

Simulating Virtual Environments within Virtual Environments as the Basis for a Psychophysics of Presence

Mel Slater*

ICREA - Universitat de Barcelona &
University College London

Bernhard Spanlang

Universitat Politècnica de Catalunya &
Universitat de Barcelona

David Corominas

Universitat de Barcelona



Figure 1: Overview of the scenario (A) rendered with real-time dynamic shadows and reflections (B) with Gouraud shading.

Abstract

A new definition of immersion with respect to virtual environment (VE) systems has been proposed in earlier work, based on the concept of simulation. One system (*A*) is said to be more immersive than another (*B*) if *A* can be used to simulate an application as if it were running on *B*. Here we show how this concept can be used as the basis for a psychophysics of presence in VEs, the sensation of being in the place depicted by the virtual environment displays (Place Illusion, PI), and also the illusion that events occurring in the virtual environment are real (Plausibility Illusion, Psi). The new methodology involves matching experiments akin to those in color science. Twenty participants first experienced PI or Psi in the initial highest level immersive system, and then in 5 different trials chose transitions from lower to higher order systems and declared a match whenever they felt the same level of PI or Psi as they had in the initial system. In each transition they could change the type of illumination model used, or the field-of-view, or the display type (powerwall or HMD) or the extent of self-representation by an avatar. The results showed that the 10 participants instructed to choose transitions to attain a level of PI corresponding to that in the initial system tended to first choose a wide field-of-view and head-mounted display, and then ensure that they had a virtual body that moved as they did. The other 10 in the Psi group concentrated far more on achieving a higher level of illumination realism, although having a virtual body representation was important for both groups. This methodology is offered as a way forward in the evaluation of the responses of people to immersive virtual environments, a unified theory and methodology for psychophysical measurement.

CR Categories: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities H.1.2 [Models and Principles]: User/machine systems—Human Factors I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality

Keywords: immersive virtual environments, presence, place illusion, plausibility, response function, Markov Chain

* e-mail: melslater@ub.edu

1 Introduction

Immersive virtual environments (IVE) are typically employed to place people within representations of physical reality - for example, for training, various forms of rehabilitation, design, and entertainment [Brooks 1999]. Yet a largely unexplored possibility is to use an IVE system to simulate what can be experienced when using another type of system. This idea was apparently first exploited in an experimental study of presence in virtual environments [Slater et al. 1994], where within a head-mounted display (HMD) delivered VE the participant was able to select and put on a virtual HMD that transferred them to a deeper level environment. In this paper we show how this capability of IVEs, the possibility of simulating one type of IVE system with another, can be used as the foundation for a psychophysical approach to the long studied concept of presence in virtual environments, introducing a new method that avoids the problems of both questionnaire studies and purely physiological or behavioral approaches to measurement.

The concept of ‘immersion’ has previously been regarded as a way to describe the technological capabilities of a virtual reality system - e.g., system *A* is more ‘immersive’ than system *B*, other things being equal, if *A* has a wider field-of-view than *B*, or, say, *A* can generate real-time shadows and reflections but *B* only ‘Gouraud shading’, or *A* has head-tracking but *B* does not [Slater and Wilbur 1997; Draper et al. 1998]. A recent review of the concept of immersion has extended this approach to the idea of a partial order over the class of IVE systems based on an immersion relation [Slater 2009].

ACM Reference Format

Slater, M., Spanlang, B., Corominas, D. 2010. Simulating Virtual Environments within Virtual Environments as the Basis for a Psychophysics of Presence. *ACM Trans. Graph.* 29, 4, Article 92 (July 2010), 9 pages.
DOI = 10.1145/1778765.1778829 <http://doi.acm.org/10.1145/1778765.1778829>

Copyright Notice

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, fax +1 (212) 869-0481, or permissions@acm.org.

© 2010 ACM 0730-0301/2010/07-ART92 \$10.00 DOI 10.1145/1778765.1778829
<http://doi.acm.org/10.1145/1778765.1778829>

The ‘immersion’ relation between systems A and B denoted by $A \supset B$ occurs when A can be used to build an application in which a participant would experience a simulation of that application as if running in system B . By definition A is at a higher level of immersion than B if $A \supset B$ but not $B \supset A$. For example, in principle it is possible to simulate the experience of being in a 4-sided Cave system using a wide field-of-view head-tracked head-mounted display (HMD): a virtual environment can be built that is delivered through such a HMD where a participant enters into a virtual Cave, sees a dynamic virtual body representation that is a likeness of him- or herself from an egocentric viewpoint, and experiences a virtual environment running in that virtual Cave. Similarly, it is possible using a HMD to simulate a virtual environment delivered by a powerwall type of display or even a desktop system. We say ‘in principle’ since clearly this is based on a series of abstractions – ignoring aspects such as the weight of the HMD compared to the weight of shutter glasses that might be used for a Cave, differences in display resolution, brightness, and so on. Of course this requires a multimodal system exploiting not just vision but the auditory and especially haptic modalities. However, it is the case that given sufficient resources, each of the above simulations would be feasible with even today’s technology. Such examples can alternatively be regarded as thought experiments but mostly they would be realizable.

The immersion relation \supset imposes a partial order over any set of IVE systems. This is based on the physical properties of each system and the corresponding set of computer programs that enable its use. However, in order to understand the utility of a system for particular applications we need also to be able to assess how participants respond to applications that are built with it. The concept of ‘presence’ has for many years been thought to provide a ubiquitous measure of one aspect of the subjective experience of being in a virtual environment that applies across different applications and systems. Presence refers to the illusion of being in the scene displayed by the IVE system, a concept developed in the early 1990s, for example [Held and Durlach 1992; Sheridan 1992]. Since then the concept has become diffuse, and has been thought of as applying to a very wide range of different types of subjective response to mediated experience [Lombard and Ditton 1997]. Moreover, there has never been a unified and generally accepted approach to the measurement of presence, rather a set of different methods have been used (questionnaire based, behavioral, physiological) each with their own set of problems.

2 Measuring Presence

Since presence has been thought of as a subjective experience, eliciting the strength of the feeling of ‘being there’ using questionnaires has been one obvious approach to measurement. The paradigm that developed was to carry out experiments where particular aspects of the virtual environment were manipulated and the resulting questionnaire responses were based on Likert scales regarding how much the participant felt themselves to ‘be there’ (e.g., scoring 1 for ‘not at all’ and 7 for ‘very much’). For examples see [Witmer et al. 2005; Lessiter et al. 2001; Schubert et al. 2001] with a review in [Sanchez-Vives and Slater 2005]. However, this approach, certainly when used alone, has several problems: it does not seem to be able to distinguish between an experience in reality and virtual reality [Usoh et al. 2000], the measurements may be unstable [Freeman et al. 1999], it has problems in actually assessing the concept itself [Slater 2004], and there are methodological problems in analyzing subjective rating data as if it were interval or ratio data [Gardner and Martin 2007; Slater and Garau 2007].

If a person feels that they are in the scenario depicted by the IVE then they should exhibit behavioral and physiological responses

concomitant with that feeling - i.e., they should have physiological responses and behaviors as if they were there. This is the basis for the application of virtual environments to real-world situations such as training, rehearsal or psychotherapy, since if the participant does not to some extent act as if they were ‘there’ then nothing much useful could be gained from their IVE experience in relation to these applications. In this type of approach an experiment is designed in order to elicit a clearly measurable physiological or behavioral response, and then assess how this changes under various experimental conditions. The typical measure used is based on the physiological characteristics of stress since this is relatively easy to identify (using arousal as measured by increasing skin conductance responses, increase in heart-rate and decrease in heart-rate variability). Skin conductance and heart rate were used to examine people’s stress response to a visual cliff in [Meehan et al. 2002; Slater et al. 2009], and also the impact of different levels of latency on the experience [Meehan et al. 2003]. The effect of social interaction in IVEs has also been studied extensively based on physiological measures, for example [Slater et al. 2006].

Although the use of behavioral and physiological responses as surrogates for presence is methodologically sound, this avoids rather than solves the problem of conceptualizing and measuring presence, since a situation must be set up in the virtual reality that would cause stress or some other clearly measurable physiological response. Not every application is amenable to that, and it does not make sense to deliberately add a stressful event into a virtual environment scenario solely for the purpose of measurement. This still leaves open the issue of a ubiquitous measure that applies across different types of application and system.

3 Deconstructing Presence

In one review of the concept it was argued that presence should be defined as the extent to which participants respond realistically to virtual events and situations [Sanchez-Vives and Slater 2005] rather than as their sense of ‘being there’. In this approach ‘response’ is considered as multilevel, from low level automatic physiological responses, through non-conscious behavioral reflexes, volitional behavioral responses, emotional responses through to high level cognitive responses and thoughts (including the sensation of being there). This response profile is what defines presence, and the more that the measured responses point in the same direction, i.e., consistent responses that point to the participant treating the virtual reality as if it were real, the greater the degree of presence.

This approach helps to solve the conceptual problem of definition and of measurement. Presence is exhibited when people behave as they would in reality, and the extent to which this occurs is measurable in principle. In other words presence is identified with its operationalization as a measurable property of the actions of people within IVEs compared with their expected or actually observed behavior within similar real-world settings. However, if we regard the actions of people as the surface manifestation of a deeper quality of subjective experience, and if this quality itself can somehow be measured, then we would have the basis for constructing a theory, one that might predict when people are likely to respond realistically. An appropriate theoretical framework well integrated with empirical studies is an essential requirement for progress in this field, and the area of presence research has been dogged by the lack of any theoretical framework that also implicitly includes how presence itself might be measured.

Moreover, lumping everything into ‘being there’ misses another very important aspect of people’s experience. In physical reality, for example, you know very well that you are there, but you can encounter events that are not what they appear to be. For example,

you enter a room and see a person standing at the far end, and you wave to them. Later you realize that there was no person there but that it was a shop dummy. This happens in IVEs - for example, a person responds realistically to a virtual character for a while, until realizing that the character engages in repetitive or inappropriate behaviors [Garau et al. 2008] and the credibility of the virtual environment situation is lost. This plausibility is separable from the sensation of being there.

[Slater 2009] put forward the thesis that there are two orthogonal components of presence to consider. The first is Place Illusion (PI) the original idea of the sensation of being in the place depicted by the VE. This is a qualia, a quality of our subjective experience, like seeing the ‘redness’ of the color red. The second is Plausibility (Psi), the illusion that what is apparently happening is really happening. Both of these are known by the participant to be illusions, but knowing that they are illusions does not extinguish them. It was argued that when there is PI and Psi then ‘response as if real’ is likely to occur.

The physical basis of PI was argued to be sensorimotor contingencies that correspond to those of physical reality. When a person perceives by carrying out actions that result in changes in (multi-sensory) perception much as in physical reality, then the simplest hypothesis for the brain to adopt is that what is being perceived is actually there - i.e., that the person is in the place depicted by the IVE. The physical basis of Psi was postulated to be the extent to which the system is programmed to produce correlations with the behavior of the participant, how much events in the IVE refer personally to the participant, and the overall credibility of the scenario (in particular in relation to how a similar situation might be in physical reality).

Another important point about the framework presented in [Slater 2009] was the fusion of PI and Psi in the notion of a ‘virtual body’. When you wear a head-tracked HMD for example, and look down towards your own body, what do you see? Unless a virtual body has been programmed, you have no body. The act of looking at yourself is a natural movement with concomitant changes in perception, and what you see determines a critical aspect of the realness of the situation. This applies also in projection type systems such as a Cave. As you move around the virtual environment depicted in a Cave, ideally you would see not only your own real body but also shadows and reflections of your body in the VE.

4 PI, Psi and Color: An Analogy

In the perception of color there is a physical basis, which is the actual wavelength distribution of the light emitted and/or reflected from a surface patch. However, the sensation of color depends on a number of complex perceptual mechanisms. In the tristimulus theory of color the average person’s response to light can be computed by integrals of the wavelength distribution times response functions for each of the ‘red’, ‘green’ and ‘blue’ cones - e.g., [Fairchild 2005]. In practice these response functions are determined empirically through color matching experiments. It is impossible to know how someone really experiences the color ‘red’, but we do know that typically normally sighted people agree on what is ‘red’. Moreover individuals can match a target color with a test patch by additively mixing three primaries in order to make the patch appear to have the same color as the target. In such experiments people are never asked questions such as ‘How red is this color on a scale of 1 to 7?’ (which has been the approach used in presence) but instead they are asked to carry out an action (e.g., turning dials on three projectors) to produce a surface patch, which has the same sensation of color as the target. When this type of experiment is repeated several times over several people, empirically based response functions can

be constructed that then are used in agreed international standards on the representation of color. This works because patches that emit or reflect light that have quite different wavelength distributions can nevertheless be perceived as the same color by an individual (these are called *metamers*). A particular color sensation can be thought of as an equivalence class over an infinite number of different wavelength distributions.

There are an infinite number of possible physical realizations of any VE application, depending on variations in the hardware and associated computer programs. Some differences between different realizations will be important, with respect to the associated PI sensation that an individual might feel. Other differences may have no influence. A particular PI sensation can therefore also be thought of as representing an equivalence class amongst the set of physical realizations that result in the same sensation of PI. The same applies to Psi. Here two realizations being in the same equivalence class would mean that the ‘average participant’ would match them.

A set of VE systems can be organized into a partial order using the \supset relation. Consider a particular realization of an application in a system S_0 such that $S_0 \supset S_j, j = 1, \dots, m$. This provides a basis for matching experiments. Suppose that an individual spends some time in the application realized with S_0 , and is asked to become aware of the corresponding feeling of PI that is experienced in that system and application. This is like looking first at the target color. Now the experimental subject has access to a set of virtual ‘buttons’ that when selected effect transitions to the various $S_j, j = 1, \dots, m$. At each S_j the participant can form an assessment, matching their sensation of PI in S_j with the feeling that they had while experiencing S_0 . If they find that the feeling is the same, then for this subject with respect to this environment, there is an equivalence between PI experienced in S_j and S_0 . We refer to that as a *match*. More to the point, if the subject is placed in some S_j and asked to choose transitions and stop whenever they achieve the same feeling of PI that they had in S_0 then we can observe the sequence of transitions that they make, in order to understand which are the most important. This can be repeated several times over several subjects, each time allowing for a different sequence of choices. From this data it would be possible to estimate the probability distributions of a match over the different configurations.

In practice attention would focus on a particular set of properties of interest that characterize a VE system $[s_1, s_2, \dots, s_m]$. For example s_1 might refer to the average frame rate achievable over the application, s_2 the number of degrees of freedom of headtracking, and so on. We use the convention that if there are two systems S' and S and $s'_i \geq s_i$ for the i th property, then S' can realize a level of this property that is equal to or higher (e.g., greater frame rate) than S . We also assume that if a system can achieve level s'_i of property i , and $s'_i \geq s_i$, then it can also achieve level s_i . If we refer to s_{ji} as the realization of the i th property under system S_j , then the ordering over the set of systems of interest implies that $s_{0i} \geq s_{ji}$, for all i, j . In other words S_0 can simulate the application of interest, and also it can simulate the application as if it were running on systems $S_j, j = 1, \dots, m$. We carried out an experiment to illustrate the methodology described above, which is described in the next sections.

5 The Experiment

5.1 Recruitment

Twenty participants (10 of them males) were recruited by advertisement through the university campus. Their average age was $27 \pm 8(S.D.)$ years. Only 2 had any prior experience of virtual reality. None of the participants had any prior knowledge of the

experiment or the general research of the group.

They were exposed in various ways to be described below to a virtual environment that consisted of a room $4 \times 4 \times 2.8m^3$ with various objects in it as shown in Figure 1.

5.2 Materials

Throughout the whole experiment a Fakespace Labs Wide5 HMD was used, which has a field-of-view $150^\circ \times 88^\circ$ with an estimated 1600×1200 resolution. The software environment was XVR [Carrozzino et al. 2005], together with a hardware accelerated avatar library (HALCA) [Gillies and Spanlang 2010]. The participant's head was tracked by an Intersense PC Tracker IS 900 system. The joystick of the Intersense system was also used in one of the conditions (see below). Tracking data was streamed to the VR system via VRPN [Taylor et al. 2001] and used to turn the avatar's head and to adapt the viewpoint in the virtual environment according to the participant's head position and orientation.

The participants wore a tight fitting Velcro suit that had retroreflective markers attached that enabled our system to track the whole body movements of our participants. The marker-based infrared tracking system was a 12 camera Optitrack system from Natural-point¹ that could track, in our configuration, a volume of approximately $2.5m \text{ width} \times 2.5m \text{ length} \times 3m \text{ height}$. 2D marker information was transferred from the cameras via USB to the Natural-Point Arena motion capture software in which the dynamic skeletal configuration of the participants was reconstructed. The movements are reconstructed at 100 Hz with millimeter accuracy. From Arena the skeletal motion data was streamed to HALCA via the NatNet protocol. The skeletal motion data was then mapped so that the avatar posture matched that of our participant to a good degree. The avatars (male and female) were from AXYZ-design².

Participants were asked to sit on the chair (which was located in the center of the volume) but were allowed to make any movement they wanted. The chair was also shown in the virtual environment and registered in the same position as the real chair. They held the joystick in their dominant hand throughout the experiment, but it was only useful for the simulated powerwall condition described below.

5.3 Properties

The property vector $S = [I, F, D, V]$, where I refers to the illumination model used (Gouraud shading, static global illumination, global illumination with dynamic changes), F the field-of-view (small or large), D the display type (simulated powerwall or HMD) and V the virtual body self-representation of the participant (none, static avatar, fully tracked avatar). It should be noted that only the Wide5 HMD was used throughout, and the system simulated each instance of S . Each of these properties is detailed below. We call each instance of the property vector a *configuration*.

(I) Illumination

- (I=0) Gouraud shading. In Gouraud shading mode the environment was rendered without taking global illumination effects into account.
- (I=1) Static global illumination. The environment was rendered with view independent global illumination effects but without dynamically changing shadows or reflections. This illumination was achieved by a light tracing and texture baking approach from Mental Images in Autodesk Maya. If the

¹<http://www.naturalpoint.com/optitrack>

²<http://www.axyz-design.com>

participant's avatar was visible it did not cast any shadows and was not reflected in the mirror.

- (I=2) Dynamic global illumination. In this mode in addition to static shadows as described in the previous mode there were dynamic soft shadows cast by the virtual character onto the environment (using the GPU based percentage closer soft shadows approach [Fernando 2005]) and the environment and the participant's avatar were reflected in the virtual mirror (using a stencil mirror approach [Kilgard 2000]).

(F) Field-of-view

The meaning of field-of-view depended on which of the two display types were used (powerwall or HMD, see below).

- (F=0) Small field-of-view. In the case of the powerwall the size was $1.25m \times 0.69m$. In the case of the HMD display the field-of-view was restricted to 60% of the full FOV.
- (F=1) Large field-of-view. In the case of the powerwall the size was $2m \times 1.1m$. In the case of the HMD the full available FOV was used ($150^\circ \times 88^\circ$).

(D) Display Type

- (D=0) A simulated powerwall display. In the simulated powerwall display the participant viewed the environment from within a virtual viewing room that had a virtual back projected powerwall on one of its walls. The scenario room was displayed on the virtual powerwall in stereo, and they were seated about 1.35m away from it. Head-tracking was used normally with respect to the viewing room, but the participant could navigate through the environment displayed on the virtual powerwall by using a joystick. The viewing room that contained the powerwall was gray and illuminated by the virtual powerwall.
- (D=1) The head-mounted display. In the head mounted display mode the participant viewed the environment from a first person view. The viewpoint was that of the avatar's eyes. The participant could look around the room using normal head-movements.

(V) Virtual Body

- (V=0) No virtual body. In this mode there was no avatar representation.
- (V=1) Static virtual body. In the static virtual body mode there was an avatar that only rotated to match the direction that the participant faced but otherwise did not move. The static avatar appeared to be in a comfortable seated pose.
- (V=2) Full body-tracked avatar. In the full body tracked avatar mode the avatar's pose was updated with the real-time whole body tracking data of the optical tracking system.

Altogether there were 36 possible configurations: 3 types of illumination \times 2 field-of-views \times 2 display types \times 3 virtual body modes. Some examples are shown in Figures 2 and 3 and on the accompanying video.

5.4 Procedures

When the experimental participants arrived at the laboratory they were given an information sheet to read, and the experimental procedures were also explained to them verbally. They read and signed an informed consent form, the experiment had been approved by the institutional ethics review committee. They then were assisted to put on the full body tracking suit and the HMD. The area of the laboratory where they wore the body suit and HMD could be



Figure 2: Examples of the different scenario properties (A) The participant wears a body suit (B) Gouraud shading seen from the first person perspective of the avatar (C) there is global illumination and the avatar is static (D) there is global illumination and the avatar is dynamic moving based on the tracking in A.



Figure 3: The powerwall simulation (A) large screen with static avatar and global illumination (B) small screen and dynamic avatar with global illumination.

closed off from the rest of the laboratory by a black curtain, so that the participants would be in darkness once the experiment started. They were seated throughout. They put on the HMD and were left to become accustomed to the displayed environment for 1.5 minutes. Then they were shown that it was possible to manipulate the properties of the environment by changing each of the illumination, field-of-view, display type and virtual body settings. They did not change these settings themselves but the settings were verbally labeled, and the experimental operator changed them on request from the subject. This continued until the participants were familiar with all the possible settings and the transitions that they could make. The verbal labels for transitions that they learned were: (I) ‘illumination’, (F) ‘display size’, (D) ‘navigation’, and (V) ‘avatar’. These were taught in the order I, F, D and V, and no participant had trouble learning them.

5.5 Transitions

After the period of acclimatization and training described above, participants had 5 trials, each of which they started from a different basic configuration, and then were encouraged to make transitions and stop whenever they had reached a level of PI or Psi that they felt was equivalent to that obtained in the full environment. The starting configurations were as shown in Table 1. Whenever they wished to make a transition to the next one, they would call out the required transition using the previously learned transition labels.

Trial	Illumination (I)	Field of View (F)	Display Type (D)	Virtual Body (V)
1	1	0	0	0
2	0	0	0	1
3	0	0	0	0
4	0	0	1	0
5	0	1	0	0

Table 1: The Basic Starting Conditions for the 5 Trials

In order to encourage participants to think carefully about their transitions, and avoid the possibility that they would straight away simply choose the full configuration [2, 1, 1, 2] (which would have made the problem trivial) we imposed the following rules:

- Transitions could only be in one direction - i.e., having chosen a higher level of one property they could not undo that and go backwards. For example, if they had made the transition from ‘Gouraud shading’ to ‘static shadows’ they could not later go back to ‘Gouraud shading’. An additional reason for this was simplicity of the task, and also to limit the total number of actual transitions that would be possible.
- Only one-step transitions could be made. For example, they could not choose to jump directly from ‘Gouraud shading’ to ‘dynamic shadows and reflections’ but would need to first make a transition to ‘static shadows’.
- In order to avoid participants choosing transitions in a random order simply to get to the initial [2, 1, 1, 2], we imposed a cost structure on the transitions. We told the subjects that they would start out with € 10. Every transition would cost them € 1. If they stopped too early i.e., before they were in the PI or Psi state, they would lose € 5. On the other hand if they reached the desired state they would get a bonus of € 5. We did not explain to them, and no subject actually asked, how we, the experimenters, would know which state they were in. The rule that we used in fact was that they lost the € 5 if they stopped in less than 3 transitions (but this was not told to them). They were informed after each of the 5 trials only whether they had passed the test or not. Their final payment for the experiment was the maximum achieved amongst their 5 trials. The final payments in fact were € 11 (3 subjects) and € 12 (17 subjects).
- After the participants had chosen the configuration at which they had made a match they were asked to continue until they had completed 5 transitions.

5.6 Experimental Design

While participants were experiencing the configuration [2, 1, 1, 2] (each property at the highest level) they were given one of the following two instructions:

- (PI) Pay attention to your feeling that you are in that room that you can see. Later we will ask you to try to get that feeling of being in that room again.
- (Psi) Pay attention to how real this feels. Later we will ask you to try to get that feeling of reality again.

Half of the participants were given the instruction PI and the other half Psi. Participants were assigned to one of the two groups by order number in which they arrived at the laboratory. Odd numbered subjects were assigned to the PI group and even numbers to the Psi group.

Overall, participants spent 1.5 minutes first exploring the virtual room in configuration [2, 1, 1, 2]. Then they learned the various possible transitions that they could make, normally taking 2-3 minutes for this.

Hence our experiment had one factor with two levels, PI or Psi, corresponding to the instruction that the participants had been given.

There were two different types of response (or dependent) variable. The first was the 4-tuple $[I, F, D, V]$ at which a participant declared a match (with PI or Psi). The second consisted of the transitions - i.e., the set of all transitions from configuration i (e.g., [1, 1, 0, 1]) to another configuration j (e.g., [2, 1, 0, 1]).

6 Results

6.1 Method of Analysis

We make the simplifying assumption that the results of the five trials were statistically independent. There is not true independence between these, however, since obviously the same person carried out each of the 5 trials, and may have learned from trial to trial. There are two reasons to suppose, however, that the independence assumption may not have been violated. First, by design each trial started from a different configuration, and therefore participants were forced to think each trial anew. Second, empirically if we let n_{ij} be the transition number at which the i th subject declared a match in the j th trial ($i = 1, \dots, 20; j = 1, \dots, 5$), we find no significant correlations between the columns of the matrix \mathbf{n} . (The highest correlation is between trials 1 and 3 with $r = 0.42$, $P = 0.06$, the next highest is between trials 2 and 3 with $r = 0.32$, $P = 0.17$, and so on). Hence although zero correlation between the trials does not prove the strong requirement of their statistical independence, the assumption is at least not contradicted empirically.

We follow two methods of analysis. First, we consider the configurations $[I, F, D, V]$ at which participants declared a match in each of their 5 trials. From these we can estimate the joint probability distributions: $P(I = i, F = f, D = d, V = v|c) = p_c(i, f, d, v)$ where $P(E|c)$ represents probability of event E conditional on c (PI or Psi). From these probability distributions we can compute any marginal or other conditional distributions of interest. In particular we define $\pi(i, f, d, v) = P(PI|i, f, d, v)$ and $\psi(i, f, d, v) = P(Psi|i, f, d, v)$ as the conditional probabilities of a match being declared when the participant is experiencing configuration $[i, f, d, v]$. These can be computed using Bayes' Theorem from $p_c(i, f, d, v)$.

Second, we consider the transitions as a Markov Chain over the configurations of the system. That is we assume that the probability of choosing a transition to any particular (allowable) configuration is only dependent on the current configuration, and not on prior history. Then using the results of all the transitions made by the subjects, we can estimate the transition matrix $P_{ij,c}$, the probability of a transition to configuration j given that the current configuration is i , where i, j range across the configurations and c is the condition of interest (PI or Psi). From the two resulting transition matrices it is easy to compute the probabilities of being in the various configurations after the successive transitions. Given the rules applied to the possible transitions, the configuration [2, 1, 1, 2] is absorbing, since once that is reached no further transitions are possible.

6.2 Probability Distributions

For any particular $[i, f, d, v]$ we can estimate the probability of a match in that configuration as the number of times that subjects stopped in that configuration over the total number of stops. Each

subject carried out five trials, and stopped in each one. Hence the denominator within each group (PI and Psi) is 50.

The two probability distributions $p_c(i, f, d, v)(c = PI, Psi)$ are shown in Figure 4A (in A only configurations with at least one probability > 0.04 are shown). A Chi-Squared test on the difference between the two distributions shows that they are highly significantly different ($P < 2.0 \times 10^{-6}$). (The Chi-Squared test combined some frequencies together to avoid values of less than 5, as is standard practice). The PI group chose the large display and HMD together more often than the Psi group (88% compared to 60%), discussed later. The PI group's most likely stopping configuration was to leave the illumination as Gouraud shading, but with the full body tracked avatar. For the Psi group the most likely stopping configuration was with no avatar, but with static shadows. The next most likely stopping configuration was with dynamic illumination, a static avatar, with the small field-of-view HMD.

The participants chose their responses non-randomly under both conditions. To see this, assume that the participants were choosing their stopping configurations randomly. Then in Figure 4A we should find a fairly uniform distribution amongst the stopping configurations reached. In fact if we carry out a Chi-squared test comparing each distribution with the theoretical uniform distribution, then random choice for the stopping configurations is an inconceivable hypothesis (in both cases the significance level $P < 1.0 \times 10^{-9}$). In fact were subjects making random choices then we would have also expected the distributions for PI and Psi to be similar, which is not the case.

Figure 4B shows the probabilities $\pi(i, f, d, v)$ and $\psi(i, f, d, v)$ where only configurations with at least one $n \geq 10$ and one $p > 0$ are shown. The meaning of Figure 4B is that, for example, for those in the PI group, the configuration [1, 1, 1, 2] was reached 16 times, and on 14 occasions the participant found a match on PI compared to [2, 1, 1, 2] leading to a probability of 0.875. It can be seen that there are some very striking differences between the two groups. For example, 89% of the 16 times that the Psi group reached [2, 0, 1, 1] there was a match for Psi, whereas there were 2 matches out of 4 in the PI group. There is a similar difference between the two groups in [0, 1, 1, 1] and [0, 1, 1, 0].

6.3 Transitions

We constructed matrices of transition probabilities for the PI and Psi groups. Each subject made 5 transitions in each of the 5 trials leading to 250 transitions for each group. Recall that they were asked to continue making transitions after they had declared their match. Given the structure of the transitions possible these are highly sparse matrices, with 58 and 61 non-zero entries respectively. There are 36 possible configurations, but our main interest is [0, 0, 0, 0] (i.e., Gouraud shading, small field-of-view, powerwall, and no avatar). From this start we consider the evolution of the configurations reached for the two groups. Let u be the 1×36 vector corresponding to starting configuration [0, 0, 0, 0], i.e. u has a 1 corresponding to this configuration and 0 elsewhere. Then if P is the transition matrix the vectors uP^n ($n = 1, 2, 3, \dots$) give the probability distributions over the configurations after n transitions.

Figure 5 shows the estimated probability distributions over the configurations at each of the transitions. (Transition 6 would be to the absorbing configuration [2, 1, 1, 2]). There are early signs of the difference between the two groups. At transition 1 all of the PI group changed either the size of the display or changed from powerwall to HMD. For the Psi group the transitions were 60% to the HMD, 30% to the powerwall, but even at this early stage 10% to improve the illumination to static shadows. By transition 3, considering the maxima of the distributions, the two groups are symmetric

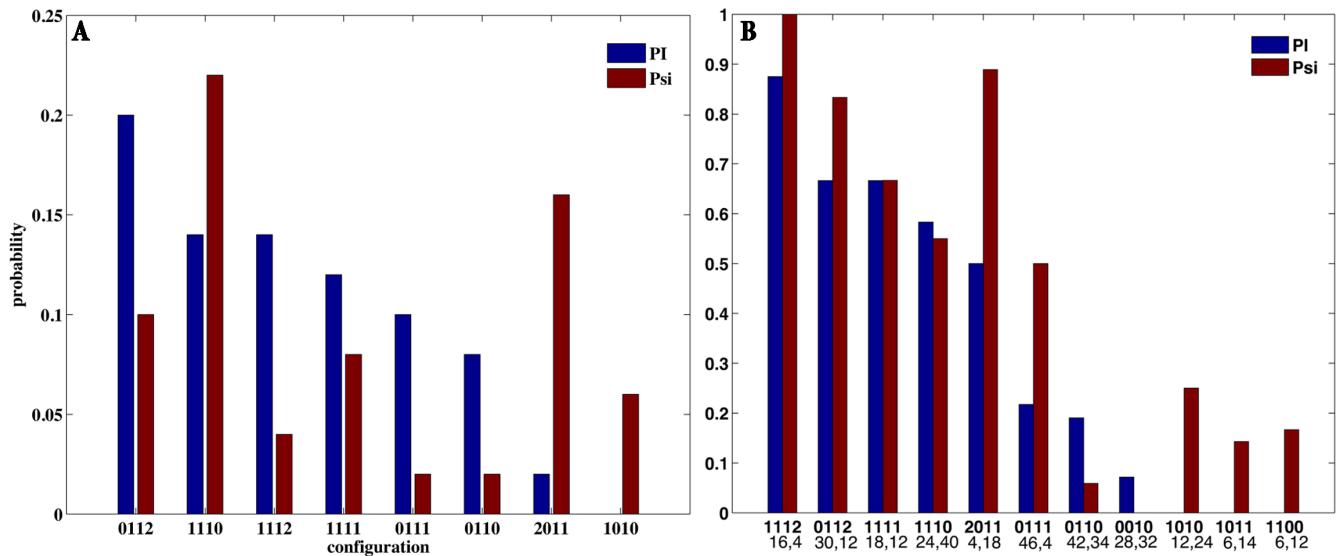


Figure 4: Response functions. (A) $p_c(i, f, d, v)$ the probability distributions over matching configurations. (B) $\pi(i, f, d, v)$ and $\psi(i, f, d, v)$, the conditional probabilities of a match in the given configuration. The pairs of numbers under the x-axis in B are the n 's corresponding to the probabilities.

- both have chosen the HMD and large display size, but the PI group has included the static avatar [0, 1, 1, 1] and the Psi group the static shadows illumination [1, 1, 1, 0]. At the fourth transition this symmetry is maintained ([0, 1, 1, 2] compared to [2, 1, 1, 0]) and again at the fifth transition ([1, 1, 1, 2] compared with [2, 1, 1, 1]). Given the cost structure of making transitions, the order is also important. The PI group tended to improve first the avatar to the best one and then turned attention to the illumination, whereas the Psi group tended to do the opposite.

7 Discussion

Here we consider the claims of the theory outlined in Section 3.

Natural sensorimotor contingencies are important for PI. In this experimental design participants could observe the virtual room on a simulated powerwall manipulating their view by using a joystick, or directly through the HMD changing their view by natural head movements. Moreover, a wider field-of-view gives a better approximation to natural sensorimotor contingencies than a more narrow field-of-view - in the case of the direct HMD interface because head movements would change the view in a way similar to physical reality, whereas with a narrow field-of-view more head movements are needed. In the PI group 88% chose to stop in a condition where both HMD and wide field-of-view were chosen, compared with 60% in the Psi group. This difference is significant ($P < 4 \times 10^{-4}$, one-sided test against the hypothesis that the proportion for PI is greater than for Psi). Additionally, considering the transitions shown in Figure 5, by the second transition the probability that both large field-of-view and HMD are chosen is 0.74 for the PI group and 0.49 for the Psi group. By the third transition these become 0.95 and 0.67 respectively. Finally $\pi(0, 1, 1, 0) = 0.1905(n = 42)$ whereas $\psi(0, 1, 1, 0) = 0.0588(n = 34)$. These are the probabilities that amongst all the times that the configuration [0, 1, 1, 0] had been reached (Gouraud shading, larger field-of-view and HMD, no avatar) that a match had been chosen in that configuration. Again the difference is significant ($P = 0.035$, one-sided test).

Correlations between self-actions and events are important for Psi. In the scenario of this experiment there were no actual events ex-

cept those caused by the participants (i.e., body movement). For these events to have counterparts in the virtual reality, the participant needs to have either a static body with reflections in the mirror [2, *, *, 1], or a dynamic body with or without reflections in the mirror [* , *, *, 2]. In these cases, movements of the participant would result in changes in the environment. For the cases with Gouraud shading, $\psi = 0.0179(n = 112), 0.0455(n = 44), 0.3571(n = 28)$ for the no avatar, static avatar and dynamic avatar respectively. For the cases with static shadows, there is no change from no avatar to static avatar (both are approximately 0.3), but again, a large change to the dynamic avatar ($\psi = 1, n = 8$). Finally in the case of illumination with shadows and reflections the values are 0.1538 (n=26), 0.5789 (n=38) and 0.8571 (n=14) respectively. (Each of these differences are significant, both $P < 0.015$). Moreover, $\psi(0, 1, 1, 2)$ is high, and so is $\psi(2, 0, 1, 1)$ (Figure 4B). Note that for Psi, SCs are less important, some participants seemed willing to sacrifice the larger display size (Figure 4, cases of the form [* , 0, *, *]).

Illumination realism may be more important for Psi. We saw in the discussion of the transitions (Figure 5) that those in the PI group tended to first establish the wide field-of-view and HMD and then gravitated towards obtaining an avatar, whereas those in the Psi group more towards improving the illumination. Additionally, in Figure 4B we can see that the second largest value of ψ occurs for condition [2, 0, 1, 1].

The fact that the illumination type was found to be less important in the PI group is consistent with other evidence. In [Zimmons and Panter 2003] a between-groups experiment assessed presence using questionnaires and physiological responses to the visual cliff (the pit-room). Subjects experienced one of 5 types of illumination model, ranging from Gouraud shading through to radiosity. Physiological responses indicating stress increased significantly once subjects saw the edge of the precipice over which they were virtually standing. However, there were no significant differences between the groups with respect to a presence rating scale nor with respect to the physiological stress responses. In other words, illumination realism apparently made no difference. However, in [Slater et al. 2009] it was found that on a presence rating scale there was a sig-

nificantly higher mean score for a group that experienced the pit room with real-time ray tracing, compared with another group that experienced it with only ray casting. This was also backed up with physiological evidence showing a greater stress response for the ray tracing group.

In this second experiment [Slater et al. 2009] participants were endowed with a very simple virtual body, but one that only partially moved in response to the participant's overall movements (e.g., leaning forward or swaying, and one arm movement). In the case of the ray tracing group movements of the virtual body were accompanied by real-time changes to shadows and reflections of that body in the environment. For the ray casting group there were only static shadows. Note that both experiments used a Virtual Research V8 HMD which has 60 degree diagonal FOV - i.e., a small FOV compared to the current experiment. To consider the experiment of [Zimmons and Panter 2003] we can compare $\pi(0, 0, 1, 0) = 0.0714(n = 28)$ (Gouraud shading, small FOV, HMD, no avatar) with $\pi(1, 0, 1, 0) = 0(n = 12)$ (static shadows, approximately equivalent to radiosity). The difference between these two is not significant. However, for [Slater et al. 2009] we need to compare $\pi(1, 0, 1, 1) = 0(n = 6)$ (static shadows, low FOV, HMD, static avatar) with $\pi(2, 0, 1, 1) = 0.5(n = 4)$ (global illumination). Of course the numbers in the second case are too small to test significance, but they are at least consistent with the results of the paper.

Additionally if we consider what might have happened if a large field-of-view had been used in the first experiment, we find $\pi(0, 1, 1, 0) = 0.1905(n = 42)$ and $\pi(1, 1, 1, 0) = 0.5833(n = 24)$ ($P = 0.0004$). In other words, a specific prediction of our method is that repeating the experiment described by [Zimmons and Panter 2003] but using a field-of-view similar to that of the Wide5 HMD would result in a significant difference between the responses to Gouraud shading compared with the radiosity-like illumination method.

The virtual body is important for both PI and Psi. It is of great interest that having a virtual body appears to be an important element for PI. This was first suspected in the early days of presence research and as reported earlier, there was one experiment that looked at this issue. Evidence that it is of importance for Psi as well was discussed above, and also Figure 4B shows that the one of the highest probabilities for Psi is in the condition [0, 1, 1, 2].

8 Conclusions

This paper is based on recent theoretical work described in [Slater 2009] that introduced the concept of immersion as a relation over virtual reality systems, forming a partial order. The basis of immersion is simulation, where one system may be used to simulate an application as if it were running in another system. It is postulated that probability distributions can be defined over a set of immersive systems, that act as response functions, for sensations such as Place Illusion and Plausibility. Each type of system can result in a certain type of qualia, and the goal of the matching experiments is to find the equivalent of 'metamers' in color science, i.e., configurations that give rise to similar feelings, just as different wavelength distributions can give rise to the same color sensation. We have shown how this can be done with a relatively simple example, only manipulating four aspects giving rise to 36 different abstract IVE systems. We have shown that depending on the matching criteria used (PI or Psi) that the probability distributions are different, and that the results fit with some previous work, and also are consistent with the theory put forward in [Slater 2009]. (Of course consistency does not imply truth).

Moreover, on the basis of the derived response functions π and ψ it

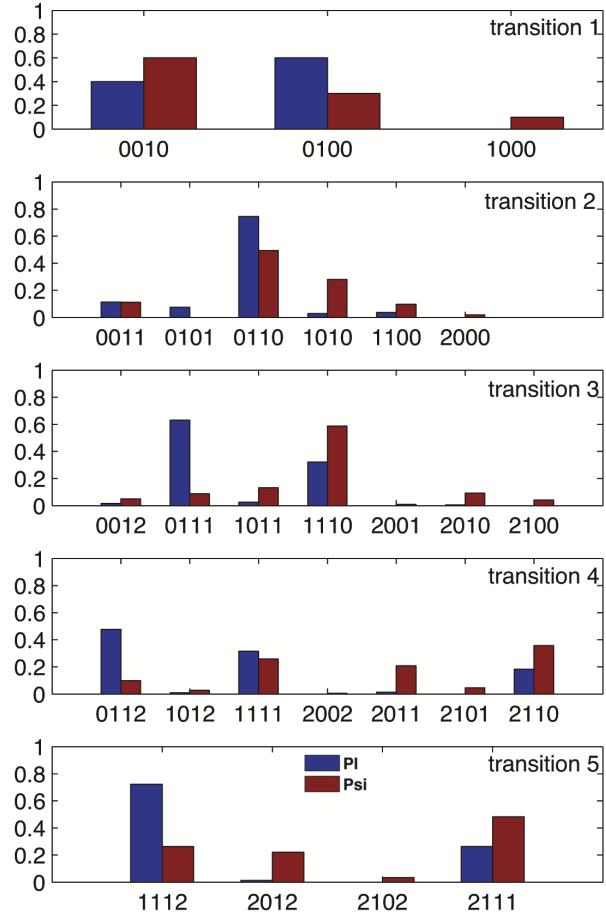


Figure 5: Probability distributions over the configurations after each transition for each of the PI and Psi groups.

would be possible to make predictions - one was made in this paper about the effects of using a HMD with a larger field-of-view on one experiment. We believe that this is the first empirically based prediction that has ever been made in two decades of research into the concept of presence in virtual environments. Many others could be made on the basis of this methodology.

The caveats are that we are working with abstract systems - we did not really use a powerwall, for example, but only a simulation of one. We cannot know whether other factors such as changes in resolution would cause any significant differences. However, the advantage of our approach is that we can compare different systems deliberately abstracting away from many confounding effects. It is not possible to do a scientific comparison of one physical system with another, since too many factors change simultaneously, not under experimental control. Moreover, the theory and associated methodology can make predictions, and these can be tested in formal experimental studies with the physical systems under investigation. It should also be noted that we are not attempting here to replace current methods, we provide empirically based response functions π and ψ that describe how the 'average participant' might respond. These can be used to make predictions. In particular studies, presence (PI and Psi) could still be measured using questionnaires and physiological and behavioral responses, and the criterion of 'response as if real'. The results of these experiments could then be compared with predictions of the underlying theory. Finally, we

put forward this methodology as a program of research, where different laboratories could collaborate in order to form an agreed set of probability distributions, based on a much wider sample of data than one lab alone can gather, and using a number of different systems.

Acknowledgements

This work was developed under the FET project PRESENCCIA (27731) and the ERC project TRAVERSE (227985).

References

- BROOKS, F. P. 1999. What's real about virtual reality? *IEEE Comput. Graph. Appl.* 19, 6, 16–27.
- CARROZZINO, M., TECCHIA, F., BACINELLI, S., CAPPELLETTI, C., AND BERGAMASCO, M. 2005. Lowering the development time of multimodal interactive application: the real-life experience of the XVR project. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, 270–273.
- DRAPE, J. V., KABER, D. B., AND USHER, J. M. 1998. Telepresence. *Human Factors* 40, 3, 354–375.
- FAIRCHILD, M. D. 2005. *Color Appearance Models 2nd Ed.*, Wiley-IS&T, Chichester, UK.
- FERNANDO, R. 2005. Percentage-closer soft shadows. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Sketches*, ACM, New York, NY, USA, 35.
- FREEMAN, J., AVONS, S. E., PEARSON, D. E., AND IJSSELSTEIJN, W. A. 1999. Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments* 8, 1, 1–13.
- GARAU, M., FRIEDMAN, D., WIDENFELD, H. R., ANTLEY, A., BROGANI, A., AND SLATER, M. 2008. Temporal and spatial variations in presence : Qualitative analysis of interviews from an experiment on breaks in presence. *Presence: Teleoperators and Virtual Environments* 17, 3, 293–309.
- GARDNER, H. J., AND MARTIN, M. A. 2007. Analyzing ordinal scales in studies of virtual environments: Likert or lump it! *Presence: Teleoperators and Virtual Environments* 16, 4, 439–446.
- GILLIES, M., AND SPANLANG, B. 2010. Real-time character engines comparing and evaluating real-time character engines for virtual environments. *Presence: Teleoperators and Virtual Environments*, in press.
- HELD, R. M., AND DURLACH, N. I. 1992. Telepresence. *Presence: Teleoperators and Virtual Environments* 1, 1, 109–112.
- KILGARD, M. 2000. Improving shadows and reflections via the stencil buffer. Tech. rep., NVIDIA.
- LESSITER, J., FREEMAN, J., KEOGH, E., AND DAVIDOFF, J. 2001. A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators and Virtual Environments* 10, 3, 282–297.
- LOMBARD, M., AND DITTON, T. 1997. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication* 3, 2, online journal.
- MEEHAN, M., INSKO, B., WHITTON, M., AND BROOKS, JR., F. P. 2002. Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics* 21, 3, 645–652.
- MEEHAN, M., RAZZAQUE, S., WHITTON, M. C., AND BROOKS, JR., F. P. 2003. Effect of latency on presence in stressful virtual environments. In *VR '03: Proceedings of the IEEE Virtual Reality 2003*, IEEE Computer Society, Washington, DC, USA, 141–148.
- SANCHEZ-VIVES, M. V., AND SLATER, M. 2005. From presence to consciousness through virtual reality. *Nature Reviews Neuroscience* 6, 4, 332–339.
- SCHUBERT, T., FRIEDMANN, F., AND REGENBRECHT, H. 2001. The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments* 10, 3, 266–282.
- SHERIDAN, T. 1992. Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments* 1, 1, 120–126.
- SLATER, M., AND GARAU, M. 2007. The use of questionnaire data in presence studies : Do not seriously Likert. *Presence: Teleoperators and Virtual Environments* 16, 4, 447–456.
- SLATER, M., AND WILBUR, S. 1997. A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments* 6, 6, 603–617.
- SLATER, M., USOH, M., AND STEED, A. 1994. Depth of presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments* 3, 2, 130–144. Depth of Presence in Immersive Virtual Environments xD;TY - JOUR.
- SLATER, M., ANTLEY, A., DAVISON, A., SWAPP, D., GUGER, C., BARKER, C., PISTRANG, N., AND SANCHEZ-VIVES, M. V. 2006. A virtual reprise of the stanley milgram obedience experiments. *PLoS ONE* 1 (doi:10.1371/journal.pone.0000039).
- SLATER, M., KHANNA, P., MORTENSEN, J., AND YU, I. 2009. Visual realism enhances realistic response in an immersive virtual environment. *IEEE Computer Graphics and Applications* 29, 3, 76–84.
- SLATER, M. 2004. How colorful was your day? : Why questionnaires cannot assess presence in virtual environments. *Presence: Teleoperators and Virtual Environments* 13, 4, 484–493.
- SLATER, M. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc Lond B Biol Sci* 364, 1535 (Dec), 3549–3557.
- TAYLOR, II, R. M., HUDSON, T. C., SEEGER, A., WEBER, H., JULIANO, J., AND HELSER, A. T. 2001. VRPN: a device-independent, network-transparent vr peripheral system. In *VRST '01: Proceedings of the ACM symposium on Virtual reality software and technology*, ACM, New York, NY, USA, 55–61.
- USOH, M., CATENA, E., ARMAN, S., AND SLATER, M. 2000. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments* 9, 5, 497–503.
- WITMER, B. G., JEROME, C. J., AND SINGER, M. J. 2005. The factor structure of the presence questionnaire. *Presence: Teleoperators and Virtual Environments* 14, 3, 298–312.
- ZIMMONS, P., AND PANTER, A. 2003. The influence of rendering quality on presence and task performance in a virtual environment. In *VR '03: Proceedings of the IEEE Virtual Reality 2003*, IEEE Computer Society, Washington, DC, USA, 293–293.