, W. M., and Jackson, R. E. (2022). Presence: A Review. *Int. J. Human-Computer Interact.* 38, 1–18. doi:10.1080/10447318.2021.1921368

Freeman, D., Lister, R., Waite, F., Yu, L.-M., Slater, M., Dunn, G., et al. (2019). Automated Psychological Therapy Using Virtual Reality (VR) for Patients with Persecutory Delusions: Study Protocol for a Single-Blind Parallel-Group Randomised Controlled Trial (THRIVE). *Trials* 20, 87. doi:10.1186/s13063-019-3198-6

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual Reality in the Assessment, Understanding, and Treatment of Mental Health Disorders. *Psychol. Med.* 47, 2393–2400. doi:10.1017/S003329171700040X

Fribourg, R., Argelaguet, F., Lecuyer, A., and Hoyet, L. (2020). Avatar and Sense of Embodiment: Studying the Relative Preference between Appearance, Control and Point of View. *IEEE Trans. Vis. Comput. Graph.* 26, 2062–2072. doi:10.1109/tvcg.2020.2973077

Galvandebarba, H., Chague, S., and Charbonnier, C. (2020). On the Plausibility of Virtual Body Animation Features in Virtual Reality. *IEEE Trans. Vis. Comput. Graph.*

Gao, B., Kim, J.-I., and Kim, H. (2018). Sensory and Perceptual Consistency for Believable Response in Action Feedback Loop. In *Proc. Comput. Graph. Int.*. New York, NY, United States: Association for Computing Machinery, 201–210. doi:10.1145/3208159.3208171

Garau, M., Ritter-Widenfeld, H., Antley, A., Friedman, D., Brogni, A., and Slater, M. (2004). *Temporal and Spatial Variations in Presence: A Qualitative Analysis, 7th International Conference on Presence*. Valencia, Spain).

Garau, M., Slater, M., Bee, S., and Sasse, M. A. (2001). "The Impact of Eye Gaze on Communication Using Humaniod Avatars," in Proceedings of the SIGCHI conference on Human Factors in computing systems.

Garau, M., Slater, M., Pertaub, D.-P., and Razzaque, S. (2005). The Responses of People to Virtual Humans in an Immersive Virtual Environment. *Presence Teleoperators Virtual Environ.* 14, 104–116. doi:10.1162/1054746053890242

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., and Sasse, A. M. (2003). "The Impact of Avatar Realism and Eye Gaze Controlon the Perceived Quality of Communication in a Shared Immersive Virtual Environment," in Proceedings of SIGCHI.

Gorisse, G., Senel, G., Banakou, D., Beacco, A., Oliva, R., Freeman, D., et al. (2021). Self-observation of a Virtual Body-Double Engaged in Social Interaction Reduces Persecutory Thoughts. *Sci. Rep.* 11, 23923. doi:10.1038/s41598-021-03373-x

Graf, S., and Schwind, V. (2020). "Inconsistencies of Presence Questionnaires in Virtual Reality," in 26th ACM Symposium on Virtual Reality Software and Technology, 1–3.()

Grassini, S., and Laumann, K. (2020). Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review. *Front. Psychol.* 11, 349. doi:10.3389/fpsyg.2020.00349

Hall, E. T. (1973). The Hidden Dimension. *Leonardo* 6, 94. doi:10.2307/1572461

Hasler, B. S., Spanlang, B., and Slater, M. (2017). Virtual Race Transformation Reverses Racial In-Group Bias. *PLOS ONE* 12, e0174965. doi:10.1371/journal.pone.0174965

Held, R. M., and Durlach, N. I. (1992). Telepresence. *Presence Teleoperators Virtual Environ.* 1, 109–112. doi:10.1162/pres.1992.1.1.109

Hofer, M., Hartmann, T., Eden, A., Ratan, R., and Hahn, L. (2020). The Role of Plausibility in the Experience of Spatial Presence in Virtual Environments. *Front. virtual real.* 1. doi:10.3389/frvir.2020.00002

Kastanis, I., and Slater, M. (2012). Reinforcement Learning Utilizes Proxemics: An Avatar Learns to Manipulate the Position of People in Immersive Virtual Reality. *Trans. Appl. Percept.* 9 (3), 1–315. doi:10.1145/2134203.2134206

Kober, S. E., Kurzmann, J., and Neuper, C. (2012). Cortical Correlate of Spatial Presence in 2D and 3D Interactive Virtual Reality: an EEG Study. *Int. J. Psychophysiol.* 83, 365–374. doi:10.1016/j.ijpsycho.2011.12.003

Kober, S. E., and Neuper, C. (2012). Using Auditory Event-Related EEG Potentials to Assess Presence in Virtual Reality. *Int. J. Human-Computer Stud.* 70, 577–587. doi:10.1016/j.ijhcs.2012.03.004

Latoschik, M. E., and Wienrich, C. (2021). *Coherence and Plausibility, Not Presence? Pivotal Conditions for XR Experiences and Effects, a Novel Model*. arXiv preprint arXiv:2104.04846.

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence Teleoperators Virtual Environ.* 10, 282–297. doi:10.1162/105474601300343612

Liebold, B., Brill, M., Pietschmann, D., Schwab, F., and Ohler, P. (2017). Continuous Measurement of Breaks in Presence: Psychophysiology and Orienting Responses. *Media Psychol.* 20, 477–501. doi:10.1080/15213269.2016.1206829

Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lect. Hum. Lang. Technol.* 5, 1–167. doi:10.2200/s00416ed1v01y201204hlt016

Llobera, J., Beacco, A., Oliva, R., Senel, G., Banakou, D., and Slater, M. (2021). *Evaluating Participant Responses to a Virtual Reality Experience Using Reinforcement Learning*. in revision. doi:10.1098/rsos.210537

Llobera, J., Spanlang, B., Ruffini, G., and Slater, M. (2010). Proxemics with Multiple Dynamic Characters in an Immersive Virtual Environment. *ACM Trans. Appl. Percept.* 8. Article 3. doi:10.1145/1857893.1857896

Maister, L., Slater, M., Sanchez-Vives, M. V., and Tsakiris, M. (2015). Changing Bodies Changes Minds: Owning Another Body Affects Social Cognition. *Trends Cognitive Sci.* 19, 6–12. doi:10.1016/j.tics.2014.11.001

Martens, M. A., Antley, A., Freeman, D., Slater, M., Harrison, P. J., and Tunbridge, E. M. (2019). It Feels Real: Physiological Responses to a Stressful Virtual Reality Environment and its Impact on Working Memory. *J. Psychopharmacol.* 33, 1264–1273. doi:10.1177/0269881119860156

Meehan, M., Insko, B., Whitton, M. C., and Brooks, F. P. (2002). "Physiological Measures of Presence in Stressful Virtual Environments," in Proceedings of SIGGRAPH. doi:10.1145/566570.566630

Mikolov, T., Grave, E., Bojanowski, P., Puhrsch, C., and Joulin, A. (2017). *Advances in Pre-training Distributed Word Representations*. arXiv preprint arXiv:1712.09405.

Mottelson, A., Petersen, G. B., Lilija, K., and Makransky, G. (2021). Conducting Unsupervised Virtual Reality User Studies Online. *Front. Virtual Real.* 2, 66. doi:10.3389/frvir.2021.681482

Murcia-López, M., Collingwoode-Williams, T., Steptoe, W., Schwartz, R., Loving, T. J., and Slater, M. (2021). "Evaluating Virtual Reality Experiences through Participant Choices," in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (IEEE), 747–755.

Nowak, K. L., and Biocca, F. (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence Teleoperators Virtual Environ.* 12, 481–494. doi:10.1162/105474603322761289

O'regan, J. K., and Noë, A. (2001). A Sensorimotor Account of Vision and Visual Consciousness. *Behav. Brain Sci.* 24, 939–973. doi:10.1017/s0140525x01000115

Ochs, M., Bousquet, J., Pergandi, J.-M., and Blache, P. (2022). Multimodal Behavioral Cues Analysis of the Sense of Presence and Social Presence during a Social Interaction with a Virtual Patient. *Front. Comput. Sci.* 4. doi:10.3389/fcomp.2022.746804

Pan, X., Slater, M., Beacco, A., Navarro, X., Bellido Rivas, A. I., Swapp, D., et al. (2016). The Responses of Medical General Practitioners to Unreasonable Patient Demand for Antibiotics - A Study of Medical Ethics Using Immersive Virtual Reality. *PLoS ONE* 11, e0146837. doi:10.1371/journal.pone.0146837

Parola, M., Johnson, S., and West, R. (2016). Turning Presence Inside-Out: MetaNarratives. *ei* 28, 1–9. doi:10.2352/issn.2470-1173.2016.4.ervr-418

Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting Yourself in the Skin of a Black Avatar Reduces Implicit Racial Bias. *Conscious. cognition* 22, 779–787. doi:10.1016/j.concog.2013.04.016

Petkova, V. I., and Ehrsson, H. H. (2008). If I Were You: Perceptual Illusion of Body Swapping. *PLoS ONE* 3, e3832. doi:10.1371/journal.pone.0003832

Petukhov, I. V., Glazyrin, A. E., Gorokhov, A. V., Steshina, L. A., and Tanryverdiev, I. O. (2020). Being Present in a Real or Virtual World: A EEG Study. *Int. J. Med. Inf.* 136, 103977. doi:10.1016/j.ijmedinf.2019.103977

Poeschl, S., and Doering, N. (2015). Measuring Co-presence and Social Presence in Virtual Environments–Psychometric Construction of a German Scale for a Fear of Public Speaking Scenario. *Annu. Rev. Cybertherapy Telemedicine*, 58–63.

Putze, S., Alexandrovsky, D., Putze, F., Höffner, S., Smeddinck, J. D., and Malaka, R. (2020). "Breaking the Experience: Effects of Questionnaires in Vr User Studies," in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–15.

Regenbrecht, H., and Schubert, T. (2021). *Measuring Presence in Augmented Reality Environments: Design and a First Test of a Questionnaire*. arXiv preprint arXiv:2103.02831.

Rey, B., Parkhutik, V., and Alcañiz, M. (2011). Breaks in Presence in Virtual Environments: An Analysis of Blood Flow Velocity Responses. *Presence Teleoperators Virtual Environ.* 20, 273–286. doi:10.1162/pres_a_00049

Ríos, A., and Pelechano, N. (2020). Follower Behavior under Stress in Immersive VR. *Virtual Real.* 24, 683–694.

Rizzo, A. S., and Kim, G. J. (2005). A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy. *Presence Teleoperators Virtual Environ.* 14, 119–146. doi:10.1162/1054746053967094

Rosenberg, R. S., Baughman, S. L., and Bailenson, J. N. (2013). Virtual Superheroes: Using Superpowers in Virtual Reality to Encourage Prosocial Behavior. *PloS one* 8, e55003. doi:10.1371/journal.pone.0055003

Rovira, A., Swapp, D., Spanlang, B., and Slater, M. (2009). The Use of Virtual Reality in the Study of People's Responses to Violent Incidents. *Front. Behav. Neurosci.* 3, 1–10. doi:10.3389/neuro.08.059.2009

Sanchez-Vives, M. V., and Slater, M. (2005). From Presence to Consciousness through Virtual Reality. *Nat. Rev. Neurosci.* 6, 332–339. doi:10.1038/nrn1651

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence Teleoperators Virtual Environ.* 10, 266–281. doi:10.1162/105474601300343603

Schwind, V., Knierim, P., Haas, N., and Henze, N. (2019). "Using Presence Questionnaires in Virtual Reality," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–12.

Sheridan, T. B. (1996). Further Musings on the Psychophysics of Presence. *Presence Teleoperators Virtual Environ.* 5, 241–246. doi:10.1162/pres.1996.5.2.241

Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence Teleoperators Virtual Environ.* 1, 120–126. doi:10.1162/pres.1992.1.1.120

Skarbez, R., Brooks, F., and Whitton, M. (2020). Immersion and Coherence: Research Agenda and Early Results. *IEEE Trans. Vis. Comput. Graph* PP, 3839–3850. doi:10.1109/TVCG.2020.2983701

Skarbez, R., Brooks, F. P., Jr, and Whitton, M. C. (2018b). A Survey of Presence and Related Concepts. *ACM Comput. Surv. (CSUR)* 50, 96. doi:10.1145/3134301

Skarbez, R., Brooks, F. P., Jr, and Whitton, M. C. (2018a). "Immersion and Coherence in a Stressful Virtual Environment," in Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (New York, NY, United States: Association for Computing Machinery), 24.

Skarbez, R., Neyret, S., Brooks, F. P., Slater, M., and Whitton, M. C. (2017). A Psychophysical Experiment Regarding Components of the Plausibility Illusion. *IEEE Trans. Vis. Comput. Graph.* 23, 1369–1378. doi:10.1109/tvcg.2017.2657158

Slater, M., Brogni, A., and Steed, A. (2003). "Physiological Responses to Breaks in Presence: A Pilot Study," in Presence 2003.

Slater, M., Cabriera, C., Senel, G., Banakou, D., Beacco, A., Oliva, R., et al. (2022). "The Sentiment of a Virtual Rock Concert," in Virtual Reality accepted. preprint. doi:10.21203/rs.3.rs-1361876/v1

Slater, M. (2004). How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments. *Presence Teleoperators Virtual Environ.* 13, 484–493. doi:10.1162/1054746041944849

Slater, M., Linakis, V., Usoh, M., and Kooper, R. (1996). Immersion, Presence, and Performance in Virtual Environments: An Experiment with Tri-dimensional Chess. *ACM virtual Real. Softw. Technol. (VRST)*, 163–172.

Slater, M. (1999). Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence* 8, 560–565. doi:10.1162/105474699566477

Slater, M., Perez-Marcos, D., Ehrsson, H. H., and Sanchez-Vives, M. (2008). Towards a Digital Body: The Virtual Arm Illusion. *Front. Hum. Neurosci.* 2. doi:10.3389/neuro.09.006.2008

Slater, M. (2009). Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Phil. Trans. R. Soc. B* 364, 3549–3557. doi:10.3389/frobt.2016.0007410.1098/rstb.2009.0138

Slater, M., Rovira, A., Southern, R., Swapp, D., Zhang, J. J., Campbell, C., et al. (2013). Bystander Responses to a Violent Incident in an Immersive Virtual Environment. *PLoS ONE* 8, e52766. doi:10.1371/journal.pone.0052766

Slater, M., Sadagic, A., Usoh, M., and Schroeder, R. (2000). Small-group Behavior in a Virtual and Real Environment: A Comparative Study. *Presence Teleoperators Virtual Environ.* 9, 37–51. doi:10.1162/105474600566600

Slater, M., and Sanchez-Vives, M. V. (2022). Is Consciousness First in Virtual Reality? *Front. Psychol.* 13, 787523. doi:10.3389/fpsyg.2022.787523

Slater, M., Spanlang, B., and Corominas, D. (2010a). Simulating Virtual Environments within Virtual Environments as the Basis for a Psychophysics of Presence. *ACM Trans. Graph.* 29, 92. Paper. doi:10.1145/1833349.1778829

Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010b). First Person Experience of Body Transfer in Virtual Reality. *PLOS ONE* 5, e10564. doi:10.1371/journal.pone.0010564

Slater, M., and Steed, A. (2000). A Virtual Presence Counter. *Presence Teleoperators Virtual Environ.* 9, 413–434. doi:10.1162/105474600566925

Slater, M., Usoh, M., and Steed, A. (1994). Depth of Presence in Virtual Environments. *Presence Teleoperators Virtual Environ.* 3, 130–144. doi:10.1162/pres.1994.3.2.130

Souza, V., Maciel, A., Nedel, L., and Kopper, R. (2022). Measuring Presence in Virtual Environments: A Survey. *ACM Comput. Surv.* 54, 1–37. doi:10.1145/3466817

Spanlang, B., Fröhlich, T., Descalzo, F. F., Antley, A., and Slater, M. (2007). "The Making of a Presence Experiment: Responses to Virtual Fire," in PRESENCE 2007 -The 10th Annual International Workshop on Presence (Barcelona, Spain.

Steed, A., Frlston, S., Lopez, M. M., Drummond, J., Pan, Y., and Swapp, D. (2016). An 'In the Wild' Experiment on Presence and Embodiment Using Consumer Virtual Reality Equipment. *IEEE Trans. Vis. Comput. Graph.* 22, 1406–1414. doi:10.1109/tvcg.2016.2518135

Steed, A., Slater, M., Sadagic, A., Bullock, A., and Tromp, J. (1999). Leadership and Collaboration in Shared Virtual Environments. *Virtual Real.* 1999, 112–115.

Tromp, J., Bullock, A., Steed, A., Sadagic, A., Slater, M., and Frecon, E. (1998). Small Group Behaviour Experiments in the Coven Project. *IEEE Comput. Grap. Appl.* 18, 53–63. doi:10.1109/38.734980

Welch, R. B. (1999). How Can We Determine if the Sense of Presence Affects Task Performance? *Presence* 8, 574–577. doi:10.1162/105474699566387

Williamson, J., Li, J., Vinayagamoorthy, V., Shamma, D. A., and Cesar, P. (2021). "Proxemics and Social Interactions in an Instrumented Virtual Reality Workshop," in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–13.

Witmer, B. G., and Singer, M. J. (1998). Measuring Presence in Virtual Environments: a Presence Questionnaire. *Presence* 7, 225–240. doi:10.1162/105474698565686