View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Revistes Catalanes amb Accés Obert

Anuario de Psicología
2009, vol. 40, nº 2, 193-210
© 2009, Facultat de Psicologia
Universitat de Barcelona

How we experience immersive virtual environments: the concept of presence and its measurement*

Mel Slater[1]
Beau Lotto[2]
Maria Marta Arnold[3]
Maria V. Sanchez-Vives[4]
[1] *ICREA, University of Barcelona*
[2] *Institute of Opthalmology, University College, London*
[3] *Instituto de Neurociencias-CSIC, Alicante*
[4] *ICREA-IDIBAPS, Barcelona*

This paper reviews the concept of presence in immersive virtual environments, the sense of being there signalled by people acting and responding realistically to virtual situations and events. We argue that presence is a unique phenomenon that must be distinguished from the degree of engagement, involvement in the portrayed environment. We argue that there are three necessary conditions for presence: the (a) consistent low latency sensorimotor loop between sensory data and proprioception; (b) statistical plausibility: images must be statistically plausible in relation to the probability distribution of images over natural scenes. A constraint on this plausibility is the level of immersion; (c) behaviour-response correlations: Presence may be enhanced and maintained over time by appropriate correlations between the state and behaviour of participants and responses within the environment, correlations that show appropriate responses to the activity of the participants. We conclude with a discussion of methods for assessing whether presence occurs, and in particular recommend the approach of comparison with ground truth and give some examples of this.

Key words: *presence, immersive virtual environments.*

---

* *Acknowledgements*: The work discussed in this chapter arose out of the Presencia Project, EU FET Program, PRESENCIA IST-2001-37927, and the PRESENCIA Integrated Project Contract Number 2773
Author's address: Mel Slater, Event Lab. Departamento de Personalidad, Evaluación y Tratamientos Psicológicos Campus Mundet. Edificio del Teatro. Universitat de Barcelona. Passeig de la Vall d'Hebrón, 171. 08035 Barcelona. Email: melslater@ub.edu
Original recibido: octubre 2008. *Aceptación final:* mayo 2009.

Como experimentamos los entornos virtuales inmersivos: el concepto de presencia y su medición

En este artículo se revisa el concepto de presencia en entornos virtuales inmersivos; es decir, la sensación de estar dentro del entorno virtual indicada por el modo de responder al mismo como si fuera real. La presencia debe distinguirse de otros fenómenos como el compromiso o la implicación. Hay tres condiciones necesarias para la presencia. La primera es un bucle sensomotor consistente y de baja latencia entre los datos sensoriales y propioceptivos. La segunda es la plausibilidad estadística, referente a que las imágenes deben ser estadísticamente plausibles en relación con la distribución de probabilidad de las imágenes en escenas naturales. Un límite para esta condición viene dado por el nivel de inmersión. La tercera es la correlación entre el comportamiento del sujeto y la respuesta del entorno. La presencia se mantiene e incrementa a lo largo del tiempo como consecuencia de la correlación entre, por un lado, el estado y el comportamiento del sujeto y, por otro, las respuestas del entorno, indicando que el entorno responde de forma adecuada a la actividad del sujeto. Se concluye con una discusión de los métodos que se pueden emplear para evaluar la presencia y se recomienda para ello, en particular, la comparación con datos obtenidos sobre el terreno; es decir, la comparación entre las respuestas del sujeto ante estímulos virtuales y las respuestas ante los correspondientes estímulos reales.

Palabras clave: *presencia, entornos virtuales inmersivos.*

In this paper "presence" is considered as the propensity of people to respond to virtually generated sensory data as if they were real (Sanchez-Vives & Slater, 2005). This encompasses their ability to act within the environment created by virtually generated sense data in a manner commensurate with how they would be able to behave if the sensory data were real. If they see an object on the floor below them, and wish to lift it, then they should be able to bend down, grab it, feel it, feel its weight, and lift it, see their limbs (or representations of these) in the act of lifting, with proprioception matched to sensory data. High presence does not demand high fidelity to physical reality, but rather that people do respond, and be able to respond, as if the sensory data were physically real. This approach makes "presence" directly observable and measurable –both with respect to observations of others, and with respect to knowledge of one's own behaviour–.

The vast majority of research on presence has represented the concept as a subjective state or feeling, including the notion of "being there", that is accessible and measurable by questionnaires, as discussed for example in Draper *et al.* (1998). Here we present a quite different view, treating presence as rooted in activity, the response of people to their surroundings and their ability to actively modify those surroundings (Flach Holden, 1998; Zahorik & Jenison, 1998). Another way to state our definition in the opening paragraph is that presence arises when there is successful substitution of real sensory data by virtually generated sensory data. The substitution is "successful" to the extent that the participant in a virtual or mixed reality forms percepts from the

© 2009, Universitat de Barcelona, Facultat de Psicologia

sense data and responds to and acts upon these as if they were real. This response is at many levels, ranging from unconscious physiological processes such as brain activation states, heart activity, breathing, skin response, through unconscious automatic behaviours and reflexes, through deliberate volitional behaviour, through to the highest level cognitive behaviour - including a 'feeling of being there'. The advantage of this definition is that it is operational, and it applies equally well to mixed reality environments as well as pure virtual reality. The "sense of being there" is but one small part of this, which in any case only applies to a virtual reality and not, for example, to augmented reality. We use the term 'reported presence' to mean that particular aspect of presence measurement that refers to people's post-hoc subjective reporting of what they felt during the experience. Hence all questionnaires at best capture aspects of 'reported presence' rather than presence.

It is important to note that this definition automatically includes interaction, and the ability to interact with the environment, especially the ability to use one's whole body in interaction where appropriate (e.g., move by really walking). The requirement to 'respond as if it were real' naturally involves interaction, the ability to pick up objects, move through an environment, avoid obstacles, and so on–these are all part of what constitutes the potentiality for a response as if the virtual sensory data were real.

Presence and Immersion

The distinction between the terms 'immersion' and 'presence' has been discussed several times before (Draper *et al.*, 1998; Slater, 1999; Slater & Wilbur, 1997). We reserve the term *immersion* to stand simply for what the technology delivers from an objective point of view. The more that a system delivers displays (in all sensory modalities) and tracking that preserves fidelity in relation to their equivalent real-world sensory modalities, the more that it may be described as being "immersive". This is something that can be objectively assessed, based on technical parameters used to describe a system. What field of view, what frame rate, what latency, is it stereo, does it have head-tracking, what haptics is supported, and so on. This is a very different matter to the human response to experiencing environments that such systems deliver. This distinction is similar to that in colour science. A colour can be described objectively in terms of a wavelength distribution. However, the *perception* and *emotional response* to a colour is an entirely different matter –and includes the notion, for example, of metamers, where objectively different wavelength distributions are perceived as the same colour by human observers–. So immersion is analogous to "wavelength distribution" –in principle it can be objectively assessed (though we may not always know how to do this)–. Following through the analogy with metamers, different immersion systems may have indistinguishable perceptual impacts on people in terms of presence.

If immersion is analogous to wavelength distribution in the description of colour then "presence" is analogous to the perception of colour. Presence is a

© 2009, Universitat de Barcelona, Facultat de Psicologia

human reaction to immersion. Given the same immersive system, different people may exhibit different levels of presence, and also different immersive systems may give rise to the same level presence in different people. Presence and immersion are logically separable, but clearly there is a relationship between them that can be discovered empirically. Part of the study of presence is to understand this relationship, and there have been many studies that have attempted this, albeit limited in scope to reported presence.

It should be clear from this discussion that presence and immersion are not the same. Remember that presence is a "response" to a system of a certain level of immersion. In order to achieve presence we could follow two different paths. The first would be to construct a system that has such a high fidelity to reality that it becomes indistinguishable from reality. A more interesting approach is to use knowledge of the perceptual system to find out what is important in our representations of reality –to deliver presence even when the level of immersion is not high. People may achieve presence with wire frame computer graphics, some approximation to auditory fidelity, low resolution, and so on. How does this work? This is a real scientific question for presence research. Knowing the wavelength distribution of light emitted from a surface informs us something about how it may be perceived in terms of colour, but it is far from the whole story. Understanding the human perceptual response to the wavelength distribution is critical in understanding colour. We know that, for example, that it is conceptually possible to reproduce the entire spectrum of perceivable colours (taking into account metamers) just by additively combining three primaries. This latter property (reducing the function space of wavelength distributions to the three dimensional space of perceivable colours) is only possible because of the way that human perceptual system works. Similarly, our anecdotal experience of virtual reality convinces us that presence can be achieved with systems that are extreme in their paucity compared to the infinitely greater detail available in perceptions of real life. We hypothesise that just as a complex wavelength distribution can be "simulated" in terms of colour perception by an appropriate additive combination of three primary colours, so the presence in a real life situation can be simulated by a virtual reality that delivers the right combination low fidelity sensory data in relation to physical reality. The question then is: what would be the equivalence to the "primary colours" in colour perception, the minimal set of elements that when combined give rise to presence in the same way that combinations of primaries can give rise to the sensation of colour?

Presence and Attention, Emotion, Involvement, Engagement and Others

There are many other concepts that are confounded with presence, whether just confused through use in everyday language or even forming part of official questionnaires –for example, «How much did the visual aspects of the environment involve you?» (Witmer & Singer, 1998)–. We can separate *involvement* from *presence*, it is at a different logical level. One can be present

© 2009, Universitat de Barcelona, Facultat de Psicologia

but not involved (as in many situations in everyday life). One can be involved but not present (e.g., watching a soap opera on tv). This does not say that we should confuse "involvement" with presence. In physical reality one can study how much different situations "involve" people. One can also do the same in virtual reality. However, this is not the same as studying presence. Listening to some quadrophonically broadcast music you might say «This is just like being in the theatre listening to the orchestra –but as for the music itself– it just doesn't interest me». This is high presence, low involvement (or interest).

Presence is orthogonal to emotional content. In physical reality you can be in a situation that results in a pattern of emotional responses. Given a similar situation portrayed within a virtual environment, the question to ask is how similar is the pattern of emotional responses generated within the VE to that which was generated in physical reality? In other words, how much is the response to the virtual sensory data as if it were real, on the emotional level. This does not imply that highly emotional events are "more presence inducing" than events that have little emotional ramifications. The question is the similarity of distribution of emotional responses, whatever the actual emotional responses may be. Moreover it is extremely difficult, and perhaps methodologically impossible, to distinguish between presence and emotion by the use of questionnaires. If participants in an experiment are subject to two different scenarios, one that is emotionally charged and the other that is emotionally neutral, of course they are going to say that the first resulted in more "presence" than the second. But here "presence" and "emotion" are deeply confounded. It is simply asking the wrong question to consider whether there is a link between "emotion" and "presence" in a virtual environment (Freeman *et al*., 2005c). Emotion, in our view, is part of the response to a virtual environment, and the important issue is whether the response conforms to what would be expected if the participants were responding to the situation as if it were real.

Correlational Presence

In a review of the concept of presence in virtual environments and associated literature (Sanchez-Vives & Slater, 2005) it was reported that experimental studies have found that factors that contribute to high reported presence are mainly concerned with the form of *how* data is displayed to participants and *how* they are able to interact, rather than with the level of realism of the displays. For example, wide field of view, low latency, high frame-rate, surround sound, haptic feedback, stereo rather than mono displays, head-tracking all seem to contribute to higher reported presence. In addition the ability to interact with the environment making use of whole body interaction in a natural way appears to favour higher reported presence in comparison with "button-press" types of interaction more suited to two-dimensional displays. What does not seem to be important is high fidelity visual realism. For example, a person giving a talk in an immersive virtual environment to an audience of virtual characters who appear to be responding to the talk is likely to react to the be-

© 2009, Universitat de Barcelona, Facultat de Psicologia

haviour of the virtual characters as if they were real –in spite of the absolute knowledge that there is no audience there, and in spite of the low level of realism of the characters (Pertaub *et al*., 2002). How is this possible?

It seems that humans have a propensity to find correlations between their activity and internal state and their sense perceptions of what is going on "out there". In studies that involve interactions between real people and virtual characters, people impute meaning to entirely contingent events. For example, as they look at a virtual character and by chance the character turns its head, people will often say something like "Every time I looked at that person she looked away" and there are many examples of this type of behaviour. This propensity to construct such stories is so strong that it was realised that virtual environments may provide an ideal laboratory for the study of paranoid ideation, and several successful experiments have now been conducted in this area (Freeman *et al*., 2008a; Freeman *et al*., 2005a; Freeman *et al*., 2005b; Freeman *et al*., 2008b; Freeman *et al*., 2005c; Valmaggia *et al*., 2007). This is one aspect of what we refer to in this chapter as *correlational presence*. An important aspect of our approach is that presence may be enhanced through statistical correlation between activity and sensory feedback at many different levels. A simple example is the required correlation between the proprioceptive feeling of head turning and the concomitant update in the displays due to low-latency head-tracking. In the interaction between people it is such correlational behaviour that establishes the reality of the situation – matched changes in posture and stance, listeners nodding (or shaking) their heads while someone is talking –a kind of unsupervised dance emerges between the participants that links them together into a higher order social entity–. In "collaborative virtual reality" applications, it is the very absence of such correlational cues that undermines the establishment of shared presence (Garau *et al*., 2005; Slater & Steed, 2000).

Correlational presence requires that when people interact their facial expressions, motor behaviours, and overall aspect including skin colour changes correlate with the actions of one another. This needs to occur with overt semi-volitional behaviour (such as a person nodding when someone talks to them, or correlated shifts in posture) (figure 1) through unconscious non-volitional behaviour, such as correlations between breathing or heart rate in one person with the actions of another person. These are aspects of interpersonal interaction that are barely noticed in everyday life, but have a profound effect when they are missing (Slater *et al*., 2000) (as is always the case to date in interactions between participants or between participants and virtual humans within VEs). This principle can be taken further, that not just the behaviours and states of representations of people correlate with behaviours of the self, but also that this can occur with multiple aspects of the environment. This extends the world of mundane reactions that actually do occur in physical reality, taking advantage of mixed reality to construct environments that adapt to the behaviours, moods and states of the individuals within it. Some preliminary work in this area is reported in (Gillies & Slater, 2005).

© 2009, Universitat de Barcelona, Facultat de Psicologia

A real person interacts with virtual characters

Mimicry is one step towards attribution of "mind"

Figure 1: When people interact with virtual characters. Correlations in behaviour between the real and virtual people are likely to enhance the probability of the real person responding to the virtual people as if they are real.

Correlational Presence and the Statistical Theory of Vision

The idea of correlational presence is rooted in one of the important currents in the neuro-scientific understanding of perception which is the statistical approach (Lotto, 2002; Lotto & Purves, 2001; Lotto & Purves, 2002; Purves & Lotto, 2002; Purves *et al*., 2000; Purves *et al*., 2004; Purves *et al*., 2001). Vision has no direct access to the visual world other than through the dynamic patterns of light that fall on its sensors. But because this sensory information conflates multiple attributes of that world, light stimuli are inherently ambiguous with respect to their underlying sources. Furthermore, the sources of stimuli are themselves behaviourally indeterminant, as they are not imbued with ecological value. As such, what we see is not a representation of the information that arises directly from the world (which is inherently ambiguous), or indeed a representation of the world that generated it (with which the brain has no direct access). Rather, the visual system has evolved to solve this problem empirically by basing its visual behaviour on the probability distribution over the past behavioural significance of retinal stimuli. We "see" correlations, correlations based on our ontogenetic and phylogenetic experience of what particular retinal images signified for visual behaviour.

This general view is not restricted to vision, since all sensory correlations are multi-modal. When participants in an immersive virtual environment place their hand inside something that is red, they will often feel some heat, even though of course nothing is actually generated by the technology. "Empirically" based per-

© 2009, Universitat de Barcelona, Facultat de Psicologia

ception results in a packet of correlated responses that form the totality of a percept –and hence the brain can automatically predict missing data based on past correlations in order to, often literally, "complete the picture"–. This can also be at the basis of the observed fact that people can become present in the most simple representations of reality –the brain is filling in missing detail, provided that there is sufficient feedback (in particular based on head-tracking).

Correlational Presence and the Sensorimotor Loop

One subject on noticing the fixed virtual left arm began to move her real left arm very rapidly, in a manner indicating panic. Another wrote «I thought there was really something wrong with my [left] arm»; ohers talked of their virtual bodies being – «a dead weight», «a useless thing», «nothing to do with me» (Slater & Usoh, 1994b).

> *Something awful's happened —she mouthed, in a ghostly flat voice— I can't feel my body. I feel weird —disembodied* (Sacks, 1998).

When we carried out one of our first ever virtual reality presence experiments in the early 1990s using a head-mounted head-tracked display, people were endowed with virtual bodies, because we had originally found that being "embodied" within a virtual reality led to a greater reported sense of presence compared to looking down and seeing no part of yourself (Slater & Usoh, 1994a; Slater & Usoh, 1994b). However, because in those days we were unable to track more than one arm, the virtual left arm was in a fixed position. Although a participant would see their virtual right arm move in more or less synchrony with their (tracked) right arm, their virtual left arm was locked in place (because there was no tracking data with which to update it). For some subjects this was a profoundly shocking experience, as illustrated in the first quotation above. In a later experiment we tried the idea of making the virtual left arm move in symmetry with the right arm –so that subjects would see their virtual left arm moving even though they might not actually be moving their real left arm–. Surprisingly we found that some subjects immediately made their real left arm movements match what they saw their virtual left arms doing. They needed to maintain a correlation between their motor actions and what their visual system was reporting.

The quotation from Oliver Sacks illustrates the profound disembodiment that occurs when someone loses their proprioception. It was remarkable that our subjects in a virtually induced mismatch between their proprioceptive and sensory data made very similar statements. We argue that "presence" in *physical reality* is only noticeable by its absence, through something that breaks the deep correlation between our motor actions and the response that we experience –a break in the sensorimotor loop– for example through brain damage due to a stroke. A fundamental requirement within a virtual reality is the maintenance of this sensorimotor loop: the continued, predictable correlation between pro-

© 2009, Universitat de Barcelona, Facultat de Psicologia

prioception and sensory data. A head-turn must result in a concomitant and predictable and appropriate change in the visual field, a movement of the body must result in the expected correlated sensory and sensed physical changes that have been learned over a lifetime. This also follows from earlier theoretical and experimental work on presence that led to the concept of "body centred interaction". This expressed the simple idea that a necessary condition for presence in an immersive virtual environment is a match between proprioception and sensory data: participant actions must be immediately and consistently represented by changes in the virtual environment. At the simplest level, for example, as the head turns so the displayed images should be updated in accordance with the changed viewing parameters. The presence literature is rife with situations where this does not happen. For example, it has been found that presence is not reported when there is low graphics frame rate or high latency.

In the statistical theory of vision it is argued that the interaction between our sensory apparatus and the world produces multi-modal "images" from which it is impossible to reconstruct properties of the world. In this view our perceptual system is like a Bayesian decision engine (Purves & Lotto, 2002). Given an "image" (a totality formed from visual, auditory, haptic, olfactory, vestibular, proprioceptive... data) the brain essentially selects its "meaning" and behavioural response according to its statistical correlation with the set of images of natural scenes from the past and their associated behavioural significance. Now when aspects of our sensory data are being generated by a virtual reality, of course the perceptual system operates in exactly the same way. So the meaning of an "image" is interpreted according to its statistical correlation with other images, and the appropriate response selected according to the past behavioural significance of similar images. For example, if the changing patterns of light that fall on the eye as the head is moved around form images that are statistically similar to patterns that had occurred in the past that were interpreted as looking over a precipice, then behavioural responses will be generated appropriate to that situation –even though at the cognitive level the participant knows for sure that there is no precipice there. In the past the evoked response may have always been one of anxiety in relation to such images, and even in a virtual reality this same response may be evoked.

Of course in a virtual reality the "patterns of light" corresponding to a precipice must be correlated with images from past natural scenes at a relatively lower level of correlation than for images corresponding to real scenes. Perhaps this leads to a lesser evoked response. However, there is also likely to be a high degree of within-image correlation –that bundles of features go together, and not just within one sensory modality but across all modalities (as a trivial example, visual images that depict a large object striking the ground are likely to correlated with auditory "images" of a loud sound). Hence the brain may fill-in missing sensory data (statistically this is equivalent to using correlation for prediction, as in regression analysis).

A statistical analogy is to consider k variables $x_1, x_2,\ldots,x_k$ and the multiple regression of each x_j on the remaining $x_1, x_2,\ldots,x_{j-1}, x_{j+1},\ldots,x_k$ based on a large set of past data. Each such analysis will produce a corresponding multiple

© 2009, Universitat de Barcelona, Facultat de Psicologia

correlation coefficient R^2(j). Now suppose (a) that for a particular *j*, R^2(j) is close to 1, but when we observe a new set of values for each variable, $| x_{j-} z_j|$ is large, where z_j is the value predicted from the regression equation. Such a result would be an anomaly. Suppose (b) that we had the observations for each variable *except for* x_j, then we could use the regression equation to predict the likely value of x_j ("filling-in"). Finally (c) if the value of R^2(j) was close to zero, with low statistical significance, then we could tolerate a large range of values of x_j irrespective of the values of the other variables.

If we think now of the observed vector $\boldsymbol{x} = (x_1, x_2,\ldots,x_k)$ as representing a multi-sensory image from a natural scene, then the response associated with $\boldsymbol{x}$ will depend on its correlation with previous examples and their associated behavioural significance. If $\boldsymbol{x}$ corresponds to a virtual reality where at least some of the x_j are virtually generated sensory data we could argue that a necessary condition for presence is when $\boldsymbol{x}$ correlates sufficiently with past image data from natural scenes and therefore a response is evoked that is similar to the response to those natural scenes –in other words, the response to the virtual data is as if it were real. Hence $\boldsymbol{x}$ must be statistically plausible in the context of the probability distribution of images over natural scenes.

In view of (c) above, this does *not* require virtual images that are high fidelity with respect to images of natural scenes –since a very high degree of tolerance may be accepted–. Anyone who has experienced immersive virtual reality knows personally that high presence may be induced even though the visual quality of the rendered scenes are poor in relation to real scenes. Also the evidence from experiments suggests that this is the case. In Usoh *et al.* (2000) subjects who experienced a poor virtual reality simulation of a laboratory did not report lower presence than another group who experienced the real laboratory. In subjects who experienced a precipice (the pit room) in wire frame had the same stress responses and did not report lower presence than those who experienced the pit in full radiosity rendering (Zimmons & Panter, 2003).

This framework also incorporates the idea of breaks in presence (BIP) (Slater & Steed, 2000). A BIP is like (a) above where aspects of the "image" are not in correlation with the rest of the image: e.g., a person walks into a wall of a Cave even though the wall should not be there according to the scene that is displayed in the Cave. It has been argued that incorrect correlations may lead to BIPs: a rose that smells like an apple would be more likely to cause a break in presence than a rose with no smell at all, since in the latter case the correlational properties of perceptual processing may lead to a "filling in" of the missing attribute (Harvey & Sanchez-Vives, 2005). This framework does incorporate the kind of perceptual "filling-in" that has been noticed within virtual environments, which is like (b) above. For example, a person will often "feel" the touch of a sharp object displayed with a virtual environment, or "feel" heat when they immerse their hand inside a red object –even though no tactile sensations are produced by the technology. More significantly non-psychotic people have reported hearing voices in a virtual reality that depicts a social scene even though none were generated by the system –which happened in the study of paranoia reported in Freeman *et al*. (2003).

© 2009, Universitat de Barcelona, Facultat de Psicologia

Necessary Conditions for Presence

In a virtual reality system there are interfaces between our sensory apparatus and the external world –for example, the display systems (stereo glasses, head-mounted display) and their properties (resolution, field of view, stereo, colour resolution and so on) and interaction devices (the extent of head and body tracking), the extent to which sensory modalities are generated at all (sound, touch smell) and of course the rendering algorithms and interaction paradigms that are supported–. Each of these impose constraints on perception. Given a narrow visual field of view with low spatial and colour resolution it *impossible* for images to have a high correlation with those in natural vision. In other words the level of immersion imposes *constraints* on how well the perceptual apparatus is able to obtain images that could correlate with those of natural scenes. Immersion, recall, refers to the number of channels of sensory data generated and the degree to which the sensory data is delivered in a form that conforms with the capabilities and expectations of our sensory apparatus. For example, visual data delivered with a wide field of view, in stereo, auditory data delivered in surround sound according to our specific head-related transfer function (HRTF) would be aspects of a high level of immersion, and so on. These are objectively measurable characteristics of a system that immersion determines the space of possible multi-sensory images.

The initiation of presence in a virtual or mixed-reality environment therefore has as a necessary condition at least two components: first the fundamental sensorimotor loop, the correlation between sensory data and propriopriocep-tion as discussed above. The second is a level of immersion that permits the generation of images that are plausible in comparison to the probability distribution over images of natural scenes, and thus can lead to responses to virtual sensory data as if it were real.

There is a third condition at a higher level of perceptual and cognitive processing which is more to do with the maintenance of presence over time rather than its initial generation. An experiment included an attempt to understand what happens with presence over time (Garau *et al*., 2008). Participants reported that their presence was very high on first going into the environment (a virtual bar scene displayed within a Cave system) and seeing solid life-sized virtual characters in stereo who looked at, gestured, changed facial expressions and talked to them. However, over time, as it was realised that these characters did not answer their questions and engaged in repetitive behaviours that did not correlate with the behaviour of the participants, this initial presence diminished. So although the "images" were sufficient to evoke an initial presence response, they were not sufficient in themselves to maintain this over time.

Thus there is a third level of correlation –between a participant's actions and the reactions to this, beyond the simple perceptual level–. This is an obvious point but overlooked in discussions of presence, since in everyday life these correlations are so obvious that they are not consciously perceived. There are many examples, ranging across the many different levels of experience, for example:

© 2009, Universitat de Barcelona, Facultat de Psicologia

1. A person's body has a shadow and it moves as they move.
2. A person's body is reflected in specular surfaces.
3. Objects that represent sentient beings (virtual humans) should respond and initiate exchanges appropriately in the context of interactions –with nods, glances, breathing changes, forward and back movements, and so on–.

Points 1 and 2 have recently been investigated and it was found that the display of real-time shadows and reflections including those of the person's virtual body significantly added to subjective presence and appropriate physiological response (Slater *et al*., 2009). In multi-person systems (where the other people are virtual) typically 3 is not implemented. Typically you can walk through a virtual character, it will not respond or try to maintain distance, it will not match shifts in your bodily posture, it will not synchronise breathing with you (often a sign of high rapport in real person-to-person communications) and so on. These correlations are so pervasive in everyday life that we only notice them through absence.

To summarise this section: our definition of presence is that it is successful combination of real sensory data and virtually generated sensory data (or in the case of virtual reality, replacement of real sensory data). It is successful when participants respond to the environment and events within it as if they were real. The response to be considered is multi-level: ranging from automatic unconscious physiological behaviours to conscious volitional actions. In between are responses that are semi-conscious, such as breathing rate, changes in gaze direction, shifts in body orientation and posture, and so on.

We argue that necessary conditions for presence are:

Condition 1. The sensory motor loop: a consistent low latency sensorimotor loop between sensory data and proprioception.

Condition 2. Statistical plausibility: images must be statistically plausible in relation to the probability distribution of images over natural scenes. A constraint on this plausibility is the level of immersion.

Condition 3. Behaviour-response correlations: presence may be enhanced and maintained over time by appropriate correlations between the state and behaviour of participants and responses within the environment, correlations that show appropriate responses to the activity of the participants.

Presence Measurement

Presence is a concept without valid measurement. The normal approach is to use *questionnaires*. Participants carry out some task within a virtual environment, and then after their experience they answer a questionnaire. The questions have ordinal scales (Slater & Usoh, 1993; Barfield & Hendrix, 1995; Witmer & Singer, 1998; Lessiter *et al*., 2001) that anchor responses between two extremes –for example 1 meaning "no presence" in the virtual environment and 7

© 2009, Universitat de Barcelona, Facultat de Psicologia

meaning “complete presence”–. The earliest questionnaires were derived from observing and listening to subjects in debriefing interviews (Slater & Usoh, 1993). Some later questionnaires were derived by factor analytic studies from earlier ones (Witmer & Singer, 1998; Lessiter *et al*., 2001; Schubert *et al*., 2001; Witmer, 2005).

Questionnaire based presence assessment methods have been shown to be unstable in the sense that prior information can change the results (Freeman *et al*., 1999). There is also evidence to suggest that typical questionnaires cannot discriminate between presence in a virtual environment and physical reality (Usoh *et al*., 2000). The use of questionnaires has been challenged through the observation that they cannot avoid a methodological circularity –that the very asking of questions about “presence” may bring into being, post-hoc, the phenomenon that the questionnaire is supposed to be measuring (Slater, 2004).

A second method for measuring presence is *behavioural*. If participants within a virtual environment behave as if they were in an equivalent real environment then this is a sign of presence. Examples include the looming response (Held and Durlach, 1992), postural sway (Prothero *et al*., 1995; Freeman *et al*., 2000), after-effects (Welch, 1997) and conflicting multi-sensory cues (Slater *et al*., 1995). These behavioural measures typically require the introduction of features into the environment that would cause a bodily response (such as swaying in response to a moving visual field, or ducking in response to a flying object).

A particular specialisation of the behavioural approach is to use *physiological measures*, such as those derived from ECG recordings or electrodermal activity. The idea in this case is that if the normal response of a person within physical reality to a particular situation is known and they exhibit the same response within a virtual environment then this is a sign of presence. The use of physiological measures as surrogates for presence has been attempted –but have been limited to situations where the physiological response is obvious (e.g., such as a response to a feared situation) and the results have been positive (Meehan *et al*., 2002)–. The drawback here is that physiological responses to mundane situations such as being in a virtual room which has a table and some chairs are not obvious.

Another method for measurement of presence is based on the idea of eliciting moments in time when *breaks in presence* (BIPs) occur (Slater & Steed, 2000). A BIP is any perceived phenomenon during the VE exposure that launches the participant into awareness of the real-world setting of the experience, and therefore “breaks” their presence in the VE. Examples include gross events such as bumping into a Cave wall, getting wrapped in cables, through to more subtle effects such as revelations that come from seeing a tree as a pixel map rather than a solid object. We proposed a stochastic model that allowed the construction of a presence measure from knowledge of moments in time when participants reported such BIPs. This estimator was shown to be correlated with traditional questionnaire measures.

BIPs plus physiological measures have also been combined. The concept is that a BIP itself can be a disturbing event that may have a physiological signature. We have carried out a complex experiment that explored this notion

© 2009, Universitat de Barcelona, Facultat de Psicologia

with positive results (Slater *et al*., 2006), where deliberate BIPs were introduced into a social virtual environment and physiological measures were recorded throughout. An analysis of a wavelet transformed representation of the skin conductance level could be used to successfully "predict" where the BIPs had occurred in the time series. Moreover a particular pattern of heart-rate change was observed through analysis of interbeat intervals, heart rate variability and power transforms of the signal.

A Measure of Presence Against Ground Truth

Following from our view of presence as response to virtual phenomena as if they were real the most appropriate measurement of presence would be to compare actions and responses of people within a virtual environment compared to their responses to similar situations in real environments. Sometimes there is available data on expectations of people's behaviour –for example, the "pit experiment" (Usoh *et al*., 1999)– there was no need to actually carry out a physical experiment with a real pit, because people's expected behaviour is well-known. The same would apply to many situations in psychotherapy. However, in other circumstances where data is not available, we must establish "ground truth" data for comparison, across multiple levels, of how people respond to virtual events compared to real events. It must be emphasised that this is not a "behaviourist" approach. We are interested across the board –in what people say, in what thoughts they report they had, how their emotions were effected, how their attention was distributed, in addition to observable behaviour, physiological responses, eye-movement patterns, brain activity and so on–. We do not give a higher priority to any of these, and results from these different levels may indeed be contradictory (e.g., people saying that they felt nothing but the physiological and observable behaviourable responses show otherwise). Moreover, this approach also includes what it is possible for participants to do, taking account of the level of immersion. Clearly if people have to use a wand to navigate through the virtual environment their experience is going to be quite different to when they can really walk. But even if they are flying through a fantastic landscape, presence is still an issue –the fundamental question is: do they treat what they are experiencing as if it were real?–.

Some recent studies have used this concept of comparison to ground truth. For example, Bidea and colleagues compared responses of people to ball throwing in real and virtual handball (Bideau *et al*., 2003). There has also been a comparison of eye scan paths in a virtual environment with what would have been expected in a real situation (Jordan and Slater 2009).

We also have carried out a detailed study of 6 subjects responding to a real person or virtual person walking by them, reported for the first time in this paper. The experiment took place in the Cave system (Cruz-Neira *et al*., 1993) –specifically a Trimension ReaCTor– which has three back-projected vertical screens (3m×2.2m) and a floor screen (from a ceiling mounted projector) (3m×3m) controlled by a Silicon Graphics Onyx 2. Participants have their head

© 2009, Universitat de Barcelona, Facultat de Psicologia

position and orientation tracked with an inertial/ultrasonic system (Intersense 900). Lightweight CrystalEyes LCD shutterglasses deliver a stereo view of the virtual world. During the experiment first a cylinder, then a virtual character, and finally a real person, walked into the Cave, stopped and looked at the subject, and then carried on walking out of the Cave. The purpose here was to compare the physiological responses of the subject to the real person walking by with the physiological responses to the virtual person (and comparison of both with the control). The results of the experiment were surprising. Subjects generally reported that they found the encounters strange, and that their response to the virtual character was not similar (subjectively) to their response to the real person. However, when we look at their physiological responses a different story unfolds.

Figure 2 shows the number of Skin Conductance Responses measured for each subject during the appropriate time period that they encountered the cylinder, avatar or real person, and also a response during a baseline period when they were standing in the Cave and nothing was happening. It is seen in each case that their response to the virtual character and to the real person was closer than the response to the cylinder. Wilcoxon (paired) sign rank tests show that the difference in responses to the real person and virtual character were not significant, the response to the cylinder and baseline were also not significantly different, whereas all the other comparisons were significantly different. Moreover with repeated exposures to similar events the SCRs should decrease through to adaptation, but here they increase, lending further weight to these results.

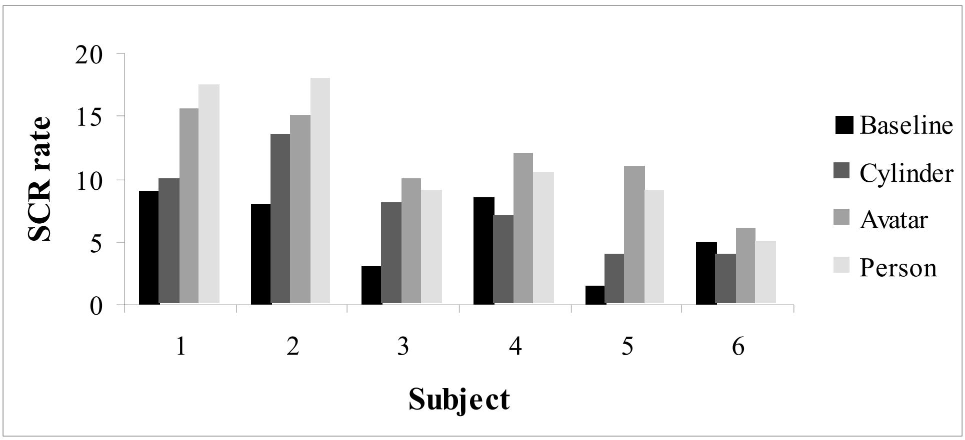

Figure 2: SCR Rate Per Minute for the 6 Subjects by 4 Conditions

We argue that this type of analysis is one component in a comprehensive approach to the measurement of presence, that requires a comparison of observable and subjective responses across multiple dimensions.

© 2009, Universitat de Barcelona, Facultat de Psicologia

Conclusions

This paper has introduced a conceptual approach that we have called "correlational presence". This approach is rooted in the idea of the brain as a correlation engine, exploiting the idea of 'the brain's remarkable ability to detect statistical correlations in sensory inputs in constructing useful perceptual representations of the world…'(Armel & Ramachandran, 2003). From this conceptual approach we have derived a notion of presence as the extent and capability of participants in a virtual environment to respond to virtual situations and events as if these were real. On this basis we argue that analysis of multi-level responses and actions of people within a VE provides a scientific foundation for the study of presence, through the comparison of these responses with known or expected responses to similar situations in reality. This does not rule out the notion of presence in fantastic non-real situations, but only that in order to understand the basis of successful presence in the first place, comparison with expectations is an essential starting point. Once we begin to understand how the form in which sensory data is displayed, and the form of interaction necessary to maintain high presence, these same ideas can be applied to content that is beyond reality.

We have a strong belief that there is too much research in this field that is non-productive: recycling and comparison of questionnaires, philosophical discussion about the true meaning of presence, and so on. It is far more difficult and time consuming to collect detailed and comprehensive data about what people actually do and how they respond within VEs. The experiments briefly detailed above on BIPs and physiological measures, and responses to a virtual character took several person-years of work to carry out and complete, for example, and there remain masses of data that are yet to be analysed (Friedman *et al*., 2005). However, we feel that if the study of presence is to get off the ground and develop as a real science then this kind of detailed difficult data collection and analysis programme is essential, as complex and time-consuming as it may be.

REFERENCES

Bideau, B., Kulpa, R., Menardais, S., Fradet, L., Multon, F., Delamarche, P. & Arnaldi, B. (2003). Real handball goalkeeper vs. virtual handball thrower. *Presence-Teleoperators and Virtual Environments, 12,* 411-421.

Cruz-Neira, C., Sandin, D.J. & DeFanti, T.A. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, 135-142.

Draper, J.V., Kaber, D.B. & Usher, J.M. (1998). Telepresence. *Human Factors, 40,* 354-375.

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

Freeman, D., Slater, M., Bebbington, P., Garety, P.A., Kuipers, E., Fowler, D., Read, C.M., Jordan, J. & Vinayagamoorthy, V. (2003). Can virtual reality be used to investigate persecutory ideation? *Journal of Nervous and Mental Disease, 191,* 509-514.

Freeman, D., Dunn, G., Garety, P.A., Bebbington, P., Slater, M., Kuipers, E., Fowler, D., Green, C., Jordan, J. & Ray, K. (2005a). The psychology of persecutory ideation I - A questionnaire survey. *Journal of Nervous and Mental Disease, 193,* 302-308.

© 2009, Universitat de Barcelona, Facultat de Psicologia

Freeman, D., Garety, P.A., Bebbington, P., Slater, M., Kuipers, E., Fowler, D., Green, C., Jordan, J., Ray, K. & Dunn, G. (2005b). The psychology of persecutory ideation II - A virtual reality experimental study. *Journal of Nervous and Mental Disease, 193,* 309-315.

Freeman, D., Gittins, M., Pugh, K., Antley, A., Slater, M. & Dunn, G. (2008a). What makes one person paranoid and another person anxious? The differential prediction of social anxiety and persecutory ideation in an experimental situation. *Psychological Medicine*, *38,* pp.1121-1132.

Freeman, D., Pugh, K., Antley, A., Slater, M., Bebbington, P., Gittins, M., Dunn, G., Kuipers, E., Fowler, D., & Garety, P. (2008b). Virtual reality study of paranoid thinking in the general population. *The British Journal of Psychiatry, 192,* 258-263.

Garau, M., Slater, M., Pertaub, D.P. & Razzaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence-Teleoperators and Virtual Environments, 14,* 104-116.

Garau, M., Friedman, D., Widenfeld, H., Antley, A., Brogni, A. & 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,* 293-309.

Harvey, M., & Sanchez-Vives, M.V. (2005). The binding problem in presence research. *Presence-Teleoperators and Virtual Environments, 15,* 616-621.

Jordan, J. & Slater, M. (2009). An analysis of eye scan path entropy in a progressively forming virtual environment. *Presence-Teleoperators and Virtual Environments, 18*, (3), pp.185-199.

Lotto, R.B. (2002). Why we see what we do: An empirical theory of vision. London: Sinauer Associates Inc. and Macmillan Press.

Lotto, R.B., & Purves. D. (2001). An empirical explanation of the Chubb illusion. *Journal of Cognitive Neuroscience, 13,* 547-555.

Lotto, R.B., & Purves, D. (2002). The empirical basis of color perception. *Consciousness and Cognition, 11*, 609-629.

Pertaub, D.P., Slater, M. & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence-Teleoperators and Virtual Environments, 11*, 68-78.

Purves, D. & Lotto, B. (2002). Why we see what we do: An empirical theory of Vision. Sunderland (MA): Sinauer Associates Inc.

Purves, D., Lotto, B. & Polger, T. (2000). Color vision and the four-color-map problem. *Journal of Cognitive Neuroscience, 12*, 233-237.

Purves, D., Williams, S.M., Nundy, S. & Lotto, R.B. (2004). Perceiving the intensity of light. *Psychological Review, 111*, 142-158.

Purves, D., Lotto, R.B., Williams, S.M., Nundy, S. & Yang, Z.Y. (2001). Why we see things the way we do: Evidence for a wholly empirical strategy of vision. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 356*, 285-297.

Sacks, O. (1998). *The man who mistook his wife for a hat: And other clinical tales.* Touchstone. New York: Simon and Schuster.

Sanchez-Vives, M.V. & Slater, M.. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*, 332-339.

Slater, M. (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. *Presence-Teleoperators and Virtual Environments, 8*, 560-565.

Slater, M. (Ed.) (2005). *Presence 2005: The 8th International Workshop on Presence*. London. University College London: Department of Computer Science.

Slater, M. & Usoh, M. (1994a). Representation systems, perceptual position and presence in virtual environments. *Presence-Teleoperators and Virtual Environments, 2*, 221-234.

Slater, M. & Usoh, M. (1994b). Body centred interaction in immersive virtual environments. En N. M. Thalmann & D. Thalmann (Eds.), *Artificial life and virtual reality* (pp. 125-148). John Wiley and Sons.

Slater, M. & 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*, 603-616.

Slater, M. & Steed, A. (2000). A virtual presence counter. *Presence-Teleoperators and Virtual Environments, 9*, 413-434.

Slater, M., Sadagic, A., Usoh, M. & Schroeder, R.. (2000). Small-group behavior in a virtual and real environment: A comparative study. *Presence-Teleoperators and Virtual Environments, 9*, 37-51.

Slater, M., Khanna, P., Mortensen, J. & Yu, I. (2009). Visual realism enhances realistic response in an immersive virtual environment. *IEEE Computer Graphics and Applications, 29* (3), pp. 76-84.

Usoh, M., Catena, E., Arman, S. & Slater, M. (2000). Using presence questionnaires in reality. *Presence-Teleoperators and Virtual Environments, 9*, 497-503.

Valmaggia, L.R., Freeman, D., Green, C., Garety, P., Swapp, D., Antley, A., Prescott, C., Fowler, D., Kuipers, E., Bebbington, P., Slater, M., Broome, M. & McGuire, P. (2007). Virtual reality and paranoid ideations in people with an "at risk mental state" for psychosis. *British Journal of Psychiatry, 191 (suppl. 51),* pp. 563-568.

© 2009, Universitat de Barcelona, Facultat de Psicologia

Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence-Teleoperators and Virtual Environments, 7*, 225-240.
Zahorik, P. & Jenison, R.L. (1998). Presence as being-in-the-world. *Presence-Teleoperators And Virtual Environments, 7*, 78-89.
Zimmons, P. & Panter. A. (2003). The influence of rendering quality on presence and task performance in a virtual environment. IEEE *Virtual Reality Proceedings*, 293-294.

© 2009, Universitat de Barcelona, Facultat de Psicologia