# Virtually There? A Vision on Presence Research

**Wijnand IJsselsteijn**,

(Eindhoven University of Technology, The Netherlands)

**Bernard Harper**

(Liverpool University, UK)

and the

**Presence Research Working Group**

**PRESENCE - IST 2000-31014**
**EC Public Deliverable (D2)**
**December 2001**
(Updated May 2002)

## Contents

## A Scenario from the Future…

It is 7pm on Wednesday May 28th 2031. You've been looking forward to this evening for the last two months, ever since you finally got around to upgrading your home Immerse-U-There© Technologies communication system as a birthday treat at the start of March.

You've noticed some real benefits since your system upgrade. The early days of teleworking using a home PC and a modem seem so primitive now, just 20 years after it became the modus operandi for the majority of the world's working population. Gone are the days of hoping that your voice communicated the right degree of warmth down the ether to your colleagues (the rudimentary video communications of the 00's didn't allow you to feel like you were really having a face to face meeting with your clients and colleagues).

Today you had a group meeting on the global communication strategy of the natural drinks company you work for. At the meeting was Jaap, the company's head communications strategist, joining direct from his small village just outside Eindhoven. Shigoro, head of marketing for Asia, has just moved to a self-sustaining village just outside Tokyo but that didn't stop him being at the meeting. Judy and Lorraine (heads of marketing communications in North and South America respectively) were in the middle of a trek through the Borneo jungle but they were able to join in too using their portable Immerse-U-There© systems. And finally, Maxine was there. Maxine is your old-timer boss, now in her late 90s, who was resolutely against 'new technologies' changing her work practices until she realised that her new Immerse-U-There© system could keep notes on everything that everyone said in the meeting, generate automatic contact reports and enable her to more or less take up residence at her local golf course, just north of Edinburgh. Constructive meeting all round. Got a bit heated when the ad agency tried to pretend their Immerse-U-There© system was broken as an excuse for joining the meeting late. You've heard that excuse from the agency at least 12 times this year - and you know its been a lie each time. Anyway, the agency came up with the goods just in time. The meeting went on for three hours - lucky really because that is the safe continuous use limit set down by the Global Standards Commission. At 2pm (World Standard Time) everyone tuned out after a hard day's work - your children still don't believe you when you tell them that only 10 to 15 years ago people were working 50 plus hours a week, and that there were different time zones around the world!

The reason you're excited though is that in about half an hour you'll be in Rio de Janeiro for the 2031 World Club Championship Final. Tickets for seats in the ground sold out over ten years ago, but with your new Immerse-U-There© system you literally feel as though you are there. You've got the best seat in the stadium, the smells of Brazil, you can feel the atmosphere (its hot and humid there tonight), you can hear every little whistle or quip that you'd hear if you were really there (and even be buzzed - but not bitten - by the local mosquitoes). And best of all, you can sit and watch with your remote friends (Jim in Hong Kong, Toby in St. Lucia, Patricia in Dubai, and Lucy in Moscow). Your kids might join you for parts of the game, but they spend most of their free time these days watching earth from the Global Space Station (through their personal Immerse-U-There© systems). Your ticket for the game was quite expensive, but this is one of the biggest sports events of the year. And you have to hand it to the system developers, they knew what they were doing. You really do feel like you're in Rio, with your friends. And it really does feel worth it. And the best thing about it all? You've hated flying all your life, and as a result of the Immerse-U-There© systems becoming the world communication standard (in 2022) you haven't had to fly anywhere for seven years. You and your partner still take holidays abroad, but there isn't the pernicious rush to get it over and done with in a fortnight, so you tend to travel by sail boat - after all you can stay in touch and be anywhere at the flick of a switch...

# 1. Introduction: What is Presence?

"*The biggest challenge to developing telepresence is achieving that sense of "being there." Can telepresence be a true substitute for the real thing? Will we be able to couple our artificial devices naturally and comfortably to work together with the sensory mechanisms of human organisms?"*

- Marvin Minsky, 1980

*"What do people do at work? They go to meetings. How do we deal with meetings? What is it about sitting face to face that we need to capture? We need software that makes it possible to hold a meeting with distributed participants -- a meeting with interactivity and feeling, such that, in the future, people will prefer being telepresent"*

- Bill Gates, 1999

With the rapid developments in the area of immersive, multisensory displays and human-machine interfaces, as well as the increased availability of transmission bandwidth, computing power and digital resources, we are able to create and experience reproductions and simulations of reality with an unprecedented sensory quality. However advanced such media systems and services may become from a technological point of view, the social acceptance and uptake (and consequent commercial success) will depend to a large extent on users' perceptions, experiences and responses towards them. Indeed a number of technological advances have become instructive failures because of their inability to connect to real user needs, or to provide true added value in terms of user experiences. A user experience that is of particular relevance in this regard is that of *presence*. Several studies have shown that users report a compelling perception of 'being there' when interacting with virtual environments, advanced broadcast and cinematic displays, teleoperation systems, and advanced telecommunication applications.

Presence is the experience of projecting one's mind through media to other places, people and designed environments. Appropriate presence technologies combine to create an *illusion of non-mediation* (Lombard & Ditton, 1997) - the closest possible approximation to "being there" without physical travel. The experience of 'being there' in a mediated environment is a subjective experience that has proven to be relevant for the design and evaluation of various interactive and non-interactive media systems. In particular through immersive, interactive and perceptually realistic media, user experiences and responses are provoked that are similar to those in nonmediated environments (see e.g. Lombard, 1995; Freeman, Avons, Meddis, Pearson & IJsselsteijn, 2000; IJsselsteijn, de Ridder, Freeman, Avons & Bouwhuis, 2001).

There is consensus that the experience of presence is a complex, multidimensional perception, formed through an interplay of raw (multi)sensory data, perceptual-motor activity and various cognitive and emotional processes – an experience in which attentional factors play a crucial role as well (Draper, Kaber & Usher, 1998; Witmer & Singer, 1998). Figure 1 depicts a diagram of media experience in general, and presence in particular.

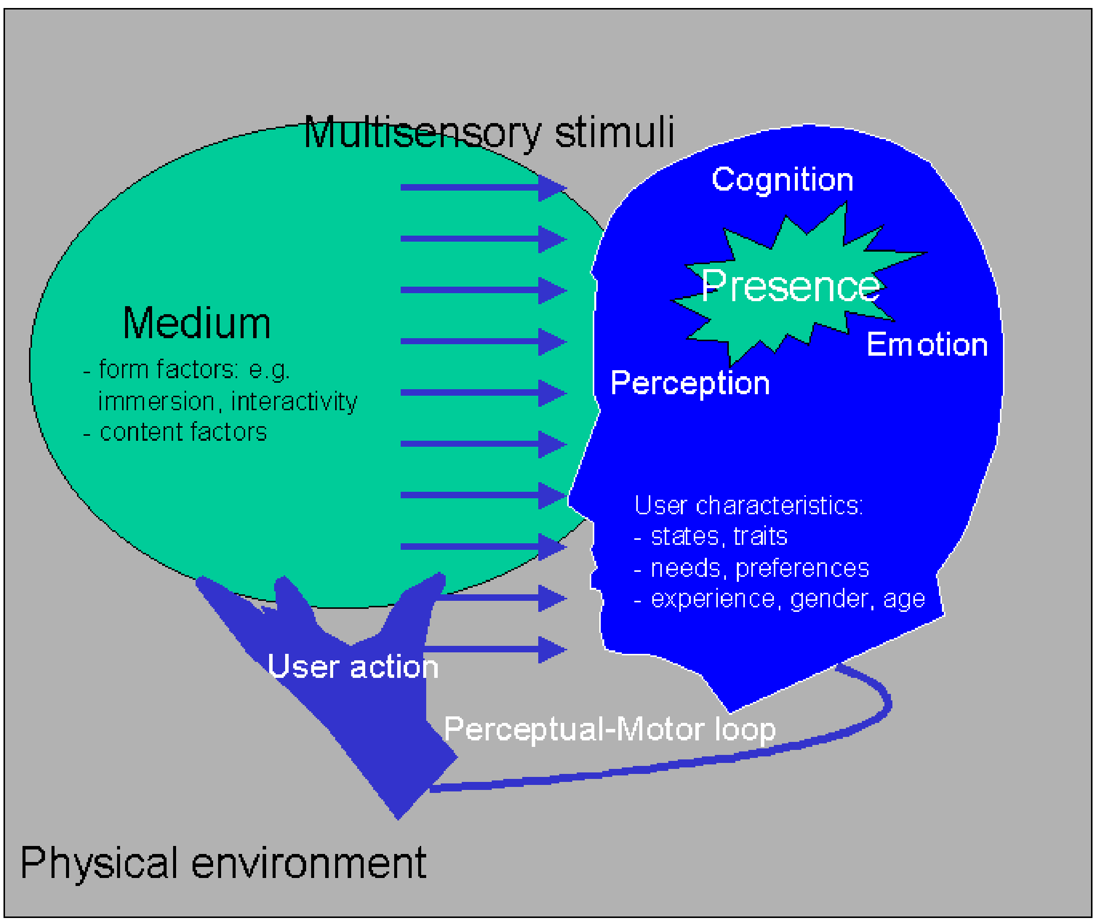


**Figure 1**. A general framework of presence

Factor analytic studies are starting to shed light on the multidimensional structure of presence. In particular, studies by Schubert, Friedman, and Regenbrecht (1999) and Lessiter, Freeman, Keogh, and Davidoff (2000) reveal very similar factor structures. Schubert et al. (1999) arrived at a 3-factor solution for the presence construct, which they termed ‘spatial presence’, ‘involvement’, and ‘realness’. Similarly, Lessiter et al. (2000) reported a 4-factor solution for presence, with three factors almost identical to the ones identified by Schubert et al.: ‘physical space’, ‘engagement’, and ‘naturalness’, and a fourth attenuating factor they termed ‘negative effects’.

As we move towards increasingly realistic media, each new development in the history of visual media can be viewed as a gradual buildup of perceptual cues that simulate natural perception and enhance the experience of presence (Biocca & Delaney, 1995). Early perspective paintings, dating back to the middle ages, only included static monocular depth information which violates most of the critical features of real-world perception. The end of the 18th century saw the introduction of panorama paintings, which stimulated large portions of the visual periphery, a principle that was also applied to great effect in the cinema of the 1950s (Cinerama, CinemaScope), and in more recent large film formats (e.g. IMAX), projection-based VR (e.g. CAVE), and head-mounted displays. The stereoscope of the 19th century allowed each eye to view the same scene from a slightly different perspective (i.e. stereoscopically), contributing greatly to the perception of egocentric distance and exocentric depth within an image.

With the introduction of cinema in the late 19$^{th}$ century, motion has been added to high-resolution photorealistic imagery as a fundamental perceptual cue. The visual system is highly motion-sensitive, and the onset of motion cannot be ignored - it demands attention and automatically elicits an orienting response. Certain camera movements provide motion parallax as a cue to depth, although it is important to note that observer movement does not transform the image appropriately. In the case of head-mounted virtual environments this viewpoint-dependent transformation is possible in real-time, although with the current state of technology real-time interactivity trades off against photorealism. Importantly, it is not clear at present how much each feature or perceptual cue contributes to the perceived realism of media, or to eliciting a sense of presence for the participant. This is an empirical question which will need to be addressed through further research.

As we are approaching the stage at which it becomes possible to develop media that can simulate the majority of critical features of natural perception (including real-time interactivity appropriately based on observer movement), the media experience is transformed from looking at pictures, to experiencing a space, to eventually visiting it as a place and being able to act within it. Despite the rapid developments it is a considerable challenge to design a realistic multisensory media environment with which we can interact intuitively and where participants can not only look, but also hear, touch, smell, taste, walk, run, pickup objects, etc.

The design goal of achieving optimal presence and making auditory, visual and haptic communications believable requires an interdisciplinary approach, integrating knowledge and ideas from disciplines such as neuroscience, multisensory perception, cognition, artificial intelligence, multimedia development, video compression or telecom engineering. In order to build future systems which can efficiently transmit remote presence, it will be necessary to incorporate and integrate ongoing insights from these fields into next-generation research for advanced, wideband multisensory services and novel telecommunications architectures. Such research would include enhancements that could amplify human perception, as well as ameliorate mediating distortions. It will allow for vastly improved mediated experiences that would bring many benefits in a number of surprisingly different ways.

## 2. Why study presence?

Historically, research interest in presence has mainly been motivated by work in three related domains: teleoperation, simulation and telecommunication. Teleoperation has been driven to a large extent by the *ideal* of telepresence, i.e. sensing sufficient information about the teleoperator device and remote task environment, and communicating this to the human operator in a sufficiently natural way, that the operator feels physically present at the remote site (also known as distal attribution).

Advanced simulation systems, such as those used for flight and combat training, have long been known to engender a sense of presence in their participants, and although the relation of presence to task performance within the simulated setting has been a matter of some debate, there is little doubt that training transfer to real-world settings benefits from a realistic simulation. More recently, computing power and rendering possibilities have dramatically increased and have made it possible to create virtual environments of ever increasing realism, at a fraction of the cost of professional simulation systems. Virtual environments can be made so convincing that they can lead to a feeling of presence in the created environment (see e.g Slater & Wilbur, 1997).

From a telecommunication perspective, social presence, or the feeling of being together, has received considerable attention for several decades already (see e.g. Short, Williams & Christie, 1976). The emergence and proliferation of email, mobile communication devices, internet chatrooms, shared virtual environments, advanced tele-conferencing platforms and other telecommunication systems underlines the importance of investigating the basic human need of communication from a multidisciplinary perspective that integrates media design and engineering, multisensory perception, and social psychology. Add to this the increasingly social nature of interfaces and the increase in mediated communications with non-human entities (avatars, embodied agents), it becomes abundantly clear that we need to develop a deeper understanding, both in theory and in practice, of how people interact with each other and virtual others through communication media. The experience of social presence within different contexts and through different applications thus becomes a concept of central importance.

A common theme linking the above examples is that presence research offers the possibility to engineer a better user experience, to optimise the effectivity, efficiency and pleasurability of the different applications. Moreover, the user-centred perspective of presence research will diminish the likelihood of expensive technology failures, such as the introduction of AT&T's PicturePhone in the 1960s and 1970s (Noll, 1992) or .....(see Loretta's doc)

From an application viewpoint, presence research will spur the development of numerous tele-applications in home and professional environments. Such tele-applications have the potential to support a great number of activities, including communication, work, education, shopping, banking, health monitoring, and entertainment. Inasmuch as these tele-applications will partially make redundant the necessity for travelling physically, they will contribute to a more sustainable use of limited resources and will put less strain on the environment in terms of pollution (see figure 2). In addition, they will enable individuals that are physically unable to travel to actively participate in activities previously beyond their reach.

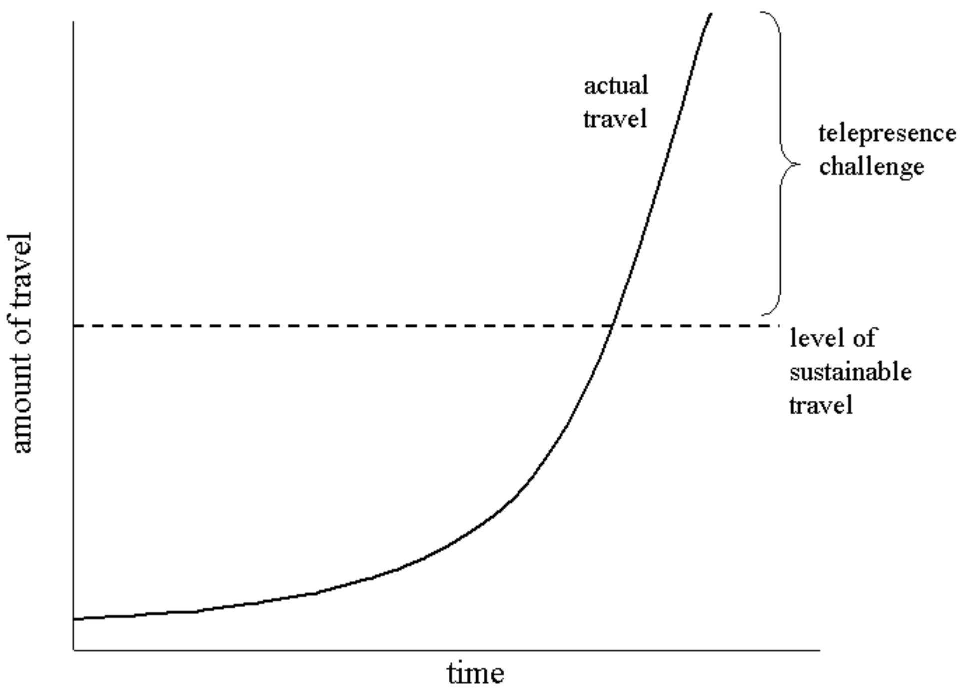


**Figure 2**. Telepresence technologies may potentially provide a solution for the challenges posed by the increase in physical travel beyond the level of environmental sustainability.

Another recent area of application with the potential of being extremely beneficial to a large number of people is the use of VE technology in the service of psychotherapy, in particular for the treatment of phobias (e.g. fear of heights, fear of spiders, fear of flying, etc.). As VEs allow full control over the feared stimulus (in contrast to in vivo exposure) at a predictable, relatively low cost (one major expenditure to equip the consulting office can be offset against the use of the equipment for many different situations, types of environment and patients), this has stimulated the use of VEs in phobia treatment programmes. However, for a virtual representation to provoke the same responses as the real situation we need to know what kind of VE is needed – is a photograph enough, is a movie enough? Also the answers to these questions might vary during the course of a treatment. Putting this another way: what is necessary in order to enhance sufficient *presence* in the VE that is representing the feared situation? What are the fundamental components in a representation that is necessary in order to achieve this sense of presence? How much like a real spider does the virtual spider have to look and behave? Presence, therefore, is of paramount importance in the successful exploitation of VEs in psychotherapy, and, of course, a successful answer to these questions in the context of psychotherapy may well have significant spin-offs for applications in other areas.

A more fundamental reason for studying presence is that it will further our theoretical understanding of the basic function of mediation: How do media convey a sense of places, beings and things that are not here? More importantly even, presence research provides the necessary bridge between media research on the one hand and the massive interdisciplinary program on properties of perception and consciousness on the other (Biocca, 2001).

Creating virtual worlds is an age-old desire of mankind, that traditionally was addressed by the world of art and imagination. It has brought the world of entertainment, the theatre, the movies, which have found their way by distribution and networks into the privacy of one's own home. Never before has it been possible to interact with the world of imagination in such a realistic way, that it renders the impression of being there, and creates the feeling of doing. In this sense the basic purpose of presence research is to increase the quality of life.

## 3. Current Issues and Research Challenges

In 1978, the philosopher Daniel Dennett presented us with an interesting and entertaining thought experiment entitled 'Where am I?' as the last chapter of his *Brainstorms* book (Dennett, 1978). Dennett recounts the story of a 'curious episode' in his life where his brain got surgically separated from his body, with each connection between them restored by placing two 'microminiaturized radio transceivers' between each input and output pathway. After the operation he, or rather his body, goes to visit his brain which was placed, in keeping with the best philosophical traditions, in a life-support vat. While looking with his own eyes at his own brain he starts to wonder:

> "*Being a philosopher of firm physicalist conviction, I believed unswervingly that the tokening of my thoughts was occurring somewhere in my brain: yet when I thought "Here I am," where the thought occurred to me was here, outside the vat, where I, Dennett, was standing staring at my brain.*" (p.312)

Dennett reasons that the location of the 'I' he was referring to in the question 'Where am I?' may be related, though not identical, to his *point of view*. He states:

> "*Point of view clearly had something to do with personal location, but it was itself an unclear notion. It was obvious that the content of one's point of view was not the same as or determined by the content of one's beliefs or thoughts. For example, what should we say about the point of view of the Cinerama viewer who shrieks and twists in his seat as the roller-coaster footage overcomes his psychic distancing? Has he forgotten that he is safely seated in the theater? Here I was inclined to say that the person is experiencing an illusory shift in point of view. In other cases, my inclination to call such shifts illusory was less strong. The workers in the laboratories and plants who handle dangerous materials by operating feedback-controlled mechanical arms and hands undergo a shift in point of view that is crisper and more pronounced than anything Cinerama can provoke.*" (p. 314-315)

What Dennett calls ‘an illusory shift in point of view’ nicely conceptualises the central idea of presence. His examples of this illusory shift are well-chosen. Cinerama, which debuted at the Broadway Theatre, New York in 1952, was one of Hollywood’s answers to the growing popularity of television in the early fifties and the resulting decline of sales at the box-office. Cinerama used three 35mm projections on a curved screen to create a 146 degree wide panorama. In addition to the impressive visuals, Cinerama also included a 7-channel directional sound system that added considerably to its psychological impact. The ads for ‘*This is Cinerama’*, the first Cinerama film, containing the famous scene of the vertigo-inducing rollercoaster ride, promised: “You won’t be gazing at a movie screen – you’ll find yourself swept right into the picture, surrounded by sight and sound.” The film’s program booklet proclaimed: “Everything that is happening on the curved Cinerama screen is happening to you. And without moving from your seat, you share, personally, in the most remarkable new kind of emotional experience ever brought to the theater” (Belton, 1992). Phrases such as these were used quite often to promote the immersive and multisensory cinema formats of the 1950’s, including 3D cinema (e.g. “It happens to YOU in three dimensions”), Todd-AO (“Suddenly you’re there…”), CinemaScope, and other formats (Lodge, 2000). Although such statements were sales pitches of the films’ marketing people, they do illustrate the fact that the aim of the cinema experience was to enhance the film’s psychological impact and entertainment value by making the viewer feel part of the movie. The once passive viewer became an ‘active’ participant. As Slater and Wilbur (1997) put it, with reference to virtual environments, the “discontinuity between the place of our current reality and the reality showing through the display” seemed to be collapsing.

Although non-interactive media can become highly immersive and realistic, engendering a compelling sense of presence for the participant, they still assume a rather passive role for the observer. Perception, of course, is more than mere 'sensing' of the environment through our various sense organs and subsequently matching these sensations against passive representations or templates of stored information. Rather, perception is a highly activity-dependent and context-dependent process (both embodied and environmentally embedded) that integrates multimodal sensory data, ongoing actions and intentions, and memory processes. Perception serves the individual's need to control behavior (action) within an environment. By allowing real-time action at a distance (through teleoperation) or in virtual space (through interactive computergenerated environments) the participant is able to control certain aspects of the mediated environment, and, as a consequence, his/her perception of the environment. In this way, the participant will become aware that he or she is an *actor* within the environment and it is likely that this experience of willful control or *feeling of doing* will greatly enhance the feeling of actually being there within the mediated environment – the sense of presence (see e.g. Welch, Blackmon, Liu, Mellers & Stark, 1996).

Dennett (1978) too states that interactive teleoperation engenders a "shift in point of view that is crisper and more pronounced than anything Cinerama can provoke." Remote-controlled manipulators (e.g. robot arms) and vehicles are being employed to enable human work in hazardous or challenging environments such as space exploration, undersea operations, minimally invasive surgery, or hazardous waste clean-up. The design goal of smooth and intuitive teleoperation triggered a considerable research effort in the area of human factors, which lies at the root of today's presence research (Johnson & Corliss, 1971; Sheridan, 1992). In fact, the term *telepresence* was first used in the context of teleoperation by Marvin Minsky (suggested to him by his friend Pat Gunkel) in his classic 1980 paper on the topic (Minsky, 1980).

Minsky's paper was essentially a manifesto to encourage the development of the science and technology necessary for a remote-controlled economy that would allow for the elimination of many hazardous, difficult or unpleasant human tasks, and would support beneficial developments such as the creation of new medical and surgical techniques, space exploration, and tele-working. The questions Minsky posed are still valid today. Although the remote-controlled economy didn't arrive in the way he envisioned, the development of telepresence technologies has significantly progressed in the various areas he identified. In addition, the arrival and widespread use of the internet brings us remote access to thousands of homes, offices, street corners, and other locations where webcameras have been set up (Campanella, 2000). In some cases, because of the two-way nature of the internet, users can log on to control a variety of telerobots and manipulate real-world objects.

It seems fair to say that the concept of presence has today become common currency in areas such as virtual environments, advanced broadcast and cinematic displays, teleoperation systems, and advanced telecommunication applications. Since the early 1990's, a growing community of multidisciplinary researchers has turned its attention to presence, looking at what causes it, how the experience may be measured, and what effects it has on the media user. Despite considerable progress, presence research is still very much in its infancy, with a great number of unresolved issues and challenges that set the research agenda of the presence community. The most significant issues are listed below:

- The *structure of presence* is still largely unclear. What are the phenomenological properties? For example, is it an all-or-none or a graded experience? Are there different forms? What are its psychological dimensions? As discussed previously, recent factor analytic studies are starting to shed light on this latter issue. Although a number of conceptual models of presence have been proposed to date, there has not been a generally accepted *explanatory model of presence* that relates current theoretical and empirical insights in perception, emotion, cognition and action to mediated experience, and has neural plausibility. A theory of presence needs to be established that has clear links not only to psychology and engineering but also to ontology, epistomology, consciousness studies and neuroscience.
- *Measuring presence* in a way that is reliable, valid and robust is still a major challenge, although this issue is currently receiving considerable attention from various research labs on both sides of the Atlantic. A good presence measure will allow engineers and media developers to identify the factors (and trade-offs between them) needed to optimise the level of presence for the media user. Moreover, a good presence measure will allow the research community to further develop its understanding and systematic investigation of the construct, which will in turn enable further refinement of measurement methodologies, and so on.
- A number of potential *determinants of presence* have been identified, and some have been experimentally validated. Studies have until now mainly focussed on medium-related form and content factors such as user-initiated control of the simulation (Welch et al., 1996), stereoscopic presentation (Hendrix & Barfield, 1996a; Freeman, Avons, Pearson, & IJsselsteijn, 1999; IJsselsteijn, de Ridder, Hamberg, Bouwhuis, & Freeman, 1998), head tracking (Hendrix & Barfield, 1996a), field of view (Hatada, Sakata, & Kusaka, 1980; Prothero & Hoffman, 1995), spatialized audio (Hendrix & Barfield, 1996b), response latency (Ellis, Dorighi, Menges, Adelstein, & Jacoby, 1997), and meaningfulness of the simulation (Hoffman, Prothero, Wells, & Groen, 1998). The relative contributions and interactions of these and other determinants are still largely unclear however. The resolvement of this issue

depends to a large extent upon the development of suitable measurement methodologies and a proper theoretical framework.

- The *effects of presence* are to a large extent still unclear. For example, under what circumstances does an enhanced sense of presence aid task performance, or learning and memory?
- What are relevant *individual differences* with respect to presence, and what impact do they have? Slater and colleagues (Slater & Usoh, 1993; Slater, Usoh, & Steed, 1994) have performed some initial studies on user characteristics, such as the extent to which a person’s field dominance (e.g., visual, auditory) affects their sense of presence. Further study is required however to investigate the potential impact of individual characteristics (traits, states, needs, preferences, experience, gender, etc.) on presence.
- What will be the *social consequences* of the introduction of certain high-presence technologies at work or in the home? How will social relationships be affected by the long-term deployment of high-presence communication applications? Are there any significant *health and safety issues* that need to be taken into account? In which *contexts of use* will presence be of most value? How does presence relate to *basic human needs* such as privacy, control, and social contact?

In fact, the defining characteristics of the concept itself are still being discussed, as is evidenced by recent debates at the Presence Workshops, the European Presence Research Conference, and on presence-related newsgroups. Presence is used in different ways by different scholars, each looking at the concept from their own perspective, applying their own emphasis or using their own specific definition. Lombard and Ditton (1997) reviewed a broad body of literature related to presence and identified six different conceptualisations of presence: realism, immersion, transportation, social richness, social actor within medium, and medium as social actor. Based on the commonalities between these different conceptualisations, they provide a unifying definition of presence as the ‘perceptual illusion of non-mediation’, i.e. the extent to which a person fails to perceive or acknowledge the existence of a medium during a technologically mediated experience.

The different conceptualisations of presence identified by Lombard and Ditton (1997) can roughly be divided into two broad categories – physical and social. The physical category refers to the sense of being physically located in mediated space, whereas the social category refers to the feeling of being *together*, of social interaction with a virtual or remotely located communication partner. At the intersection of these two categories, we can then identify 'co-presence' or a sense of being together in a shared space, combining significant characteristics of both physical and social presence. Figure 3 illustrates their relationship with a number of media examples that support the different types of presence to a varying extent. For example, while a painting may not necessarily support physical presence to any great extent (although trompe l'oeil and panorama paintings may be examples to the contrary), interactive virtual reality (VR) technology has the potential to engender a high sense of physical presence.

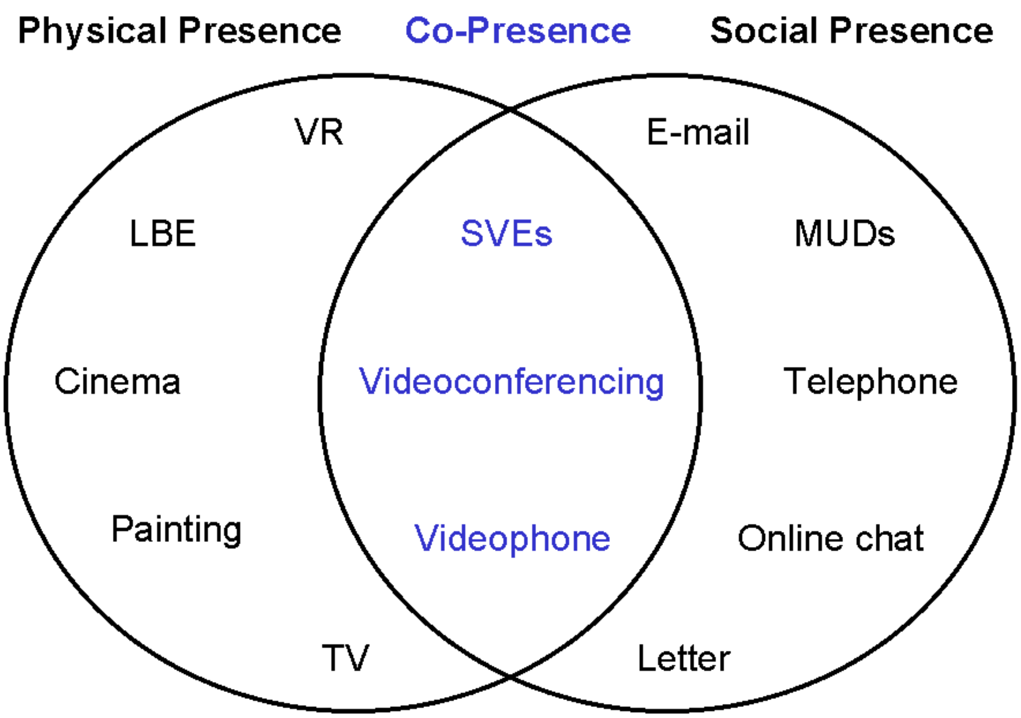


**Figure 3.** A graphical illustration of the relationship between physical presence, social presence and co-presence, with various media examples. Abbreviations: VR = Virtual Reality; LBE = Location-Based Entertainment; SVEs = Shared Virtual Environments; MUDs = Multi-User Dungeons.

## 4. Presence and Perception: Media Problems and Research Solutions

There has been an unchallenged assumption amongst many observers that the technologies underlying the current generation of audio and visual media are soundly based and have been thoroughly developed. However, recent psychophysical and perceptual discoveries strongly suggest that existing photography, film, television, audio recording/reproduction and digital transmission technologies are inadequate in important areas. Not least is the problem that some sensory modalities like olfaction, haptic, vestibular and “somatic” sensations are rarely recorded or replayed at all. There is now evidence these “information transmission” inadequacies can adversely affect very many users. This in turn can be demonstrated to cause inefficiencies, misrepresentations, misinterpretation, and even widespread public health issues.

Our basic understanding of human perception has improved significantly in recent years. But many of the standard psychophysical tests used by the audio and visual industries to design and evaluate their products were produced over 50 years ago and have led to very significant failings in current design and performance. Ask yourself a simple question: “Have I ever seen an image or heard a sound that I instinctively believed to be “real”, when it was in fact produced by a machine?” Surprisingly, it is not only the general public would say “no” to this. It is almost unheard of for someone to walk up to an object and try to touch it, only to discover that it is an image on a screen. And so far as can be discerned, no one has recorded a natural sound and replayed it so that the reproduction cannot be discriminated from the original source. The reason for this is that the information we sense in our immediate environment is currently recorded and replayed in ways that are not fully compatible with the human perceptual systems. The qualitative, quantitative and temporal events that we experience in the real world are only crudely captured and replayed by existing media technologies. If such information is correctly recorded and replayed, then it is likely that people will no longer be able to discriminate between real and virtual stimuli, and they would report that they are in the *presence* of a “real” event or scene.

*Primary sensory stimulation* happens when a person experiences stimuli in reality, rather than through a mediated technology. In addition to the five main senses, our other senses (involving the vestibular/somatic systems) can detect body orientation, acceleration, impacts, temperature, air pressure, water immersion, electrical activity and many others. New sensory capabilities are being discovered with surprising regularity (e.g. Todd & Cody, 2000). All of these sensory inputs need to be combined correctly in the brain to give us the sense that we are present in a real environment and allow us to experience stimuli as if we were really there.

*Artificial sensory simulation* is the process by which current audio/visual technologies re-present captured environmental stimuli to our senses. There are many areas where inadequate environmental information is captured and/or corrupted so that the medium distorts the message in ways that impair its reality. This corruption of environmental information is why we never normally perceive sound and visual images so that they can be easily confused or compared with reality. Our brain's natural ability for pattern recognition means that when current audio/visual technologies provide only a poor approximation of the original stimulus, it can still recognise the source. But it will almost always categorise this source as an obviously artificial reproduction. If mediated artificial stimuli can activate *primary sensory stimulation* (in that it is perceived in the same way as a real event) then our abilities to discriminate the real event from an artificial sensory stimulus would be diminished to the point at which presence could be achieved through the mere perceptual realism of the stimulus. In the next paragraphs we describe two examples of how presence research may bridge the gap between fundamental insights in the mechanisms of perception and the way media represent reality in order to overcome inadequacies and distortions intrinsically present in current media technologies.

**Why we need presence research: The case of colour perception and reproduction**

Colour discrimination is an important part of human existence. Whether it is choosing between unripe or ready fruits, deadly or tasty mushrooms, live or neutral wires, stop or go lights, toxic or safe substances, colour perception is a fundamental aspect of most peoples everyday lives. Yet the RGB colour space that we currently use to artificially represent natural colours can only show approximately ten percent of the over one million brightness and colour combinations that most people can discriminate (Hurlbert, 2001). A simple but surprising example of this inadequacy can be found on almost any packet of flower seeds. People often buy seeds using the colour photograph on the packaging as a guide. They expect and even rely on the fact that the photograph will show an accurate colour reproduction of the adult plant. For obvious commercial reasons, it is vital that this image is a good match. Colour photography however is very poor at providing adequate representation of subtle colourations and natural hues. Photographers call this phenomenon "subject colour failure" and accept it as a fundamental limitation of their medium. Seed growers overcome the problem by using hand colouring over a monochrome rendition of the adult plant. They compare each colour on the adult plant against a physical colour chart, often from the Pantone printing ink colour space. These colour values are conveyed to the printers who then ensure that the correct dyes are printed over the monochrome photograph to ensure a good match. Commercial colour printers rarely use the RGB colour space, as they have understood its limitations for decades. Most photomechanical reproduction systems are based on the CMYK colour space, yet can have up to sixteen colour layers in the final print. Even then, it is known that there are some combinations or types of colours that the best systems still cannot reproduce.

Many people have assumed that the success of home shopping using colour catalogues would undermine the previous statement. However, catalogue shopping has often been expensive because of its returns policy. If a customer rejects a product and returns it because the colour of the garment or product they bought does not match the colour they thought they were buying, *then the cost of that failed transaction is carried by the company ( and therefore the clients of the company)*. It is a core inefficiency that undermines the profitability of home shopping. The Internet uses the restricted sRGB colour gamut to convey its photographic images. So it is hardly surprising that clothing does not sell well over the Net. If computer based digital imaging used a colour space that was derived from a fundamental understanding of the human visual system, then it would be possible for people to remotely sense stimuli that previously could only be experienced by direct contact. Doctors could diagnose many illnesses without physically being there, if imaging systems could convey the information they require. They use vision and haptic sensation in conjunction with question and answer interaction for many of their diagnoses. If they could do this remotely, then doctors could be telepresent almost anywhere and the time saved between accidents and treatment could save many lives. Military or police observers could use their highly trained visual acuities without exposing themselves to the danger of attack. Working on or above a location (in an aircraft) exposes many observers to danger as a normal part of their jobs. Bomb disposal experts, fighter pilots, submarine pilots, geographers, geologists, biologists, power line workers, safety engineers and countless other professions have to take risks to use their senses when it is now possible for them to do their work from a safe location. Presence research is currently the only route by which all of our sensory thresholds can be identified and realistically stimulated.

**Why we need presence research: The case of body image distortion through photography**

Poor colour discrimination is not the only problem caused by inadequacies in imaging technology. Recent research has indicated that conventional single lens photography is surprisingly poor at conveying images of people and objects with their correct size and shape information. Photographers are occasionally aware that scenes they see with their normal direct vision will change significantly in their 2D representations when they are seen on the page or screen. One of the best known, and most disconcerting changes, is the fattening effect of photography (Harper & Latto 2001). Models and actors are often chosen for slimness in the attempt to combat the fattening effects of photography. While looking “normal” in the photographic image, they often look frighteningly slim in real life. The reason for this is that cameras and people see the world from a single-point perspective. By the process of cyclopean vision (Julesz, 1971) we perceive through a “virtual eye” that generates an artificial viewpoint from a location mid-way between each real eye. However, recent experiments suggest that a single-lens camera cannot convey the way in which we can focus/fuse on an object with two eyes and simultaneously see both diverging and converging optical paths. Close-up objects viewed stereoscopically occlude less of the background than their monocular equivalents and therefor are seen as significantly slimmer. Failure to reproduce this perceptual geometry in a display medium is a major cause of the fattening effects and misperception in conventional photographic, film and television images.

The public health issues relating to these distortions were discussed in a major review document “Eating Disorders, Body Image and the Media” by the British Medical Association. One of the many findings of this report was that actresses on average are 20% percent slimmer than non-actresses of the same age. It is now generally accepted that that there has been an obsession with slimness in photography, film, television and print media. Yet recent work by Professor Rose Frisch of Harvard University has confirmed that body fat and fertility is intimately linked in pre-menopausal women. She says[1] “We can think of body fat as *Sex Fat*, because it provides the

[1] From the Observer newspaper interview, 26-0502 p.7. on the publication “Female Fertility and the Body Fat Connection.”

energy required for reproduction. What I found as most astonishing is that there is a razor-thin borderline where losing just 3lb can tip a normal sized woman into infertility. And many women who maintain the body shape made popular on the catwalks are completely infertile. These women can, however, continue having periods and not guess that anything is wrong until they try, and fail, to get pregnant. If low body fat switches off a woman's reproductive system for too long, it may be too late to regain fertility simply by putting on weight. Women need to gain weigh at a younger age. A 5ft 3" girl must weigh at least 90lb if she is to become fertile, while a pre-menopausal woman of the same height needs to over 101lb if she is to continue to ovulate."

There is now widespread public health concerns regarding healthy people slimming to the point at which their health is seriously impaired. Early mortality is an inevitable outcome for large proportion of people who restrict their food intake in this way. It is probable that the fattening effect of photography has not previously been linked to anorexia nervosa and related illnesses because until the Harper and Latto studies, no one had proposed a mechanism by which it might occur. Their results suggest that anyone who cannot form a cyclopean view in the brain is likely to perceive images of people (with their direct vision) as fatter than observers with normal stereo vision. And surprisingly, many people are completely unaware that they do not have normal stereovision. Paradoxically, the photographic images that they see in magazines and television would appear to have a similar type of body image information as their own direct vision would give them. This could lead them to trust images that are already distorted because they have not been made orthostereoscopically. Orthoptists estimations vary on the amount of stereo acuity failure that exists within the general population, but figures of between one in six to one in nine are mentioned anecdotally. These people may have two eyes, each with excellent vision. But the mechanism for allowing them to fuse a stereo image in the brain is encumbered in various ways that effectively gives them zero measured stereo acuity. This leads to the important possibility, as yet not investigated empirically, that people with stereo acuity failure are more likely to develop a slimming disease. Presence in the media would allow remote viewers to perceive unhealthy levels of slimming as clearly as if they were in the presence of the real person, rather than using the "flattened and fattened" 2D image they see today.

## 5. A Vision of the Future

It is inevitable that new imaging and media technologies will be developed that will be more realistic than currently available systems. But effort and resources will be wasted on these developments until there is a fundamental understanding of how we perceive the world around us. Once we understand the thresholds, mechanism and interactions of all of the normal human sensory modalities, we can artificially deliver compelling and realistic experiences through the development of new media technologies. Current efforts to produce compression algorithms have failed to find solutions that do not throw away data vital to realism. This is almost certainly due to the fact that the psychophysical research on which their designs were based is outdated and inadequate. MPEG compression technology is now known to discard motion and colour data that it had been wrongly thought we could not see. Audio compression too has been hindered by the same theoretical failures, leading to the observation by audiologists that every compression system so far developed introduces artefacts that can clearly be perceived when compared to the original source. When the promised "unlimited bandwidth" data transmission systems arrive, heavy compression will be unnecessary and the prospects for presence will be significantly enhanced. Understanding how the brain produces such rich percepts through the seemingly impoverished sensory abilities given to us by nature will also allow for reductions in the amount of data required to reproduce reality.

The forces of motion itself can also now be effectively reproduced in a simplified immersive environment. Accelerations and forces can easily be recorded. And now they can be replayed without physically moving the subject. Galvanic Vestibular Stimulation (Scinicariello, Eaton, Inglis, & Collins, 2001) works by using tiny currents behind each ear to selectively stimulate the vestibular system. A prototype device can now "fine tune" sensation to make balancing easier, or "move" a subject so that they feel as if they are on a motion platform. Expensive motion platform style simulations could soon become commonplace once the need to "move the room" has been replaced by a tiny headset. Olfaction too is now on the list of things we can realistically simulate. DigiScents Technologies (now bankrupt) claimed to be able to record smells and replay them from a small palette of source scents that could be mixed to reproduce the original stimuli. The most pertinent challenges in this area are the reliable capture of smell (i.e. an artificial "nose") and the swift removal of the smells once they have been delivered.

Much of the travelling that people commit themselves to in a lifetime is out of duty rather than choice. Presence research shows that we do not need to travel to experience social presence or to use our senses to experience a remote stimulus directly. In the future, humans will physically travel to a destination because they want to, rather than need to do it. For instance, we would be able to attend conferences or interact physically and socially at a remote location without leaving our homes or workplace. It would be almost like the teleport technology so beloved by popular science fiction writers. However, it is likely that the software projection holograms found in science fiction films will not be easy to reproduce in the near future. In the film "Star Wars," a low-resolution holographic image of the Princess is projected in 3D space as a vital part of the plot. This simple effect is as science fiction now as it was 25 years ago when the scene was filmed. While some groups have reported success in capturing still holographic images without the need to use laser illumination (Marks, Stack, Brady, Munson, & Brady, 1999), no one has yet speculated on what type of display could reproduce the data or make it move realistically. The data rates required to do this are an order of magnitude above what is possible using envisioned developments of computer technology. Fortunately for presence research, many of the characteristics of software holograms have already been approximated using existing lens based imaging systems.

Realistic remote sensing and interaction is not the only possibility offered by presence research. Extended "hypersensory" sensation would also be possible too. Most lens-based imaging can easily have its sensitivities increased to include ultraviolet, infrared, and polarised light. It is feasible that these emissions could be presented within an immersive multisensory display so that engineers could sense stress patterns, hot spots and microscopic details that they could not see with direct sensory stimulation. Perhaps sensitivity could be extended to x-rays, gamma rays and other useful electromagnetic spectra. These could be added to expert computer systems that could be of great assistance to many professional disciplines.

In practice however, the most widespread use of presence technologies is likely to be in museums, encyclopaedias, entertainment, and the mass media industries. The enhancements that this technology could bring to people's lives are immense. It could become so pervasive that it might one day become difficult to imagine life without presence as part of our day to day reality. We might have life-long and intimate relationships with people that we have never physically met. The possibility to "travel" to places that are impossible for us to physically visit, or might not even exist any more is also enticing. The film industry is currently investing in the development of "synthespians." These are virtual actors who look, sound and appear to interact as if they were real. For studios that own the likenesses of long-dead Hollywood stars, the prospect of a "new" Marilyn Monroe, Humphrey Bogart or Bruce Lee film is too tempting to ignore. If this technology was coupled to a high-speed rendering engine and an interactive artificial intelligence program, then we could "meet" people in virtual environments who do not actually exist. The possibilities of this type of development are almost endless.

It could be argued that human perception is so sensitive that technology could never fool it into completely believing that even the finest artificial stimulus is in fact real. This argument has now been challenged by recent research using Transcranial Magnetic Stimulation (Cook & Persinger, 1997). By using tiny rotating magnetic fields (1 microtesla) a sense of human (and sometimes even "god-like") presence is induced in people who are in a sensory deprivation experiment. It works in over 80% of subjects at signal levels barely above the background noise in the brain. Subjects report that this sensed "physical presence" is vivid, exciting and completely believable. It really does not seem like science fiction to imagine a small TMS headset that one would wear while watching a next generation multi-sensory display. It could help us to suspend our disbelief that what we are experiencing is not real. It could also allow lifelike perception to become a reality, even in subjects who might be thought to be resistant or sensitive to artificial sensory stimulation. It might also allow less than lifelike resolutions, colour reproduction and sensory interactions to be perceived as completely "real." The combination of precisely gathered environmental information (that is conveyed in an uncorrupted way to the user) with this lowered reality threshold could make presence become a reality that would enhance the quality of life for millions of people. We simply need to do the basic research first and then creatively apply its findings.

## References

Belton, J. (1992). *Widescreen Cinema*. Cambridge, MA: Harvard University Press.

Biocca, F. (2001). Presence of M.I.N.D. Invited talk at the European Presence Research Conference, Eindhoven, 9-10 October 2001.

Biocca, F. & Delaney, B. (1995). Immersive virtual reality technology. In: Biocca, F. & Levy, M. (eds.), *Communication in the Age of Virtual Reality*. Hillsdale, NJ: Lawrence Erlbaum, pp. 57-124.

Campanella, T.J. (2000). Eden by wire: Webcameras and the telepresent landscape. In: Goldberg, K. (ed.), *The Robot in the Garden*. Cambridge, MA: MIT Press, pp. 22-46.

Cook, C.M. & Persinger, M. (1997). Experimental induction of the “sensed presence” in normal subjects and an exceptional subject. *Perceptual Motor Skills 85*, 683-693.

Dennett, D.C. (1978). *Brainstorms. Philosophical Essays on Mind and Psychology*. Brighton, UK: Harvester Press.

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

Ellis, S.R., Dorighi, N.S., Menges, B.M., Adelstein, B.D., & Jacoby, R.H. (1997). In search of equivalence classes in subjective scales of reality. In: Smith, M.J., Salvendy, G., & Koubek, R.J. (eds.), *Design of Computing Systems: Social and Ergonomic Considerations*. Amsterdam: Elsevier, pp. 873-876.

Freeman, J., Avons, S.E., Pearson, D.E., & IJsselsteijn, W.A. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments 8*, 1 - 13.

Freeman, J., Avons, S.E., Meddis, R., Pearson, D.E. & IJsselsteijn, W.A. (2000). Using behavioural realism to estimate presence: A study of the utility of postural responses to motion-stimuli. *Presence: Teleoperators and Virtual Environments 9*, 149-164.

Harper, B., & Latto, R. (2001). Cyclopean Vision, Size Estimations and Presence in Orthostereoscopic Images. *Presence: Teleoperators and Virtual Environments 10*, 312-330.

Hatada, T., Sakata, H. & Kusaka, H. (1980). Psychophysical analysis of the "sensation of reality" induced by a visual wide-field display. *SMPTE Journal 89*, 560-569.

Hendrix, C., & Barfield, W. (1996a). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments 5*, 274 - 289.

Hendrix, C., & Barfield, W. (1996b). The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments 5*, 290 – 301.

Hoffman, H.G., Prothero, J., Wells, M.J., & Groen, J. (1998). Virtual chess: Meaning enhances users' sense of presence in virtual environments. *International Journal of Human-Computer Interaction 10*, 251-263.

Hurlbert, A. (2001). Colour perception in the natural world: From quanta to qualia. Invited talk at the 'Reality and our Sense of Presence" conference, IST 2001 Event, Düsseldorf, 4 December 2001.

IJsselsteijn, W.A., de Ridder, H., Freeman, J., & Avons, S.E. (2000). Presence: Concept, determinants and measurement. *Proceedings of the SPIE 3959*, 520-529.

IJsselsteijn, W.A., de Ridder, H., Freeman, J., Avons, S., & Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion and screen size on subjective and objective corroborative measures of presence. *Presence: Teleoperators and Virtual Environments 10*, 298-311.

IJsselsteijn, W.A., de Ridder, H., Hamberg, R., Bouwhuis, D., & Freeman, J. (1998). Perceived depth and the feeling of presence in 3DTV. *Displays 18*, 207-214.

Johnson, E.G. & Corliss, W.R. (1971). *Human Factors Applications in Teleoperator Design and Operation*. New York: Wiley-Interscience.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2000). Development of a new cross-media presence questionnaire: The ITC-sense of presence inventory. In: W.A. IJsselsteijn, J. Freeman, & H. de Ridder (eds.)**,** *Proceedings of PRESENCE 2000 - 3rd International Workshop on Presence* (CDROM), Delft, 27-28 March 2000.

Lodge, N. (2000). What can we learn from the cinema of the 1950's? Invited talk at PRESENCE 2000 - 3rd International Workshop on Presence, Delft, 27-28 March 2000.

Lombard, M. (1995). Direct responses to people on the screen: Television and personal space. *Communication Research*, 22: 288-324.

Lombard, M., & Ditton, T.B. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2). [online]

Marks, D.L., Stack, R.A., Brady, D.J., Munson Jr., D.C., & Brady, R.B. (1999). Visible Cone-Beam Tomography with a Lensless Interferometric Camera. *Science 284*, 2164-2166.

Minsky, M. (1980, June). Telepresence. *Omni*, pp. 45-51.

Prothero, J. & Hoffman, H. (1995). Widening the field of view increases the sense of presence. HITLab Technical Report R-95-5. [online]

Schubert, T., Friedman, F., & Regenbrecht, H. (1999). Decomposing the sense of presence: Factor analytic insights. Paper presented at the 2nd International Workshop on Presence, University of Essex, 6-7 April 1999.

Scinicariello, A.P., Eaton, K., Inglis, J.T., & Collins, J.J. (2001). Enhancing human balance control with galvanic vestibular stimulation. *Biological Cybernetics 84*, 475-480.

Sheridan, T. (1992). *Telerobotics, Automation and Human Supervisory Control*. Cambridge, MA: MIT Press.

Slater, M. & Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments 2,* 221-233.

Slater, M., Usoh, M. & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environment 3,* 130-144.

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.

Todd, N.P.M. (in press). Evidence for a behavioural significance of saccular acoustic sensitivity in humans. *Journal of the Acoustical Society of America.*

Todd, N.P.M. & Cody, F. (2000). Vestibular responses to loud dance music: A physiological basis for the "rock and roll threshold"? *Journal of the Acoustical Society of America 107*, 496-500.

Welch, R. B., Blackmon, T.T ., Liu, A., Mellers, B. A., Stark, L. W. (1996). The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence: Teleoperators and Virtual Environments*, 5: 263-273.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments,* 7: 225-240.