International Society for Presence Research

24-26 August

The 9th International Workshop on Presence

Proceedings

Edited by
Cheryl Campanella Bracken, Cleveland State University, USA
and
Matthew Lombard, Temple University, USA

International Program Committee

Mariano Alcaniz (Universidad Politécnica de Valencia, Spain)
Carlo Alberto Avizzano (Scuola Superiore S. Anna, Italy)
Ann-Sofie Axelsson (Chalmers University of Technology, Sweden)
Jeremy N. Bailenson (Stanford University, USA)
Rosa Baños (University of Valencia, Spain)
Woody Barfield (University of North Carolina at Chapel Hill, USA)
Edwin Blake (University of Cape Town, South Africa)
Cristina Botella Arbona (Universitat Jaume I, Spain)
Doug Bowman (Virginia Tech, USA)
Cheryl Campanella Bracken (Cleveland State University, USA)
Prabu David (Ohio State University, USA)
Theresa Ditton (Strategic Business Research, Inc., USA)
Jonathan Freeman (Goldsmiths College, University of London, UK)
Doron Friedman (University College London, UK)
Luciano Gamberini (University of Padua, Italy)
Marco Gillies (University College London, UK)
Feliz Ribeiro Gouveia (Universidade Fernando Pessoa, Portugal)
Ilona Heldal (Chalmers University of Technology, Sweden)
Wijnand Ijsselsteijn (Technische Universiteit Eindhoven, Netherlands)
Matthew T. Jones (Temple University, USA)
Nicole Kraemer (University of Cologne, Germany)
Rita Lauria (North Carolina Agricultural and Technical State University, USA)
Kwan Min Lee (University of Southern California, USA)
Jane Lessiter (Goldsmiths College, University of London, UK)
Matthew Lombard (Temple University, USA)
Celine Loscos (University College London, UK)
Claus Madsen (Aalborg University, Denmark)
Katerina Mania (University of Sussex, UK)
Fabrizia Mantovani (Istituto Auxologico Italiano, Italy)
Miriam Reiner (Technion: Israel Institute of Technology, Israel)
Albert Skip Rizzo (University of Southern California, USA)
Daniela Romano (University of Sheffield, UK)
Roy Ruddle (University of Leeds, UK)
Maria Victoria Sanchez-Vives (Instituto de Neurociencias de Alicante, Spain)
Ralph Schroeder (Oxford Internet Institute, UK)
Thomas Schubert (University of Jena, Germany)
Melissa E. Selverian (Temple University, USA)
Mel Slater (University College London, UK)
Michael Smyth (Napier University, Scotland
Anna Spagnolli (University of Padua, Italy)
Susan Turner (Napier University, Scotland
Aleksander Väljamäe (Chalmers University of Technology, Sweden)
Paul F.M.J. Verschure (Institute of Neuroinformatics, Switzerland)
Vinoba Vinayagamoorthy (University College London, UK)
John Waterworth (Umeå University, Sweden)

Welcome from Conference Co-Chairs

We are very pleased to welcome you to Cleveland and this 9th Annual Workshop on Presence. Thanks to the scholars and practitioners who submitted their work for blind peer review, and the dedicated members of the international Program Committee who did the reviewing, we expect another productive set of presentations and discussions this year.

We've made an effort to create a more diverse and discussion based program this year to take advantage of our collective physical presence at Cleveland State University. While most panels feature standard presentations (which will include time for audience questions and responses), this year two new High Density sessions will feature brief introductions to a larger number of projects followed by separate informal discussions and demonstrations. There are also two sessions ("Session 6: Pushing Presence: Have We Gone Far Enough?" and "Session 8: Presence Controversies") with formats designed to allow us to consider important issues in our field through (hopefully) productive discussion and debate. And we have an outstanding keynote speaker, Janet Murray, Georgia Institute of Technology professor, author of *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*, and an internationally recognized interactive designer.

We want to thank everyone who has assisted in the planning of the conference, including the Program Committee and CSU graduate students who have volunteered their time. Special thanks to student Idalita Raso who conducted a wonderful public relations campaign here in Cleveland. We also want to acknowledge the financial support of both the Cleveland State University College of Liberal Arts and Social Science and the School of Communication.

We've both found previous PRESENCE conferences particularly enjoyable and intellectually stimulating and we hope this event provides that experience for you!

Matthew Lombard, Ph.D.
Temple University
Philadelphia, Pennsylvania
President, International Society for Presence Research

Cheryl Campanella Bracken, Ph.D.
Cleveland State University
Cleveland, Ohio
Conference, Co-Chair

About ISPR and the PRESENCE Workshops

The PRESENCE Workshops began with a small, informal gathering of scholars and professionals interested in presence theory and research in June 1998 in Suffolk England. Thanks to the hard work of conference organizers and the support of the growing presence community, the Workshops have been held each year since then and have retained their single track format and productive but informal and very pleasant character. Next year will mark the conference series' first decade:

PRESENCE 1998 - BT Labs, Suffolk, England

PRESENCE 1999 - University of Essex, Colchester, England

PRESENCE 2000 - Eindhoven University of Technology, Delft, The Netherlands

PRESENCE 2001 - Temple University, Philadelphia, Pennsylvania, USA

PRESENCE 2002 - Universidade Fernando Pessoa, Porto, Portugal

PRESENCE 2003 - Aalborg University, Aalborg, Denmark

PRESENCE 2004 - Polytechnic University of Valencia, Spain

PRESENCE 2005 - University College London, UK

PRESENCE 2006 – Cleveland State University, Cleveland, USA

Following the 2001 conference, the International Society for Presence Research (ISPR) was founded to coordinate the annual conferences, sponsor other conference panels and events, and provide a variety of resources to those who conduct research, develop theory, write about, or simply are interested in, the concept of presence. The ISPR web site, at http://ispr.info, provides many of these resources (as well as a list of the organization's current Board of Directors). ISPR also sponsors the presence-l listserv (see ispr.info for details).

The ISPR Board is planning future international presence conferences, enhancing the resources it provides online, developing procedures to offer organizational memberships, and considering the development of a journal; look for announcements about all of these at ispr.info and on the presence-l listserv. ISPR exists to serve the presence community and we always welcome questions, comments and suggestions at help@ispr.info (or directly to me at lombard@temple.edu).

Matthew Lombard, Ph.D.
Temple University
Philadelphia, Pennsylvania
President, International Society for Presence Research
http://matthewlombard.com
http://ispr.info

Content

Presence: A Network of Reciprocal Relations

Pericle Salvini
IMT - Institutions Markets Technologies, Lucca, Italy
{pericle.salvini@imtlucca.it}

Abstract

In this paper, I propose an alternative explication of presence based on the theoretical framework provided by theatrical presence. In my opinion, a unified and consistent discourse about presence has emerged since the earliest theorizations and explications of telepresence in the '80s. Scholars, scientists, engineers, psychologists, philosophers and virtual reality experts have defined remote and virtual presence according to a common denominator: the assumption that presence is the result of subjective sensory rich experience mainly given by immersion and unilateral activity (vision and action). On the contrary, according to the explication proposed in this paper, presence occurs when natural and/or technological conditions allow for reciprocal relationships. Finally, political and social issues related to presence and remote presence are taken into account.

Keywords--- **Theatrical presence, presence, remote presence, ethics.**

1. Introduction

More than ten years ago, Thomas B. Sheridan, one of the firsts scholars to investigate the concept of presence, maintained: '[a]t present, we have no theory of presence, let alone a theory of telepresence or virtual presence. This is in spite of the fact that students of literature, the graphic arts, the theatre arts, film, and TV have long been concerned with the observer's sense of presence' [1].

In this paper, therefore, I propose to pick up Sheridan's suggestion and look at one specific art form in particular – the theatre – in order to see how it can contribute to the explication of presence.

Since the earliest theorizations and explications of telepresence and teleoperation, I argue, the underlying notion of presence has been understood and described mainly as a subjective or ego-centric sensory experience, provided by unilateral interaction (action and vision) and immersion. According to Marvin Minsky – one of the pioneers of artificial intelligence – telepresence is the name given to robotic devices, such as mechanical hands or movable platforms, which can be remotely controlled to accomplish various kind of tasks: '[t]o convey the idea of these remote-control tools, scientists often use the words teleoperators or telefactors. I prefer to call them telepresences' [2]. Ten years later, in his musings on telepresence, Sheridan defined telepresence as an experience deriving from teleoperation, namely the 'human control of vehicles, manipulators and other systems using video, audio, kinaesthetic and tactile feedback from the remote site' [1].

For Paul Virilio, telepresence, telerobotics and teleoperation are all interconnected; in fact, he defines telepresence as 'an advanced form of teleoperation in which the robot operator gets a sense of being "on location", even if the robot and the operator are miles apart; control and feedback are done via telemetry sent over wires, optical fibers or radio' [3]. Lev Manovich's definition is exemplary in confusing both teleoperation with telepresence. He defines telepresence as 'the ability to remotely manipulate physical reality in real time through its image. […]. *A better term would be teleaction. Acting over distance. In real time*' [4, original emphasis]. More recently, and more specifically, the notion of presence has been described as the 'perceptual illusion of non-mediation' [5], as 'the observers subjective sensation of 'being there' in a remote environment' [6], 'the experience of being there in a mediated environment' [7]. Furthermore, presence has been explicated as 'tied to one's successfully supported action in the environment' [8], and as 'a series of moments when cognitive and perceptual reactions are closely tied to current sensory impingements' [9].

Conversely, theatrical presence, I argue, offers an alternative model for explicating presence, which I believe could be usefully implemented also in the context of remote and virtual environments. Basically, the explication of presence proposed here is not based on the dichotomous relation between an active subject and a passive object. According to the model proposed here, presence occurs when natural and/or technological conditions allow for reciprocal relationships. Furthermore, and most importantly, the approach to presence based on reciprocity brings to the fore an aspect which is still missing in current research: namely, the ethical and political implications of presence.

In section one, I will introduce the reader to current varieties of theatrical presence. In section two, I will propose a general explication of presence. In section three, drawing on telepresence art, I will provide the reader with two examples of remote presence based on reciprocity. Finally, in the last section, I will take into account the social and political issues concerning presence.

2. Theatrical presence

Theatre has always been considered as the mirror of the world. Calderon de la Barca's 'el gran teatro del mundo' and Shakespeare's 'all the words is a stage' are just two of the most popular statements confirming the relationship between life and theatre. No surprise, therefore, if in this paper I will se theatre to cast some lights on the notion of presence.

Indeed, the concept of presence is one of the tenets of theatre art, and with the passage from "mechanical reproduction" [10] to "technological presentation" [11] its relevance has become even more crucial. The literature about theatrical presence is rich and vast. However, in what follows, I will sum up the main varieties of theatrical presence by drawing on the analyses and insights of Roger Copeland [12].

Copeland explains that 'the word "presence" means different things to different people – and […] some of these meanings are mutually exclusive' [12]. Among the many conceptualizations of theatrical presence, Copeland lists "stage presence", also known as the performer's *charisma*, which he further subdivides into two other kinds of presence: "possession", namely the capacity to project a fictional character by the actor; and "authenticity", that is, 'the performer's ability to […] reveal her "authentic" self' [12]. According to Copeland there exists another conception of theatrical presence, which has nothing to do with either "charisma" or "authenticity", and consists of "being in the presence of". In other words, this kind of theatrical presence takes place when performers and spectators simply 'share a certain amount of time together in the same space' [12]. As Copeland remarks, in this case, the only condition for presence to take place is represented by the *possibility* of touching the performers: '[p]erhaps being "in the presence of" a performer means that we could, if we so desired, reach out and touch (that) someone' [12]. Finally, Copeland proposes a last conceptualisation of theatrical presence, according to him, the most appropriate definition of theatrical presence. He maintains that presence in the theatre has to do with the sense of *reciprocity* taking place between actors and spectators, namely 'a sense that what transpires onstage – in contrast to the movies – is affected almost as much by what happens in the audience as the other way around' [12]. According to Copeland, therefore, 'presence in the theatre has […] to do with […] the way in which the architectural and technological components of the performance space either promote or inhibit a sense of "reciprocity" between actors and spectators' [12].

I would like to conclude this section with some remarks concerning the varieties of theatrical presence that I have just discussed. There is one fundamental assumption underlying all kinds of theatrical presence: actors and spectators are physically sharing a space. However, in the variety of stage presence, sharing a space does not seem to be determinant in eliciting presence. As a matter of fact, presence, in both "authenticity" and "possession" varieties, is produced by (and is the prerogative of) the performers only. On the one hand, presence depends on the performer's ability to remove all forms of theatrical mediation and reveal her nude, true self. On the other, presence results from the performer's ability to create an illusion, namely, that the characters is there, on stage. Therefore, in both varieties of stage presence, "authenticity" and "possession", presence is unilaterally determined by the actors, whereas the spectators are not involved in the production of presence.

On the contrary, in the other two varieties, i.e. "to be in the presence of" and "presence as reciprocity", sharing a space acquires more relevance. In the former case, sharing a space is the only determinant of presence, and to stand up and touch the performers is the proof-test. However, even though sharing a space is a fundamental condition for presence to occur, nonetheless, presence is still understood as a unilateral event, this time the prerogative of spectators only (i.e., they are in the presence of). Conversely, in the latter case, "presence as reciprocity", presence is given by the reciprocal relations taking place between actors and spectators. Hence, presence does not characterize just one party, but both; neither is it a subjective experience determined by the abilities of a person – whether authenticity or possession – nor the result of an experiential fact (I touch it therefore I am present). Presence has become a bilateral event, a condition given by external, objective, circumstances, that is, by the architectural and technological components of the theatrical space. Compared to all other varieties of presence, this last definition is more holistic and transitive.

3. Towards a general theory of presence: Presence as a network of reciprocal relations

Drawing on Copleand's last definition, I propose the following explication of presence: *the way in which natural and/or artificial (i.e. technological) factors/conditions either promote or inhibit a sense of reciprocity between two or more people or between an environment and a person.* I agree with Giuseppe Mantovani and Giuseppe Riva's assumption that 'the meaning of presence is closely linked to the concept we have of reality, i.e., to the ontology that we more or less explicitly adopt. Different ontological stances support different criteria for presence, telepresence, and virtual presence' [13]. According to my ontological stance presence is a natural condition pertaining to all living organisms inhabiting an environment. We all live in a condition of presence even though we are not aware of it. In my opinion, being-in-the-world means being caught up in this network of reciprocal relations. Of course, there might be natural or technological factors preventing us from being present or reducing our condition of presence, as in case of physically impaired people or when technology is purposefully used to inhibit the reciprocal relations that characterize our relations with the environment and the other human beings, as for instance when handling hazardous material by using tele-operation systems or sending an email instead of having a face-to-face meeting.

However, technology can also be used in order to extend and enhance the condition of presence, as for instance in all those cases in which it provides us the means to cross spatial boundaries, allowing us to share a space with somebody, as for instance when making a telephone call.

When talking about presence in remote or virtual environments, scholars have pointed out the necessity of designing and developing more transparent media, disappearing interfaces, tactile sensory feed-back devices and so on. Although I consider these very important research areas, I believe that there can not be presence without reciprocal relations. As a matter of fact, impressive technologies such as those used in warfare, commonly described as telepresence technologies, though providing a sense (or illusion) of being there, they do not allow reciprocal relations and this elicit abstraction and moral disengagement. [14].

According to my understanding, remote presence can be achieved only when technologies keeps intact the presence condition: i.e. the network of reciprocities. Presence always implies a double flow and the awareness of being in a reciprocal relationship with some-*body* or some-*thing*, to be trapped in a *becoming* which affects the other and myself at the same time. Therefore, to be present is to be subjected to the contingency and randomness of a situation. As soon as I shield myself and escape or avoid one or all these conditions, I am no more present, or, if we want to quantify presence, I am less present.

3.1. Maurice Merleau-Ponty's Phenomenology of Perception

The phenomenological philosophy of Maurice Merleau-Ponty, in my opinion, points out the indissoluble relation that connects presence and reciprocity. Merleau-Ponty considers reciprocity a fundamental aspect of existence. This is especially true, if we consider his notions of the *intertwining* or the *chiasm*. Merleau-Ponty explains: '[m]y body as a visible thing is contained within the full spectacle. But my seeing body subtends this visible body, and all the visible with it. There is reciprocal insertion and intertwining of one in the other' [15]. And this is particularly true as regards vision and touch: '[t]here is vision, touch, when a certain visible, a certain tangible, turns back upon the whole of the visible, the whole of the tangible, of which it is a part' [15].

Presence, therefore, seems characterized by a sort of Narcissism, the one Merleau-Ponty speaks of about vision, which can be understood as reciprocal reflection:

'[T]here is a fundamental narcissism of all vision. And thus, for the same reason, the vision he exercises, he also undergoes from the things, my activity is equally passivity – which is the second and more profound sense of the narcissism: not to see in the outside, as the other sees it, the contour of a body one inhabits, but especially to be seen by the outside, to exist within it, to emigrate into it, to be seduced, captivated, alienated by the phantom, so that the seer and the visible reciprocate one another and we no longer know which sees and which is seen' [15]. According to Patrick Burke, for Merleau-Ponty visibility, 'as well as all the other modes of present-ability, is thus characterized by a reversibility of the seer and the seen, of the touching and the touched, and is not a function of either term but of both, insofar as they are originally unified in the flesh' [16].

On the whole, the philosophy of Merleau-Ponty, I argue, describes being in the world as relational, more precisely as the intertwining of the subject with other subjects or between the subject and his/her environment. On the contrary, as seen earlier, telepresence in its ordinary use tends to foster an understanding of presence as based on unilateral activity and subjective feeling.

What is missing, therefore, in the discourse about presence and telepresence brought about by teleoperation, telerobotics and some virtual reality applications, is a fundamental aspect of existence, which can be explained by drawing on another notion introduced by Merleau-Ponty: the 'flesh'. Merleau-Ponty maintains that the flesh 'is not matter, is not mind, is not substance' [15]. According to him, the flesh is an 'element', 'in the sense of a general thing, midway between the spatio-temporal individual and the idea. [...] The flesh is in this sense an "element" of Being' [15]. The relevance of flesh for my argument is implicit in this passage: 'For if there is flesh, that is, if the hidden face of the cube radiates forth somewhere as well as does the face I have under my eyes, and coexists with it, and if I who see the cube also belong to the visible, I am visible from elsewhere, and if I and the cube are together caught up in one same "element" (should we say the seer, or the visible?), this cohesion, this visibility by principle, prevails over every momentary discordance' [15].

Therefore, the importance of flesh is that within it the one who sees is also the one who is seen and the one who touches is also the one who is touched. The network of reciprocal relations is intact.

Let's stop here and see how a definition of presence based on reciprocity can be put into practice in remote presence applications.

4. Telepresence Art

4.1. Telephonic Arm Wrestling

This artwork was presented by Canadian artists Norman White and Doug Back in 1986 and took place in two locations simultaneously: at the Canadian Cultural Centre in Paris and at the Artculture Resource Centre in Toronto. As its title tells, the piece consists in an arm wrestling. What distinguishes it from an ordinary arm wrestling is that the competition takes place between two wrestlers located miles and miles away from each other. The piece attempted to reproduce the intensely physical and bodily experience which characterised the arm wrestling, by producing an equivalent physical experience not based on the here and now but on distance. How is that possible that two geographically-distant people can engage in an arm wrestling? Each competitors seated in front of the same identical technological apparatus, consisting of two kinematically equivalent masters, made of boxes, cables and an aluminium stick similar to that of a joystick. The materials used were very simple and low cost: steel, Plexiglas, motors, and custom electronics. The two platforms were connected via telephone line. Using motorized force-transmitting systems interconnected by a telephone data link, the force signal inputted by the hand of the participant in location A was transformed by a modem into an output signal which was sent to another modem in location B which ri-converted it in a force signal on the joystick, and vice-versa. The joystick can be considered as the prosthetic interface or the vicarious arm used by the remote competitors.

Although Edward Shanken explains that '[i]t was impossible for the competitors to really have much of a fight' [17] because of the time-delay in the telephone link, White maintains that the lever was remarkably sensitive: 'You could almost feel the pulse of the other person [...] it was uncannily human-like – the sensation of sinews and muscle – not at all like feeling a machine' [18].

Shanken reports that the idea to build an arm-wrestling device came up to Back and White while talking in a bar about the arms race, a topic very popular in the '80s. Back said: 'Wouldn't it be great if [the arms race] could be resolved by arm wrestling?' [18]. Behind Back's ironic proposal, I believe possible to see the origin a different understanding of technology. Instead of seeing technology as a means to compensate for the loss of physical capabilities or for augmenting existing ones, the idea behind Back's words is that technology can be used to help mankind to settle controversies by fostering antagonistic relationships not based on hierarchical disposition. Technology, therefore, appears as a tool for negotiations and dialogue. The attitude implicit in Back's statement is one of reciprocity and responsibility. The notion of presence deriving from the *Telephonic Arm Wrestling* is based on interactivity. However, interactivity here is not meant as a *unilateral transmission* of actions. In other words, interactivity consists of a *bi-directional exchange* of actions. The robotic platforms do not only transmit action, but they are implemented so as to receive it as well. Presence and telepresence therefore, are the result of a communica(c)tion, namely, actions taking place in a dematerialised and ri-materialised plane of dialogic communication.

4.2. Telematic Vision

Telematic Vision by British artist Paul Sermon took place for the first time at the ZKM Museum in Karlsruhe, Germany, within the context of the ZKM Multimediale 3 exhibition, in 1993. *Telematic Vision* can be described as an installation taking place simultaneously in two identical sets located far away from each other. Each set is furnished with a large blue sofa placed in front of a TV screen which occupies the middle of a room. Inside the room, there are also a blue carpet and two other screens placed at both sides of the sofas. A video camera hangs above the central TV monitor, facing the sofa. *Telematic Vision* works by linking together the two remote locations. The cameras in both locations record the scenes and fed it into a video mixer by ISDN line. Both images are mixed together by using the Chroma-key technique, and then the resulting composite image is sent back to the TV screens in both locations. The blue colour of sofas and carpets is necessary for combining the two scenes together into a single image. The performance begins as soon as a participant recognizes the scene – which Sermon calls 'domestic interface' – and therefore starts to behave accordingly: '[t]he viewers in both locations assume the function of the installation and sit down on the sofas to watch television' [19]. However, instead of seeing an ordinary broadcasting television screen, participants are offered a spectacle in which they are the actors: the audience becomes the scenario and theater of its spectacle. Not only do they see themselves in the screens, but furthermore, they see themselves seated next to somebody which is not physically there and which they do not know. 'At this point they enter the telematic space, watching a live image of themselves sat on a sofa next to another person' [20].

In *Telematic Vision* there was no possibility to exchange vocal messages. Audio was purposefully cut off by Sermon in order to avoid participants to interact only with voice. In so doing, Sermon compelled them to use other forms of interaction. Shanken offers some examples of interactions taking place among participants: mapping oneself onto the body of another participant, hiding some of the parts of one's own body using the pillows or other objects, melting, touching, caressing, dancing and making shapes together [20].

Apparently Sermon's Telematic Vision could be understood as another case of telepresence resulting from immersion and interactivity. As a matter of fact, the sense of immersion and illusion are very high and play a relevant role in the experience.

Yet, I argue, interaction and immersion are different from those implicit in the current discourse about presence and telepresence. Interaction is not unilateral activity. Actions are exchanged bi-directionally and synchronously among participants. The telepresence space created by Sermon is a space where it is possible to touch some other participant, but also to be touched by him/her; it is a space where it is possible to see the other but also to be seen by the other. Presence is not the result of being able to act or spy a target-participant, but is the result of relations of reciprocity. On the other hand, immersion is not the aim of Sermon's piece. In my opinion, immersion and illusion are secondary and functional to providing participants with the means of extending their consciousness in a virtual environment. The illusion deriving from immersion is reduced and lessened due to a sort of "mirror-effect", which is heightening the sense of awareness – in Brechtian's terms, it produces an "estrangement effect". In other words, participants can not only see and be seen, but they can also see themselves while seeing or being seen. Such kind of immersive technique, I believe, rather than aiming at illusion and embodiment, invites the participant to reflect critically on his/her actions. Consequently, the participants' presence is not so much the result of immersion and interaction but it arises from a sort of responsible attitude which derives from participating in Sermon's piece.

5. Presence: Ethics, politics and powers

The condition of presence, I argue, implies always an ethical dimension – one of responsibility – arising from the awareness of being in a reciprocal relation with the other and the environment. Tom Lombardo maintains, '[r]eciprocity has not only served as a primary mechanism for the creation of biological and social complexity, it provides a universal principle upon which human values and ethics are defined. Reciprocity is the foundation of the concepts of justice, equity, and perhaps even human care and kindness' [21].

Indeed, power is given by the lack of reciprocity, or as Jean Baudrillard reminds us, by the lack of "responsibility", that is, the impossibility to respond: 'power belongs to the one who can give and cannot be repaid. To give, and to do it in such a way that one is unable to repay, is to disrupt the exchange to your profit and to institute a monopoly' [22].

Therefore, besides an ontological condition, presence is also an inherently political condition, characterized by an equal balance of powers and responsibility (in both sense, e.g. the possibility to reply and a careful attitude). For instance, the possibility to act or feel, without exposing or showing one's own body, but from a safe and shielded standpoint, implies a different politics and distribution of powers.

To be in control of an environment, a thing or a human being, to have the power of affecting and acting on that thing, environment or human being – whether remotely or in the here and now – without being reciprocally open to the effects, consequences, or will of that "other" thing, in my opinion, cannot be defined as a situation of presence.

Tele-operation and other technologies sometimes unites through domination, namely they implies always a dichotomous relationship between a subject and an object. On the contrary, presence and remote presence unite through exchange and negotiations. When the network of reciprocal relations which characterize presence is limited or reduced either by technological or natural conditions, then presence is also negatively affected.

Conclusions

In this paper I have questioned the current understanding of presence, and proposed an alternative explication. Drawing on the theoretical framework provided by theatrical presence and drawing on the artworks of a group of artists doing telepresence art, I have argued that another explication of presence is possible: one consisting of networks of reciprocal relations naturally and technologically inhibited or promoted. Finally, I have also pointed out the political and social issues related to presence.

Of course, there are also non technological factors that prevent us from being in a condition of presence. Cruelty, hate, vengeance, racism and other kinds of discriminations, poorness, sickness, and many more, have in many occasions turned human beings in mechanical devices of destruction and violence, towards the other, the different, the animals and the environment. They have erased every kind of empathy and reduced our presence in the world.

References

[1] T.B. Sheridan. Musing on telepresence and virtual reality. *Presence* 1 (1). 1992.

[2] M. Minsky. Telepresence. *Omni*. June. 1980.C. Reeve, Presence in Virtual Theatre. *In Proceedings of the BT Workshop on Presence in Shared Virtual Environments*, BT Labs, Martlesham, Ipswich, UK, 10-11 June 1998.

[3] P. Virilio. *Open Sky*. trans. by Julie Rose, London and New York, Verso. 1997 [1995].

[4] L. Manovich. To lie and to act: Potemkin's villages, cinema and telepresence. In *The Robot in the Garden: Telerobotics and Telepistemology in the Age of the Internet*. K. Goldberg (ed.). Cambridge, (Mass). The Mit Press. 2000.

[5] T. Lombard and M. Ditton. At the heart of it all: the concept of presence. *Journal of Computer Mediated Communication*. 3 (2), 1997.

[6] W. Ijsselsteijn. Presence in the Past: What can we learn from Media History. In *Being There: Concepts, effects and measurement of user presence in synthetic environments*. G. Riva F., Davide, W.A. Ijsselsteijn (Eds), Amsterdams, The Netherlands, Ios Press. 2003.

[7] W. Ijsselsteijn and D. Riva. Being There: The experience of presence in mediated environments. In *Being There: Concepts, effects and measurement of user presence in synthetic environments*. G. Riva F., Davide, W.A. Ijsselsteijn (Eds), Amsterdams, The Netherlands, Ios Press. 2003.

[8] P. Zahorick and R.L. Jenison. Presence as Being-in-the-World. *Presence* 7 (1). February. 1998.

[9] C. Heeter. Reflection on Real Presence by a Virtual Person. *Presence* 12 (4). August. 2003.

[10] W. Benjamin. The work of art in the age of mechanical reproduction. In *Illuminations*. New York: Schocken Books 1962 [1936].

[11] M. Costa. L'estetica dei media. Avanguardie e tecnologie. Roma: Castelvecchi. 1999

[12] R. Copeland. The presence of mediation. *The Drama Review* 34 (3) 1990.

[13] G. Manotvani and G. Riva. "Real" presence: how different ontologies generate different criteria for presence. *Presence* 8 (5). October 1999.

[14] P. Salvini and C. Laschi. From Robotic Tele-Operation to Tele-Presence through Natural Interfaces. Paper presented at *Biorob 2006. The first IEEE / RAS-EMBS International Conference Conference on Biomedical Robotics and Biomechatronics*. Pisa, Tuscany, Italy. February 20-22. 2006.

[15] M. Merleau-Ponty. The Intertwining - The chiasm. In *The Visible and the Invisible*. C. Lefort (Ed.), trans. Alphonso Lingis, Evanston, Ill.: Northwestern University Press. 1968 [1964].

[16] P. Burke. The flesh as urprasentierbarkeit in the interrogative: the absence of a question in Derrida. In M. C. Dillon (ed.) *Ecart and Difference: Merleau-Ponty and Derrida on Seeing and Writing*, New Jersey: Humanity Press. 1997.

[17] E.A. Shanken. Tele-agency: telematics, telerobotics, and the art of meaning. *Art Journal*, 59 (2), Summer, 2000.

[18] D. Back and N. White. [Online] Available: http://www.normill.ca/artpage.html)

[19] P. Sermon. From telematic man to Heaven 194.94.211.200. Public lecture 1997. [Online].

[20] P. Sermon. Private interview by Pericle Salvini. 2004.

[21] T.J. Lombardo. (1987) The reciprocity of Perceiver and Environment. The evolution of James J. Gibson's ecological psychology. Hillsdale, NJ: L. Erlbaum Associates

[22] J. Baudrillard. Requiem for the media. In *For a Critique of Political Economy of the Sing*, trans. by Charles Levin, Telos Press: St. Louis, Mo, 1981.

What Production of Presence and Mimesis have in Common

Jan Söffner
University of Cologne, Germany
{amr69@uni-koeln.de}

Abstract

The aim of this paper is to open up the discussion on presence to questions about the influence of the presumed ontic status of the perceived on presence experience. In order to develop an approach considering this item, experiences of tele-presence have to be distinguished from those of 'pure' presence on the one hand and 'fictional' presence on the other: Fictional presence depends on an 'as-if' experience that is due to a communication act, whereas 'pure' presence is due to the belief in an un-biased perception that allows also non-communicative items to occur. Tele-presence is a grounds for intersection between these two issues that allows an experience of two realities to 'take place' in the same time and place. The relations between these three concepts among each other and regarding virtual presence (as a broader concept) are discussed and a definition for hyper-presence is introduced.

... a presence to be felt and known.
(Percy Bysshe Shelley, Adonais)

1. Introduction

The Vatican recently decided to declare a papal blessing valid if watched in real time transmission – and void if watched from a recording. Evidently Roman Catholicism not only counts on the ontological premises of a divine presence in its rituals – but also on its transmission (a belief that evidently has a tradition in iconological theories of real presence deriving from the middle ages). The concept underlying the papal definition, though, also occurs in more secular contexts; e.g. the sense of watching a recorded repeat can profoundly affect the experience of a spectator – even if the sensual stimuli and information factor are the same. An analogous difference can also occur in the case of interactivity – when 'being there' is equated to 'doing there' (cf. [27]) in real time, and thus if phenomena of immersion (cf. [14], [18], and [25]) are concerned. It is problematic to distinguish a given person's ontology from her/his perception, for the simple fact that we cannot see them as two psychological entities completely separated from one another. This means, that in order to describe the whole phenomenon of presence experience, the conditions of involvement (understood in Singer and Witmer's sense as "a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events" [42], p. 227) and immersion (the psychological "state of perceiving oneself to be enveloped by, included in, and interacting with an environment" [ibid.]) are not sufficient as long as one does not consider the presumed "reality status" (an issue that, at least according to Erwing Goffman, is implicitly present in any framing process organizing experience – cf. [17], p. 8). In other words: Research on the experience of 'being there' must question the status of this 'there' – it must ask, whether it is considered to be a 'nowhere', an 'elsewhere', an 'everywhere' or even something else – as in the first example (because Christians believe God to be elsewhere *and* here at the same time, as He is beyond earthly concepts of time and space).

Evidently these premises are in contrast to the branch of presence theory that regards tele-presence experience as due to "willing suspension of disbelief" (cf. [10]). In saying this I do not want to question the excellent work that has been done according to this definition – though I think that it regards only one part of possible presence experiences. Indeed, "willing suspension of disbelief" (cf. Samuel Taylor Coleridge, *Biographia Litteraria*, chapter XIV) being one of the most famous definitions of the conditions for "poetic faith" to occur, the kind of presence defined this way cannot be but fictional (for a discussion of this term see chapter 4 and 5). To put this question in a more concrete way: The fact stated by Aristotle (*Poetics* 4; 1448b), that we can feel pleasure in watching the most accurate imitation of something we would find horrible and disgusting in reality, can still be considered proof of this fact and remains open as a question – especially for presence research: If one of the participants of a video conference should suffer a heart attack, it should at least be hoped, that nobody on the other side of the transmission would see this with pleasure – even if she/he is a fan of horror movies and violent video games.

2. Presence Ontology

2.1. Philosophical approach

Of course the question about ontological aspects of presence have been broadly discussed. The simplest approach has recently proposed by Luciano Floridi [13] who put forth a definition of tele-presence – opposed to approaches discussing tele-presence according to what he calls "epistemic failure". In order to give a different solution his attempt is to base presence theory on what he calls "successful observation" – conceived as openness towards information (he gives the example of a motion detector), with the condition of the 'observer' being a "property bearer" and/or "source of interaction" in a remote area. Presence and tele-presence for him are thus not ontologically different from one another – they differ only by means of remoteness. Accordingly tele-presence, for Floridi, is to be distinguished less from simple presence and more from tele-epistemics (the latter being the ability to get to know something about the remote area while not necessarily being present there as an acting or property bearing entity).

Indeed this model seems to be an accurate description of the ontological side of presence devices. But as (on purpose) it lacks any kind of psychological issues, some questions may arise to its usefulness for describing presence experience. Floridi's reason for this limitation is his purpose to avoid Cartesian subjectivity. Though one could object, that not all psychology is Cartesian (for another option see chapter 2.5.). And lacking this consideration one can also argue, that Floridi's proclaimed "non-Cartesianism" (ibid. p. 666), if applied to psychological questions about presence, even shows some strong affinities to Cartesian thought, which non-Cartesian psychology could avoid. This fact can easily be shown, by comparing Floridi's concept of "successful observation" with James J. Gibson's thoughts about successful action (cf. [16]). Gibson's theory sees perception as a part of action in an environment – and this state of embedding is due to avoiding distinctions between sensual perception and conceptual representation. By distinguishing an 'observing' from an 'observed' instance, Floridi undermines this state – and clings to essential aspects of Descartes' thought: In talking about observation we still have to deal with something that changes the mere impulse of something present which requires space (a *res extensa*) into something represented and therefore abstract (which is the task of an *ens cogitans*). To say it in Niklas Luhmann's terms: A "first order observer" (about whom Floridi is implicitly talking) excludes himself from the observed as a "blind spot" (and it will be the "second order observer" to observe the first order observer in this status of being excluded). And if one starts to talk about psychology, an 'observer' (be she/he a first order or a second order observer) must therefore be defined as the very entity, that *does not* feel its being imbedded into its environment – otherwise she/he would be an immersed participant and not an observer. Accordingly the likely misunderstanding of equating Floridi's "flow of information" (ibid. p. 666) with the famous sense of 'flow' in presence experience must be avoided: The first is a metaphor for the process as a whole – the second one for a particular experience from within that can also be due to other issues.

2.2. 'Ecological' Approach

Instead of trying to describe only the 'real' ontology of presence therefore also the ontologies of given people have to be considered. This fact has been brought up by Giuseppe Mantovani and Giuseppe Riva – who also tried to scrutinize theories on the base of 'epistemic failure' – though this time by stating that "actors, that move within VE [virtual environment] are aware of the fact that they are interacting with a synthetic environment" ([30], p. 542) – and that therefore one cannot do without ontology (or better: "ontologies", as they correctly put it). To solve this problem they developed an "ecological" theory about presence based on a constructionist epistemology. This theory though, is only apt for describing ontologies as issues in the interaction of persons and their environment from a sociological point of view. This point of view though, cannot be presumed to be the meta-epistemology of all given persons possible – and therefore the trap about this approach is the risk of confusion between theoretical description parameters, and judgments about presence experience of the people acting 'from within': And this way they lack consideration about the issue where sense of presence mainly takes place.

Very promising seems the approach of Zahorik and Jenison [27]. Convincingly in the beginning these authors put forth a definition of presence as 'being-in-the-world'. In introducing Martin Heidegger's thoughts about "throwness" (Geworfenheit) and "readiness-to-hand" (Zuhandenheit) into presence theory, they allow for a distinction between presence as existence in a worldly context defined as state of acting on the one hand, and on the other interpretation requiring stable (mental) representation that takes place necessarily outside action (for a closer discussion see chapter 2.5.). The problem about this approach is though, that by afterwards equating not only presence, but also 'being' with 'throwness' and 'readiness-to-hand' Zahorik and Jenison exclude fundamental aspects of Heidegger's ontology (or even ontological concerns as such) from their interest. 'Being' for Heidegger is not only due to immanent acting, but also to transcendence taking place in this immanence. Without considering this aspect of ontic 'openness' (Erschlossenheit and Entschlossenheit) we are indeed much nearer to what Heidegger called "Benommenheit" (takenness): a pure state of attention in which only the fact of 'inhibitness' or 'uninhibitness' (Enthemmung) towards possible actions count (cf. [21] p. 352f.; 361). And it only is this missing consideration that allows for the equation of the prior with James J. Gibson's ecological theory about perception. And – even more important for the aim of my paper – it is also due to this lack of ontological 'depth' that Zahorik and Jenison do not and cannot distinguish between environments presumed as 'virtual' and those presumed as as 'real' (cf. [27], p. 87) – a distinction that was unproblematic for Floridi.

2.3. Semiotic Approach (with many thanks to Michael Cuntz)

Umberto Eco has proposed a different solution (cf. [11], pp. 337-392, especially 363-386). Similar to Floridi's theory of successful observation, Eco defines television at its "pure state" (understood as 'closed-circuit' real time presentation) as having to be sharply distinguished from cinema and photography – but not from mirrors. The reason is the following: If mirrors were (hypothetically) placed in a kind of chain formation, images of something that could not be seen from the place where this chain ends, could be 'transmitted'. Accordingly real time TV and mirrors are *prothesises* of human perception: They show things in a state of presence (things that would vanish from the screen if physically taken away). Cinema and photography are instead *communication media*: They are in a position to imbed images taken from the physical world into a communication act, because they 'freeze' them – and by doing so they make them refer (indexically) to something *absent*.

Evidently, speaking in Floridi's terms, Eco here is dealing only with tele-epistemics and not with telepresence; but this distinction is not valid for Eco's view: His definition of presence and absence is not grounded on the ability to be interacting or to bear properties in a remote area – but on a semiotic model of the sign as "something that stands for something else in its absence" (ibid. 368): Presence for him is accordingly defined negatively as a fact of non-signification. Everything that can be perceived is thus present as long as it is perceived as itself and not in the state of referring to something else.

Nevertheless this solution causes another problem: And this problem is about presence as an effect: Even the most hypothetical 'chain' of mirrors cannot present, say, Jupiter other than as the small spot which as it can be seen by the naked eye. For the logical restriction to categorical questions based on the sharp distinction between presence and representation, this is no problem: For Eco Jupiter is present – and this is all there is to say about it. For the question of the experience of immersion and involvement though, the factor of improved perceptibility is essential: Presence experience will begin only, when Jupiter is looked at through a telescope – and the question is, whether a toy telescope be sufficient for this experience, and how sense of presence could be improved by using, say, the Hubble-Telescope. Analogous questions can arise to the TV image's ontological presence considering its capability to lie (capability essential for Eco's concept of a sign): Of course it is true that a closed circuit real time TV transmission is ideally not able to lie about existence (because it is "produced only in the presence of the object") – or its 'thisness' (or *haecceitas* to use the term put forth by medieval philosophy). Nevertheless, it *is* able to lie about its qualities or properties (its 'whatness' – *quidditas*) by means of presenting it as being located elsewhere (i.e. near), lacking the third dimension and olfactory issues, being bigger, smaller, distorted, black and white, and so on. Eco's limitation of defining presence only according to thisness or *heacceitas* makes any whatness or *quittitas* only a question of more or less *further* information, that has nothing to do with presence itself. Presence research is about gradual questions – this definition instead only answers a categorical one.

But Eco's approach is not only problematic for the psychology of presence – it is also problematic for semiotics of presence: The capability to present half-truths by means of biasing the presented items indeed opens a realm not only between deception and reality, but, by doing so, it also shows a space between sign and non-sign – an ambiguous realm, which Eco fails to consider. For example: As is well known, so-called real time transmissions always present a certain delay. This delay is physically reducible only to the speed of light, and it can be extended as much as the director likes. So where does 'freezing' begin and where does presence end? Is a soccer match transmitted with a ten second delay already lacking the status of presence and introducing the status of communication? The answer would probably be yes if we have to deal with a human intervention delaying the transmission and no if we have not. This answer, though is not very convincing if one compares this case with the 'presence' of Proxima Centauri, which without human intervention shows up here with a four year delay – a knowledge that can deeply influence presence experience (everybody knows the experience of asking whether a given star that she/he sees still exists). A third problem arises again from the theory of the observer implicated here (or even the lack of such a theory). This problem can easily be seen when considering the case of a *fata Morgana*: Technically speaking this natural phenomenon can be considered a mirror – it is thus a 'mirage'. Its more mystical name though, due to the fabulous enchantress Morgana, derives from the fact that for an unprepared person it could easily be understood as a miracle close to visionary experience (and thus, according to Eco's own presumptions as a sign). Now indeed when speaking about psychological phenomena of involvement and immersion the question is not, what a *fata Morgana* is in optical theory, but as what it is perceived. And, unusually for his thought, Eco's theory of presence – like Floridi's – lacks the consideration of differing ontological presumptions.

2.4. Anti-Hermeneutic Approach (with many thanks to Michael Cuntz, Florian Mehrltretter and Andreas Kablitz)

Hans Ulrich Gumbrecht [19] recently proposed a different concept of the sign. From the title of his work *Production of Presence – What Meaning Cannot Convey* it is clear, that his thoughts see signs not only in the function of simple signification. As an alternative option he opens also the process of "coupling […] a substance (something that requires space) and a form (something that makes it possible for the substance to be perceived)" (ibid. 81-82) as a condition for a thing to become a sign. One has accordingly to conclude, that there are two ways for sense to arise: Either it occurs in a purely mental way (according to semiotic theory) or as located between sensual perception and mental cognition. This second concept of a sign allows therefore for a condition, in which the "spiritual" aspect is inherent to a material one, so that it cannot be conceived as absent but rather as present: If for Catholic ontology Christs spiritual substance is present in the consecrated host, for a believer this spiritual issue can be experienced in its material form. Another example would be the presence of, say, the Boca Juniors in their soccer stadium: Put in a Saussurian way, the players would simply *be* the Boca Juniors – or *signify* the Boca Juniors in their absence. Though both solutions are not very convincing: Nobody would go to a stadium to see something absent – nor would anybody believe the players to simply *be* the ideal 'spiritual' essence of the team (if they play badly, spectators will insist that they be substituted in order to make the true Boca Juniors more real). For Gumbrecht's second kind of a sign, though, presence would be a question of a form that is fulfilled in a more or less felicitous way (so that the 'real' Boca Juniors would be present only in their best actions and in the most intense interactivity between the players, the audience, and, most important: Diego Maradona). This example shows, how open this concept is towards gradual questions of presence. Though it also shows that the conception of form ("something that requires space") has to be opened to temporal aspects in order to be relevant for presence theory: It also requires time – be it durative or be it ephemeral.

This concept of 'signification' offers some advantages for presence theory. First: It cannot be without consequences for the theory of the observer: Gumbrecht's sign standing for something in its presence sees not only the signified, but also the understanding entity (the *ens cogitans*) as being in the same ontological space – and unlike in Floridi's theory, Gumbrecht does not imply a theory of the observer that would necessarily have Cartesian presumptions if applied to psychological questions: His Anti-Cartesianism is not based on the materialization of the *ens cogitans* (as was Floridi's), but on the materialization of signification as such. Second: Being the process of sense occurring now due to the distinction of material being and formal qualities it has no problems regarding the 'whatness' (quidditas) of presented items. Therefore questions about bias factors etc. are even essential and not accidental to it. The problems mentioned about bias in tele-presence (even the delay factor in a real time transmission) can now be scrutinized as facts of making presented items *more or less* present. Third advantage: According to this theory there are even no problems regarding perceptive issues about presence: If for Eco the possibility to "relate a sign token to a type" ([11] p. 368) is constitutive for a signification taking place afterwards, Gumbrechts sign makes 'signification' already take place inside this primary cognitive process. By opening a) the signified towards perception and b) signification towards presence, this approach promises therefore *vice versa* to open phenomena of tele-presence towards aspects of signification.

When considering this kind of signification, some questions can arise though, whether this sign is really a sign – or rather an ordinary token-to-type cognition: What can it be that makes Gumbrecht's "spiritual" aspect exceed the simple mental representation of a type? Indeed I think that this 'sign' has nothing of a sign, unless a special kind of ontology can be presumed when perceiving it – an ontology changing cognitive types from something purely mental into something "spiritual", universal or true: A similar ontology does not seem very close to our modern eyes – even though it shows at least affinities to e.g. the epistemology allowing a conception of a physical law to literally take place within a given phenomenon. A more striking example for such a kind of ontology is that of certain medieval philosophers, who conceived some 'types' as 'universals' – or platonic thought conceiving them as 'ideas'. Though there are still existing contexts (or 'frames' – cf. [17]) in which similar ontologies can occur: E.g. the soccer stadium of the Boca Juniors.

The most serious problem for presence research is though again the purpose to do without psychology also present in Gumbrecht. While reading his text, at times I could not figure out, when he was talking about ontology and when he was talking about effects of ontology on experience. I myself feel neither capable of developing a complete ontology – nor is this my aim. When speaking in terms of psychology though, some issues about presence have to be put more clearly (e.g. token-to-type cognition cannot only occur in cases of involvement or immersion – but it can also take place in quite a skeptical way; it therefore has not necessarily to deal with presence experience.

2.5. A Psychological Approach to Presence (with many thanks to Klaus Martin Schulte and Hans-Georg Soeffner)

To sum up: Presumed ontology cannot be searched for without considering the 'ecological' aspects 'being-in-the-world' – nor without taking into account effects of transcendence. The latter aspects though, cannot be considered without at least thinking about signification and signs. Vice versa, though, theories of the signs must be grounded on a psychology open for ontic as well as for ecological questions – otherwise they remain irrelevant for presence theory. The centre of consideration must therefore be psychology – and to show the ways a similar psychology could go I now want to recur to some aspects of the phenomenological psychology of Erwin Straus [45], i.e. his distinction between 'livingly experiencing' (erleben) and experience of 'understanding'/'getting to know' (verstehen).

This distinction is quite close to (and partly deriving from) Heidegger's distinction between 'being-in-the-world' and interpretation. The most important difference between the two options for Straus is about constructing time and space: The Euclidian space with its three dimensions and their being differentiated from time (at least in the non-Einsteinian everyday knowledge about the world) for Straus is open only to our understanding – and this understanding is accordingly defined by an abstract and extraterritorial point of view. Living experience is instead marked by the symbiosis of a being (be it human or an animal) that is capable of motion with its at least potentially moved environment. In this latter experience time and space are not divided one from the other. 'Living experience' occurs in and as a continuum without dimensions – or with only one: distance. The concept is therefore about 'being able to touch' or 'being able to be touched' and the length of the time-space this event would require to happen. This does not mean, though, that this continuum of experiencing was simple: Evidently there is not only one distance, but many of them. The experiencing person therefore finds herself or himself lodged in to several options of motion, like a spider in its net. The border of the perceptive horizon is to be seen as the realm from where things coming *appear* and into which things passing away *disappear*. And this coming and going is all there is about past and future: 'Living experience' takes place only in a present that is in continuous change, into which always something can intrude, and in this way its field never can be overlooked a whole: Instead it remains in a constant state of 'un-concludedness' ("Un-Ganzheit"). And (like in Heidegger's concept of 'throwness') it is only this condition of "Un-Ganzheit" the reason, why 'living experience' is not open to 'understanding' and *vice versa*: Only concluded issues can be 'understood' – and this concludedness cannot occur as far as the being is included in the space of living experience. Understanding therefore requires to get but rather extraterritorial to the understood (condition for the Euclidian space differentiated from time to occur); it necessarily divides a knowing subject from its known object (or an *ens cogitans* from a *res extensa* – to say it in Cartesian terms). "We live in the present tense and know in the perfect" ([45], p. 393) – is Straus' conclusion – but this metaphor is not quite precise: Perhaps it should be reformulated as follows: 'We live in sentences without – and we understand in sentences with a finite verb.'

These psychological insights, further developed by Merlau Ponty ([31] pp. 314-315) who theorized 'living experience' as the "space of the present" (champ de present), have been brought to new actuality by neuro-psychology, where a very similar concept of a 'primary' space of senses has reappeared under the name of "motor space" (cf. [38]). Considerations have already been fruitful for the theory of presence and immersion (cf. [25] and in a certain sense also [14] p. 268). Also Frank Biocca and Jin Kim Yung Choi [6] seem to be referring to a similar conception when they state that the "human sensorimotor system is designed to experience the world as whole, merging and synthesizing input form different sensory modalities in an ongoing and dynamic form". The term 'sensorimotoric' though, has to be understood in a broader sense of the word: The described state of experiencing evidently results from cognition without which a constitution of a "form" would be impossible. We therefore have to deal with a complex construction of an environment – a one-sided one, unable to reflect its being split into multiple realities – but nevertheless highly complex: It is built of interferences from perception, motor activities, and results of learning (e.g. the spontaneous 'knowledge' about stable and unstable surfaces or 'mental mapping' (cf. [23]). This relatedness of motor cognition to a sort of knowledge (or better: a set of presumptions) is – so to speak – the 'missing link' between ecological theories of perception and action and the immanently experienced items Heidegger would have conceived as ontological openness (even near to his concept of "unhiddeness" – "Unverborgenheit"; an etymological translation of the Greek term *alêtheia*). Though not wanting myself to get too metaphysical, I would like to recur to a more psychological and concrete way of describing the phenomenon – i.e. Edmund Husserl's concept of "appresentation" ("Appräsentation" – cf. [24] pp. 111-123): Husserl with this term refers to a status of "also-there" ("Mit-da") in perception, i.e. things that spontaneously occur, even though they are not perceived in a physical way. He himself gives the example of the reverse or far side of things (that occurs as 'present' also in motor perception – even though a camera would not be able to 'see' it). Accordingly Husserl conceives appresentation not as due to an "act of thinking", but as part of "primordial perception" – a state in which everything 'other' can only occur as part of the 'own' (of course it is this very theory that has deeply influenced Heidegger's 'being-in-the-world' as well as Straus' 'living experience'). Though Husserl does not limit his thoughts to far sides, presumed weight, softness or so on – but he puts this concept at the base of the perception for describing others as 'subjects' that occur as presumed bases of their own "original sphere"; i.e. highly complicated phenomena of 'depth' that can be presumed as inexhaustible – and for which Husserl accordingly uses the term "horizon". To say it again: These phenomena occur already to primordial perception, to which in his terms everything can appear only as part of the same continuum: Therefore appresentation for Husserl is a fact of "immanent transcendence". Straus himself was expert on Husserl's phenomenology. And he drew according conclusions about the construction of 'lived' reality: 'Living experience' for him is a grounds for the intersection of sensorimotoric and ontic issues in the experience of being embedded in a spatiotemporal continuum or, again 'ongoing form' – i.e. the status of un-concludedness mentioned above.

2.6. Being there

Similar distinctions about conceptual understanding and immersive perception, interpretation and being-in-the-world or even perception for action and perception for conceptual cognition (for the latter see [32]), can have far-reaching consequences for the psychology of signification. 'Concludedness' as a property of signification and its relatedness to understanding has accordingly been discussed by Andy Clark [9] who described the process of signification in a way, that by means of "'freezing' our own thoughts in the memorable context-resistant, modality-transcending format of a sentence, we [...] create a special kind of mental object – an object, that is amenable to scrutiny from multiple angles, is not doomed to alter or change each time we are exposed to new inputs or information, and fixes the ideas on a high level of abstraction from the idiosyncratic details of their proximal origins in sensual input" (ibid. p. 210). This kind of redefinition of Pierce's concept of "thirdness" as a generalized property made independent from its property-bearer by mental means, is based on the assumption that understanding experience is open to a semiotic concept of signification only – being a 'signified' ideally something ideally concluded and therefore absent: A "mental object", that is freed from its links to sensual input by the same means also constitutes the entity which thinks the object in a state of being 'extraterritorial' to it (because of its being concluded) – and at the same time independent from sensimotoric issues. The (semiotic) sign is accordingly closely related to the condition of the Cartesian *cogito* and of Straus' "understanding": The distinction between a signifier and a signified (i.e. the independence of the signified from sensual data) can be seen as the basis for the extraterritorial status of observing, which is necessary for both of these concepts. Indeed Clark therefore conceives this linguistically formed approach to the world as an approach of "second-order cognition".

Nevertheless to control "real-time, real-world responsiveness" ([9] p. 8) (i.e. things occurring inside 'living experience') not only direct "tuning of basis responses" (ibid. p. 4) plays a role. To give a simple example: The sense of 'knowing' that God is helping (i.e. being present as an agency inside the own actions) can strongly influence experiences and even actions: The fire-man who saved the veil of Veronica at Turin stated afterwards that he would not have succeeded without the help of God (i.e. in my mind without believing God to be present in his actions). Similar experience is due to ontic 'depth' that can occur as appresented in 'living experience': If the ontic status of a given 'reality' is due to its depth (i.e. its not being concluded), this ontic status is not only open to epistemology (i.e. the possibility of controlling the ontic status of a reality by means of sensual objects already stated by William James – cf. [26] p. 301): The relatedness to 'the' world can indeed occur also as being co-present (appresented) as an additional dimension of depth inside the environment that in a certain way makes a second reality occur inside „motor space", "champs de present" or "living experience".

In my eyes only the Gumbrechtian theory conceiving some signs as being able to reach a kind of generality or 'thirdness' taking place *inside* the same sensual world, allows for the conception of the full range of this interrelation between presumed ontology, sensual cognition and 'living experience'. And it does so in theorizing signification as a work on the form – in a very broad sense of the word: As stated above: 'Form' is for Gumbrecht is the mere condition of something being perceived – and this means: it can also be perceived in a state of 'living experience'. And to finally let the cat out of the bag, this work on the form or shape is at the base of my definition of tele-presence experience: *A non-representational (a-semiotic) coexistence of (at least two) different beings – or better: the experience of two concurrent presences overlapping in the same continuum of time and space by means of one taking the form of the other.*

The relation between these two presences can be manifold, and indeed is not a problem, but the most important issue about it: It can be the concurrence between the senses of distance (optical and acoustical) and the senses of unmediated spatial contiguity (the haptic and the gustative – the olfactory sense is to be seen as intermediate) – linking some spiritual issue to the former (for an example see chapter 3). It hypothetically could even be the concurrence of a complete being 'there' involving all the senses while only knowing to be 'here' (e.g. the hypothetical case of a pilot knowing to be steering a real aircraft by means of remote control from the perfect copy of a 'real' cockpit procuring all sensual stimuli the 'real' aircraft provides). It can be, though, also the concurrence between a purely spiritual item and a corporal one (as in the case of Christ's presence in the sacred host that nobody can feel or taste – if no miracle occurs). The only four conditions about this kind of presence taking place are:

1) That the present and the tele-present item can be separated from one another (because otherwise we would have to deal with simple presence) – and that distance plays a role for the distinction (otherwise the prefix 'tele' would be senseless).
2) That the relation between the two items is not of mere signification but of realization of a form, so that both can be seen as present at the same time and in the same place (and not the one to be the absence of the other). This allows therefore for a gradual condition of the fulfillment of presence because of the fact that a form can be realized in a more or less perfect way.
3) Tele-presence must have access to 'living experience' allowing for immersion to occur.
4) The presented item must have a presumed ontic 'depth' that is appresented. This factor can be lived as a presumed relation to simply a remote part of 'reality', but also as a relation to 'truth', to the 'infinite', or else; and it can e.g. occur as the "un-hidden"

(Heidegger) or of a human being as a presumed other "original sphere" (Husserl – cf. [24] p.113) or agency (cf. [15]).

Indeed virtual communities can be seen as a striking example for the last of these conditions. In discussing this phenomenon Mark T. Palmer ([36]) limits his considerations to phenomena of communication between humans – excluding relationships between humans and machines, which he considers as deceptive. Though it is well known that machines have no intentions – and therefore they cannot lie. Humans can lie – also and especially if they construct avatars. So why should deception be an aspect of machines? Considering the so called "Eliza Effect" (cf. [4]) of a machine behaving in a way that people would believe it to be steered by a human, the reason, though, gets quite clear: Deception occurs here as a falsely presumed of 'depth' vs. 'surface': The machine has nothing 'behind' the appearance, it only has a surface – whereas the presumed human agency would have another ontic status due to a human avatar being unpredictable – i.e. relatedness to something un-concluded.

3. Antique Theories of Mimesis

Evidently this option of defining presence especially the aspect of two presences being there at the same place and moment without a fact of representation occurring seems quite unusual for our modern eyes (and I think, that this is also the reason, why it has been taken into consideration only by Gumbrecht). Indeed very few theoretical research has been done into this direction and the concept accordingly seems quite vague or week. I nevertheless hope to find a remedy for this problem in my attempt to reanimate antique theorems.

The oldest example of tele-presence can be seen in the Homeric conception of the bard (aoidós). When for example the bard of the Iliad invokes the present Muses to sing about the Danaeans coming to Troy he states that ten throats, ten tongues, an indestructible voice and a heart of iron would not be sufficient to reveal the number of the heroes (cf. *Iliad* 2,484-93): Evidently the mere information plays no decisive role here, but rather, as Grace M. Ledbetter stated, a „divine knowledge that has the immediacy and pleasure of sensory experience“ ([33], p. 13) having to present a „past object in a way that present objects are known“ (ibid. 21). The suggestive or even synaesthetical power of chant can also be concluded from the Sirens episode the Odyssey (12,39-54 and 12,158-200). Their song makes their island – the shores of which in 'reality' being covered with the fleshless bones of their victims – seem beautiful and inviting. The Sirens are c indeed limited to two senses – the optic and the acoustic one, i.e. the 'bodyless' senses of more 'indirect' presence. And also their existence seems to be beyond the present: They know the past and the future, but they evidently know nothing about Ulysses being bound to the mast of his ship and his companions having their ears anointed with wax. Accordingly they address Ulysses not, as usual, by calling him 'many-sided and artful' (polymetes), but "great glory of the Archeans" (what he was in the battle of Troy and "renowed" (what he will be in the song of the Odyssey). The lack of the present is exposed as a 'higher' realm beyond concrete time and space. A parallel to this state can be observed in Ulysses' journey to the Hades. The dead he encounters there are similarly restricted to the optical and the acoustical: They can be seen and heard, but they cannot be hugged. And like the Sirens they know the past, the future and things far away – but they do not know the present. The status of non-corporality seems to be a great problem for the dead; indeed they are very eager of the blood of animals sacrificed to them evidently giving them a rest of haptic experience. Corporal presence evidently is the realm of true being in the Homeric world, whereas the hereafter is defined as a lack of corporality – even though it is related to knowledge inaccessible to living people: i.e. the dead seem to see things in their absence. The Sirens accordingly reflect a state of pure chant *as* a state of death – i.e. a state without a body that can be equated with absence from the present and at the same time presence of the absent. And this is possible as a fact of doubling presences by means of dividing the senses: The Sirens use song in a deceptive way trying to pull Ulysses' body into the space of their chant. That which is haptically present for the bound Ulysses contradicts the space offered visually and acoustically to him.

Though this quality of chant doubling presence can also be inverted into the positive: In singing and acting a bard can literally lend his body to the dead: He restores these optical and acoustical creatures with that what they lack – and therefore they will be lived as half-present, as can be seen from too emotional reactions (cf. Odyssey 1,337-344 and 8,485-545). Voice, mimics, and gestures have to be as close as possible to the truth (not the reality) of the heroes. This fact is made possible by a divine help not opposed to human technology – but at the base of it: The gift of the Muses (daughters of Mnemosyne, the Goddess of Memory) is a presence-based kind of memory, the place of which is performance. By means of metric rhythm (with a close link to the cultic dance) and melody of the language itself (due to the musical accent of ancient Greek) they grant a sacred order of time and space in song accessible to human practices (cf. [2] 61f.): A kind of divine high tech called mousikê ('the Muses' technique') allowing for tele-presence as accordance to a true form. The process of making present, some centuries later, has been given the name mimesis (mímêsis).

As Plato he discusses questions of education he accepts the concept of mimesis in all its facets (even musical issues can play a role for education – cf. *Laws* 664a-667e): E.g. in his *Republic* (392a-397b) a warden-to-be becomes a true warden by means of mimesis that leads him to correspond his behavior to the 'idea' of a warden – and therefore the ontic qualities of wardenship become his own. On the other hand he is very eager to avoid mimesis in its quality to lead into a double presence (as described above for the Homeric concept about it). His argument – best exposed in his *Ion* – is very simple: If in a bard's chant the presented persons really were present, the bard would have to have access to all knowledge and techniques of the persons exposed by him – which evidently he has not (bard is not a warrior – otherwise he would know, how to fight). Mimesis therefore is good, if leads to 'pure' presence; it is bad and deceptive, if it is there for itself and doubles presence (i.e. if we have to deal with tele-presence).

Aristotle takes up aspects of both the Homeric and the Platonic concept: Like them he conceives mimesis as presence of the presented characters in the presenting body of the actor – by means of perfectly forming it into the form the presence of this other requires: According to him by this means a "this one [the actor] is becoming that one [the presented character]" (*Poetics* 4,19); This doubledness of presence though, is no more ontologically problematic as it is in Plato: Like in the Odyssey it enables pleasure to be felt and even moral psychological and physiological purging (the catharsis). This fact is only possible though, because the reality status of mimesis is no more due to tele-presence in the presented items – but to the author's knowledge about probability (that on its side depends on truth). Instead of performance, it is creation, instead of relatedness to 'ideas' it is epistemic control that grants the ontic aspects: Truth is taking place in the probability of the things presented – whereas the presence factor is no more granting its ontology but is now a mere effect. Presence is no more connecting, but *actualizing* truth.

4. Presence and Mimesis

4.1. Fictional Presence and Tele-Presence

The third of these options is evidently much closer to modern theatre theory than are the first two. Indeed from Aristotle's theory of mimesis to a concept of mere theatrical representation it is only a small step. Already Roman theatre performance often was interrupted in order to have the 'best' scenes repeated – a practise evidently apt to undermine the experience of 'appresentation' of ontic truth necessary for presence experience. And I myself do not think it to be a simple coincident that Roman entertainment on the other hand had to recur to simple presence: In order to experience the presence of heroes, Romans had to watch real gladiators fighting.

Up to now corporal presence of an actor and represented role are understood as categorically distinct from one another in theatre theory (so that the one can only be seen as the absence of the other; and in the best of the cases this fact causes a switching of two according states of perception – cf. [12]). This distinction has also had effects on media theory. Of course it is true that cultural sciences nowadays deeply scrutinize the concept of mimesis (as imitation – slightly distorted meaning of the word deriving from translation into Latin) and even replace it by the notion of simulation (and thus focus the creativeness – poíêsis – instead of the ontological background); and it is also true, that concepts of fiction have essentially changed as they are often replaced by the notion of virtuality (focusing the status of construction instead of the ontological status as grounds for 'reality'). Though not finding a way back to a concept of presence based on the notion of accordance to a form, this epistemology is still based on the dialectics of presence vs. representation, so that a doubled presence cannot be discussed.

It is also on this grounds that presence research has partly followed the problematic way to define tele-presence as an 'as-if' experience – i.e. as 'epistemic failure, as willingly taking something represented for present by means of 'suspending disbelief': Because of this 'as if'-condition, I would like to call this kind of presence experience *fictional* presence. To avoid misunderstandings: Unlike in novels or movies the term 'fiction' here does not refer to what the things re-presented communicate, but only to the status of the 'presence' in which they are presented. E.g. a neat virtual reconstruction of a historic environment in a museum is a fact of fictional presence of (nearly) non-fictional issues, whereas a live transmission of a theatre performance of *The Tempest* is a fact of tele-presence of a fictional issue. In the case of fictional presence, presence experience is therefore due to 'surface' perception – i.e. to what in cinema would be caused by 'special effects': We have to deal with a presence working only on perception without including apperceptions of ontic 'depth'.

Even though both can occur in virtual environments this kind of presence experience has to be distinguished form facts of making present something distant (i.e. when talking about tele-presence. Here indeed our mediatic reality re-actualizes aspects of the more antique form of mimesis – i.e. the reality status depends not only on a truth communicated inside the presented world – but also on ontic depth in presence experience: The question about the simulation of the presented reality in this case is not only about its being constructed, but more than this it is about making a remote form present in a near substance. An example taken from fiction: In Jean Philippe Toussaints recent novel *Fuir* (Paris 2005) the protagonist, while having sex with a Chinese woman in the toilet of a Chinese train, is disturbed by the ringing of his cell phone. As he leaves the rest room to respond, his wife, drowned in apathy by the news of her father's death, describes the Paris sky she is seeing at the same time (cf. pp. 45-58). Like in Homer the acoustic presence of the voice (as such) is important in its fidelity, the exactness of its form: The presence experience would have been impossible via a Morse-Alphabet. Like in Homer this voice evokes a mental image causing a second presence concurring with corporal presence of the 'here and now' to arise – the narrator feels close to the emotions of his partner even though (or because) her apathy causes her to talk about something else. And like Ulysses upon hearing the Sirens' voice the narrator fails to construct a coherent reality of the double binding, he cries, and at the end he hugs his Chinese lover with a hug that is destined for his wife. The present woman is giving a corps to the absent, as the bard did to the dead (and in fact the narrator states, that to him the telephone has always been close to death – cf. ibid. p. 44).

By this means modern technology of tele-presence seems to show strong similarities to the Muse's technique. It is again the factors of the presented not being concluded and the taking place as a realization of a form (in Toussaint's novel the voice) allowing both to be present at the same time. And again it is the (presumed) ontic status of something being present in this form even if not completely 'here' that allows this presence to literally 'take place' – only that this ontic status is now much closer to our own secular everyday life than is the ontic status of the dead being present in the body of a singer or Christ's spiritual body being present in a piece of bread.

4.2. Tele-Presence and Hyper-Presence

Of course there is still an important difference between these issues: The cell phone makes present only something far away, whereas the Bard, the sacred host, and Plato's Warden make present something beyond space and time. Considering this item obviously requires a rediscussion about spatial issues. Stating the advent of a new spatiality that is due to changes in media technology is not very new or original: Similar things like those I am going to deal with in this chapter have been stated not only (but much more intensely) since the notion of the 'virtual community' came up (cf. [37]). The problem about these notions is that sometimes it is not very clear, if the interrelations between the 'virtual' place (for example a chat 'room') or space (for example the notion of 'links' and 'hyper-links') and perceptive models of space experience: E.g. the notion of cyberspace can refer either to a metaphorical space of the medium itself as "all-embracing power" (cf. [22], p. 91) – or to the concrete spatial experience of its users; the notion of a virtual place can either be understood metaphorically as a multi-user dungeon (because of certain analogies to e.g. a "corner bar" – cf. ibid. p. 24) – or it can be considered as being due to concrete perceptual experience of places "recreated" by technical means (cf. [46], p. 205).

The most interesting issue about these frequent confusions seems to me the interrelations they mix up – i.e. spatial closeness and abstract or spiritual contiguity. According to Euclidian parameters this is evidently impossible, because neither the multi-dimensional structure of e.g. the Internet nor its accessibility to everything structured by the number of 'clicks', nor its relatedness to the arbitrariness and predilections of the use can easily be translated into three dimensions. The fact, that nevertheless spatial metaphors, half-metaphors and concrete terms are constantly used to describe this 'space', though, makes it very probable, that a non-metaphorical 'sense of space' is involved in these media themselves (i.e. *not* in what they represent, but in what their technique as such is experienced by their users) – a fact similar to the ancient Greek's thoughts about music (the Muses' technique) as means of presence.

Generally this space beyond Euclidian spatiality is conceived as an "extension of our mental space" (cf. [1], or [22] p. 91—98). Though I think that it also has to be discussed in a way more open to 'living experience', 'champ de présent', 'motor space' or simply: To presence. The striking issue about this fact has (though without considering presence experience) stated by Rebecca Bryant (cf. [8]): According to her Cyberspace is structured by distances that have to be considered as temporal – a concept very close to Straus' definition of spatiotemporality in 'living experience'. And I now would even like to go so far as to define the spatial notions about technically transmitted presence as extensions of this kind of 'champ de présent' – in just the same way as Marshall McLuhan has defined media as extensions of our body.

This means though a redefinition of the concept of spatiotemporality – as quite different issues enter into the mere concept of distance: Contiguity or nearness are brought as much as possible in the line with the human arbitrariness of felt or desired affinity restructuring space and time experience. Even though one could even say that a similar kind of presence is also functioning in any kind of prayer, modern technology is constructing a new space of living experience that is in a much sharper concurrence with the material or Euclidian space than it has ever been. For the audience of a Greek bard or for a Catholic the experience of doubled experience was a quite unusual one. For us nowadays it is a part of every day life. E.g.: As I was traveling some moths ago in a train, a lady, with whom I had chatted to could not resist to the temptation to call the convent she had spent her youth in while the train passed it. Somewhat later she called her husband while passing a place they had passed together an important time in their lives. Modern communication technology creates new coherences in time space (opening new possibilities for emotionally coherent space and time experiences) and new discrepancies (see the example put forth by Toussaint's novel) that are not accessible to Euclidian 'dimensions', but to a space, in which the multiple distances can count also in their mediated form.

The problem for defining this kind of presence experience according to terms of tele-presence is the difficulty to talk about physical distances, about things 'far away' and thus far-present: *tele*-present. More than this we have to deal with another kind of spatiality making these distances even obsolete and allowing presence to take place in a higher, more spiritual space. Though again no new theory has to be developed, but an old (pre-Cartesian) one has to be adapted to our reality in order to describe this space. Again Plato plays an important role here; this time, though, we have to talk about the realm of the ideas as such. In his *Phaedo* Plato describes how it is possible for the souls to loose their heavenly status. By being hurt their form is damaged, and this loss of accordance to the forms of truth is equated to spatial distance: A hurt soul is no more able to follow the flight of its true essence, and it falls on earth, where it has to find a body. The realm of ideas is a realm where affinity is contiguity and discrepancy is distance. Of course this 'myth' has some metaphorical aspects. But it is difficult to say how metaphorical this equation of spatial and conceptual distance is. In fact also astrology counts on this kind of structure: The heavenly 'real' distances of the planets are to be seen as ever changing affinities (whereas the contingent distances of the sublunar world are simply casual). Also theories about paradise can show a similar structure: When Dante describes his *Paradiso*, or when Bernardus Silvestris his *Cosmography* in the realm of the heavens, this higher form of spatial coherence is the structuring principle (beatitude is conceived as spatio-conceptual nearness to God). Even more interesting for my purpose is the case of the highly influential (and equally underestimated) Italian renaissance philosopher and medic Marsilio Ficino. He conceptualized not only a similar 'truer' space (as the space of the platonic ideas), but also the means by which the material world could get in contact to it; and again this is due to a perfection of form and embodiment: The purging of the inner spirits (that – according to contemporary medical theory – are produced by the liver, purged by the heart and a miraculous 'web' in front of the brain in order to then move mental activities) through dietarily means, makes the mind more receptive to the higher forms (cf. *De vita* II). And if these forms are then transmitted by means of according the mind to them and thus perceive them from within, this forming process can be equated to an approximation in the other, truer, spiritual space: Therefore at the same time cosmological spirits enter – a procedure Ficino (*In Platonis Ionem*) defines as 'inspiration'. The inspired person then is able to 'be in' the higher reality – a state Ficino calls 'fury' (translating the Greek term 'mania' Plato used in his *Phaedrus*) – and to produce a higher, inspired, poetry that makes accessible the higher truth also to others.

Now I do not consider it to be a coincidence that, when Marconi invented the radio, the 'space' of the radio waves was called by the same term that not only astrologers and natural philosophers, but also Aristotle, Plato, the Platonists and the Neo-Platonists had used when talking about the fifth essence of intellect that governs the heavens – i.e. the word Ether. Also for the radio the Ether marks a second space beyond Euclidian distances, accessible to everybody everywhere – even though located nowhere. And in a way it is also structured by affinities – only that these affinities now are no longer the affinities of a higher truth, but rather those of human arbitrariness and predilections (switching on or off, or choosing the right channel). It thus links affine people to a common space beyond Euclidian dimensions in a less metaphysical way.

I would now like to call this special case of presence 'hyper-presence': First because it is structured by affinities accessible to and constituted by human arbitrariness as is the so-called hypertext. Second because (e.g. by people who want to marry on TV) it often seems to be experienced as the space of something 'above' things real. Third because it shows some similarities Baudrillards concept of "hyper-réalité"[1] (though I do not want to follow his concept of simulacra derived from Saussurian roots for phenomena of presence): In fact by means of special effects it nowadays is possible to make things much more present than presence itself (cf. [43]). Television is often closer to the action and has more perspectives on it than a real spectator ever could be or have – a fact made evident by the necessity of large video screens in stadiums that increase presence experience (even though they contradict the celebration of physical presence).

6. Conclusions: Presence – Fictional Presence – Tele-Presence – Hyper-Presence

The experiences of Presence (a sense of 'straight-forward' sense of 'being there'), fictional presence (the experience of 'being there' due to a presumed 'as-if' ontology), tele-presence (the doubling of two presences at the same place by overcoming spatial distance), and hyper-presence (the doubling of two presences at the same time by creating a new space) often cannot be distinguished neatly from one another. For example in Homer the dead heroes are tele-present in the bard if one considers the Hades as an island – and they are hyper-present, if one thinks of the Hades to be beyond sublunar spatiality. A more striking example of the interference between these concepts can also be seen in the complex phenomenon of virtual reality: Nowadays even a concludedly constructed virtual environment cannot often be overlooked, as it is designed by many different people who often do not communicate with each other about what they are doing. And it is by having less and less effects of concluded constructions that virtual reality becomes more apt to ontic presumptions (as a 'second universe') due to the fact that this way more 'depth' can occur as appresented. Even more interesting, though, is still the case of interpersonal virtual environments: I would even go so far as to state that if (as Mark T. Palmer states) "[interpersonally used] VR provides the user with control over the shapes, forms, and textures of his or her world" in order to "express mental and emotional states" in a better/more adequate way ([36] p. 294), this refers not only to what I define as the ontology of tele-presence: It even is a re-definition of Platonic ontology on a personal basis: The more a *form* is accorded to an essence considered as 'true' (i.e. in most of the cases also good and beautiful), the more present a given person feels as (i.e. in) her/his avatar. Virtual reality therefore seems to be a technical means of realizing (and trivializing), what Plato conceived as only spiritually possible (cf. [22], pp 91-86-91). On the other hand again this sense of the 'real' is also related to an appresented ontic 'depth', supposed in humans.

The latter of these conditions also allows presence

[1] In this latter way, the term has also been used by Waterworth and Waterworth ([47], p. 509).

technology for the adoption of "effects of the real" (effets de réel) once theorized by Roland Barthes [3] for fiction – but easily translatable into questions on virtual reality: The "effect of the real" for Barthes consists in the fact that a given item seems to be not invented. This "referential illusion" according to him is due to elements that escape a smooth construction of a text – allowing for what Goffman would have called a "frame shift" from a frame of fiction to real-world frame [cf. 17]. Evidently this kind of an effect is due to signification (in the semiotic sense of the word): Like in Aristotle the reality status is searched for as accordance of the represented to truth. Other kinds of "effects of the real" though, have been introduced e.g. into telenovelas long ago – and phenomena of tele-presence occur here: For example, in the case of extrapolations of an election taking place at the same time as the transmission (not as the production) being embedded into the fictional reality in real time (or nearly). Also in virtual reality 'effects of the real' can accordingly consist in dysfunctionality of a given item for construction (Barthes), in a sense of being related to something inexhaustible, and also in openness to influences from another inexhaustible reality (i.e. of elements or traces of tele-presence).

The problem for presence experience is though, that it can only be perfect where it is lived and (ontologically) presumed as non-mediated. Perfect experience of presence is therefore only ideally possible in the realm of tele-presence (in its concrete realization it would mean the perfect sensual and presumed ontological accordance of the two presences to one another – making the experience tele-presence therefore become an experience of presence itself). The same can be said for the concept of Hyper-Presence: In order to be really one with Christ's body (i.e. without this body also being a piece of bread), the believer would have to enter Paradise – in order to enter the 'true' realm of hyper-presence the bride and groom in a TV-wedding should see only their presence in the virtual place on the screen (but in fact they are in a place filled with cameras and spotlights). Perhaps it is this double lack (presence being less present in one way than hyper-presence and more present in the other), that generates the most sophisticated combinations of these three forms of presence experience – trying to give a body to hyper-presence and to make bodies more present than present:

While writing this paper I could not avoid getting involved in the euphoria the soccer world cup. And having spent one of the best years of my life in Florence, in the beginning 'my' team was not only the German, but also the Italian one. Now, as is also well known in the United States, the sense of being scattered all over the world is a very important aspect of Italian patriotism; therefore Italian TV does not only show the soccer games, but also – later in the evening – the reactions of the 'italiani all'estero' (the Italians abroad). My city has one of the most important Italian colonies, that uses to gather half a block away from my place to watch Italy's matches. So at the beginning of the world cup I took part in the following kind of event: In an Italian bar a large TV-set was placed onto the open street, where (important issue for the form of the 'real thing') the Italian (and not the German) real time transmission was shown. During the transmission – as in front of every maxi screen during the world cup – people would behave as if they themselves were in a stadium, while Italian TV-cameras were taking this event. Evidently some people behaved even more enthusiastic knowing that they would perhaps be shown on TV. Perhaps, one could suspect, they believed the players of the Italian team were going to watch their cheering after the match and thus be comforted and motivated for the next match – but this solution does not consider the fact, that people would behave just the same, when TV cameras were off. So where were people acting? To answer this question, I think it should be put in a different way, i.e.: Where was Italy? 1) On the soccer ground that was *tele-present* to the spectators. 2) In the *present* bodies of the Italians acting. 3) In the *hyper-presence* of television making a 'truer' Italy arise where all Italians are ethereally linked. The total and holistic 'Italy' this way evidently was present nowhere, but therefore it had to be presented everywhere – in the most coherent form one could give it on a grounds so scattered. I got more and more enthusiastic in 'getting into' this event.

This fractal kind of presence though, can also be a reason for failure of presence experience: When I returned to the same bar to watch the final, many things had changed in the meantime: Everybody knew about this place and its events, therefore too many people who just wanted to watch people watching were there, and I myself had been rather disappointed by the Italian way of winning matches in order to forget scandals: The particular form of Italy realized by this particular Italian team was one I did not want to be as present as the supporters wanted to make it. As I finally understood my disappointment leaving no solution, I went home – only to find out, that Italian TV was about ten seconds faster than the German one: From the cheering Italian crowds I always knew everything important already before it would occur on my television. Too many concurring senses of presence took place at the same time – blocking the last bit of coherent experience. Indeed I rarely felt so disappointed about production of presence as I did in the night when I finished this paper.

References

[1] Anders, P. 2001. Anthropic Cyberspace: Defining Electronic Space from First Principles, Leonardo 34, 409-416.

[2] Barmeyer, E. Die Musen – Ein Beitrag zur Inspirationstheorie, Fink 1968.

[3] Barthes, R. L'effet de reel, in: R.B. – Œuvres completes II, ed. By E. Marty, Seuil, pp. 479-484.

[4] Bates, J. 1992. Virtual Reality, art, and entertainment, in: Presence Presence: Teleoperators and Virtual Environments 1, pp. 133-137.

[5] Biocca, F. 2001, Inserting the Presence of Mind into a Philosophy of Presence: A Response to Sheridan and Mantovani and Riva, Presence: Teleoperators and Virtual

Environments 10, pp. 546-557.

[6] Biocca, F. and Jin Kim Yun Choi 2001. Visual Touch In Virtual Environments: An Exploratory Study of Presence, Multimodal Interfaces, and Cross-Modal Sensory Illusions, in Presence: Teleoperators and Virtual Environments 10, pp. 247-265.

[7] Böhme, Hartmut. Zur Thelogie der Telepräsenz, in F. Hager: KörperDenken. Aufgaben der Historischen Anthropologie, Dietrich Reimer, pp. 237-248. 1996.

[8] Bryant, R. 2001. What Kind of Space is Cyberspace? Minerva 5, 138-55.

[9] Clark, A. Being There. Putting Brain, Body, and World Together Again, MIT Press 1997.

[10] Ditton, Th. and Lombard, M. 1997. At the Heart of it All. The Concept of Presence", JCMC 3 [on-line serial]. http://www.ascusc.org/jcmc.

[11] Eco, U. Kant and the Platypus, New York, NY: Harcourt Brace. 2000 [orig.: Kant e l'ornitorinco, Garzanti. 1997].

[12] Fischer-Lichte, E. Ästhetik des Performativen, Suhrkamp. 2004.

[13] Floridi, L. 2005. The Philosophy of Presence: From Epistemic Failure to Successful Observation, Presence: : Teleoperators and Virtual Environments 14, pp. 656-667.

[14] Friedmann, F./ Regenbrecht, H./ Schubert, Th. 1999. Decomposing the Sense of Presence: Factor Analytic Insights. Paper presented at the 2nd International Workshop on Presence, University of Essex, 6th and 7th April, 1999.

[15] Gell, A.: Art and Agency. Anthropological theory, Clarendon Press, 1998.

[16] Gibson, J.J. The Ecological Approach to Visual Perception, Erlbaum Associates, 1979.

[17] Goffman, E. Frame Analysis. An Essay on the Organization of Experience. Northeastern UP 1986.

[18] Grau, O. Virtual Art: From Illusion to Immersion, MIT Press. 2003.

[19] Gumbrecht, H.-U. Production of Presence. What Meaning Cannot Convey, Stanford U.P. 2003.

[20] Heidegger, M. Sein und Zeit, Niemeyer, 1993.

[21] Heidegger, M. Wegmarken. Vittorio Klostermann. 1967.

[22] Heim, M. The Metaphysics of Virtual Reality, Oxford U.P., 1993

[23] Heyman, K. 2006. The Map in the Brain – Grid Cells May Help Us Navigate, Science 312 (2006), 680-681.

[24] Husserl. E.: Cartesianische Meditationen, ed. by Elisabeth Ströker, Hamburg 1987.

[25] Ijsselstein, W. Elements of a Multi-Level Theory of Presence: Phenomenology, Mental Processing and Neural Correlates, Proceedings of PRESENCE 2002, pp. 245-259. 2002.

[26] James. W. Principles of Psychology II, New York: Dover Publications 1950.

[27] Jenison, R.L. and Zahorik, P. Presence as Being-in-the-World, Presence: Teleoperators and Virtual Environments 7, pp. 78-89. 1998.

[28] Kablitz, A. 2003 Die Kunst des Möglichen. Prolegomena zu einer Theorie der Fiktion. in: Poetica 35 pp. 251-273.

[29] Klimmt, C. and Vorderer, P. Media Psychology "Is Not Yet There": Introducing Theories on Media Entertainment to the Presence Debate, Presence: Teleoperators and Virtual Environments 12, pp. 346-359. 2003.

[30] Mantovani, G. and Riva, G.: "Real" Presence: How Different Ontologies Generate Different Criteria for Presence, Telepresence, and Virtual Presence, Presence: Teleoperators and Virtual Environments 8, pp. 538-548. 1999.

[31] Merlau Ponty, M. Phénoménologie de la perception, Gallimard, 1945 [engl. Transl. London: Routledge and Kegan Paul 1962].

[32] Moran, D.W. / Reina, A./ Schwartz, A.B. „Differential Representation of Perception and Action in the Frontal Cortex", Science 303 (January 2004), 380-383.

[33] Murray, P and Wilson: Music and the Muses – The Culture of Mousike in the Classical Athenian City, Oxford 2004.

[34] Nancy, J.L. 1993. The birth to presence, Stanford U.P.

[35] Otto, W.F. Die Musen und der göttliche Ursprung des Singens und Sagens, Diederichs 1956.

[36] Palmer. M.T. Interpersonal Communication and Virtual Reality. Mediating Interpersonal Relationships, in: Frank Biocca (ed.): Communication in the Age of Virtual Reality. Lawrence Erlbaum 1995, pp. 277-299.

[37] Rheingold, H. The Virtual Community. Homesteading on the Electronic Frontier, MA: Addison-Wesley, 1993.

[38] Rizzolati, G./ Fadiga, L./ Fogassi, L. and Gallese, V. 1997. The Space Around Us, Science 277, pp. 190-191.

[39] Schutz, A. 1973. "On Multiple Realities," in: Collected Papers of Alfred Schutz, vol. 1, edited by M. Natanson. Martinus Nijhoff.

[40] Seel, M. 2000. Ästhetik des Erscheinens, Suhrkamp.

[41] Sheridan, T.B., Musings on Telepresence and Virtual Presence, Presence: Teleoperators and Virtual Environments 1, pp. 120-126. 1992.

[42] Singer, M.J. and Witmer, B.G., Measuring Presence in Virtual Environments: A Presence Questionnaire, Presence: Teleoperators and Virtual Environments 7, pp. 225-240. 1998.

[43] Slater, M. and Usoh, M. 1994. Depth of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments, 6, 603-616.

[44] Steiner, G. 1989. Real Presences. Faber and Faber.

[45] Straus, E. 1963. The Primary World of Senses. A Vindication of Sensory Experience, Collier-Macmillian [german orig. Vom Sinn der Sinne, 1935/1956].

[46] Turner, P and Turner, S. 2006. Place, Sense of Place and Presence, in Presence: Teleoperators and Virtual Environments 15, pp. 204-217.

[47] Waterworth, J.A. and Waterworth, E.L. 2003: Being and Time: Judged Presence and Duration as a Function of Media Form, in Presence: Teleoperators and Virtual Environments 12, pp. 495-511.

Presence in Technologically Mediated Environments: A Research Framework

Chang Nam and Steve Johnson
Department of Industrial Engineering, Department of Computer Science & Computer Engineering
University of Arkansas, Fayetteville, AR 72701, U.S.A
{cnam@uark.edu, sajohns@uark.edu}

Abstract

The paper outlines a research framework that can serve as a test-bed to presence research. Few system architectures are still available for presence researchers to use as a test-bed in order to investigate research issues involved in perceived presence in technologically mediated environments. The framework allows researchers to manipulate and investigate system configurations and parameters (e.g., dynamically changing agent roles, autonomy level, etc.), as well as psychological and social factors, which may hinder or facilitate the user's perceived presence in multi-agent contexts. The framework is an adaptable multi-agent system that interfaced intelligent, autonomy-based agents to a human commander within a real-time simulation environment. The research framework developed in this paper would be able to provide a basis for more flexible and systematic study on presence in technologically mediated environments.

1. Introduction

Advances in immersive, interactive technology, combined with its increasing availability and quality, have resulted in a practical concern with the manner in which people interact with technologically mediated environments such as virtual environments, haptic systems, and 3D online games. Terms like presence [1, 2], telepresence [3], and virtual presence [4] are used interchangeably to describe the extent to which people perceive that they are actually present in the artificially created virtual environment. Many attempts have been made to define such experience - *the feeling of being there that is created by media technologies* - and identify its determinants. Researchers have often developed an experimental system for their presence studies. More recent studies, on the other hand, utilize commercial off-the-shelf systems (such as 3D online games) as a test-bed.

These systems tend to limit the research scope of a presence study, however, because they do not allow the researcher to easily modify system configuration in a way that is cohesive to the research goal(s). That is, there are few system architectures available for presence (and even copresence) researchers to use as a test-bed in order to investigate research issues involved in perceived presence in technologically mediated environments. In addition, little attention has been paid to the presence of the user who is interacting with intelligent, autonomy-based agents - as in adaptable multi-agent systems.

This paper aims to establish a research framework that allows researchers to manipulate system configurations (e.g., task type, agent characteristics, etc.) and parameters (e.g., dynamically changing agent roles, autonomy level, etc.), as well as psychological (e.g., personality, trust, etc.) and social factors (e.g., collective efficacy, etc.), which may hinder or facilitate the user's perceived presence in multi-agent contexts. The framework is an adaptable multi-agent system that interfaces intelligent, autonomy-based agents to a human commander within a real-time simulation environment. The research framework developed in this paper would be able to provide a basis for more flexible and systematic study on presence in technologically mediated environments.

The paper begins with an outline of the UT2003 game environment, which is a major system component of the framework. A development strategy to implement intelligent agents, represented as Java Bots in the UT2003 environment, is then set out. Java Bots can dynamically change their roles according to the situations (e.g., as a team member, as an individual helper, or as a team helper). A specific mechanism that interfaces Java Bots to UT2003 follows, with the aim of developing agents that can dynamically change their autonomy levels (e.g., 0%, 50%, or 100%). Finally, the paper concludes with a description of how the framework proposed in the study can be used to investigate various factors affecting perceived presence.

2. The UT2003 Game Environment

The framework breaks down into three components:

1. **The Java Applet** – this Applet is used to add new agents/bots into a game session and to communicate with them.
2. **The Startup File** – this program is used to initialize the server and a game to certain specifications. It can also be used as a log by agents, which can be viewed by the human player(s).
3. **UT2003** game environment – also referred to as the Unreal Tournament 2003 Game.

Figure 1 shows the architecture of the framework developed in this study. UT2003 is a game that was co-developed by Epic Games, Digital Extremes, and Atari. UT2003 provides a simulated 3-D world in the form of pre-designed, customizable 3-D maps. Each map is customizable through the use of the UT2003 map editor *UnrealED* that allows a level designer to custom build a terrain based environment containing elements such as doors, elevators, and water. A map also contains AI routes, referred to as navigation points, which an Agent can traverse when it needs to travel from one location to another. A finished map is then simulated through the UT2003 physics engine, which is capable of adjusting the gravity of a simulated environment.

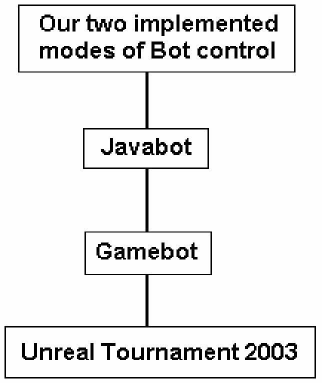


Figure 1 – The architecture of the framework

There are several ways in which one can play UT2003. The focus of our framework, however, is *Capture the Flag (CTF)*. In Capture the Flag, two teams (Red and Blue) attempt to outscore one another by capturing each others flag (see Figure 2). Each team has a *Home Base* in which their *Home Flag* resides. A team scores a point when they capture an enemy's flag, and return it to their home flag's spawn location.

The key to a team's success lies in maintaining the delicate balance between offense and defense. A team must be able to fortify its own strengths while exploiting the opposing team's weakness. In other words, a team's members must work together to ensure that their home flag remains safe while the opponent's flag is constantly under siege. In order for a team to work together, each member must have a line of clear communication to all other members. Through communication, a team is capable of maintaining an offensive and defensive balance via the division of offensive and defensive "roles" amongst its members. A typical successful team contains members that are willing to adjust offensively and defensively together according to the opponent's tactics.

Figure 2 – Example of *Capture The Flag* game

3. Intelligent Agents

In a UT2003 environment, Agents are referred to as Bots. Each Bot has been modified from it original UT2003 design to properly integrate into our framework. Each individually created Bot has its own intelligence, based on a uniquely designed team-oriented algorithm. In order to increase the probability of a team's success, each Bot is capable of serving as any of the following six unique offensive and defensive roles in our *Capture the Flag* game framework:

- ***CaptureEnemyFlag*** – A Bot will attempt to Capture and Enemy Flag to score a point for its team.
- ***DefendOurFlag*** – A Bot will travel to its Home Base Flag and wait until an approaching enemy is seen, at which point it will fight that enemy.
- ***DefendOurBase*** – A Bot will travel to the outskirts of its home base and wait until an approaching enemy is seen, at which point it will fight that enemy.
- ***SearchAndDestroy*** – A Bot will roam the map, increasing its Inventory and Health until an enemy is seen, at which point it will fight that enemy
- ***CoverMe*** – When a Bot issues this order, all available Bots will follow and protect the issuer until it dies.
- ***HoldThisPosition*** – When a Bot issues this order, all available Bots will travel to the position issued and wait until an approaching enemy is seen, at which point it will fight that enemy.

Crucial to the Bot AI framework is a Bot's ability to change roles. The core of a Bot's AI, the *Brain*, assesses each team member's role and surrounding environments, and makes a decision on what role would best serve the team. Additionally, Bot's are able to communicate clearly with one another – affecting Bot role assessment.

While in a specified role, a Bot will perform tasks related to that role. Under certain environmental situations, it is necessary for a Bot to reassess what actions it is performing in relation to its team members. Figure 3 demonstrates a code sample utilized in a Bot's brain.

```
if(DOF == 0 && CEF == 0)
{
    r = frand();
    if(r < 0.5) gotostate('DefendOurFlag','Begin');
    else gotostate('CaptureEnemyFlag','Begin');
}
else if(CEF > DOF || DOF == 0) gotostate('DefendOurFlag','Begin');
else if(DOF > CEF || CEF == 0) gotostate('CaptureEnemyFlag','Begin');
else if(DOB > DOF || DOB > CEF || SND > DOF || SND > CEF)
{
    r = frand();
    if(r < 0.5) gotostate('DefendOurFlag','Begin');
    else gotostate('CaptureEnemyFlag','Begin');
}
else
{
    r = frand();
    if(r < 0.5) gotostate('DefendOurBase','Begin');
    else gotostate('SearchNDestroy','Start');
}
```

Figure 3. Code sample from a Bot's Brain

The extent to which a Bot can independently reason is dependent upon a Bot's autonomy level. Before discussing Bot autonomy, let us first consider the manner in which Bots are interfaced to UT2003 and to a human user.

3.1. Interfacing to UT2003

Interfacing Bots to UT2003 is performed using the contributions made with the *GameBots* and *JavaBots* projects, as shown in Figure 4. *Gamebots* is a system which allows the UT2003 Bots in game are controlled using network sockets connected to clients. *JavaBots* provides a selection of Java packages that are designed for handling low level communication to the *GameBots* server. Essentially, the combination of *JavaBots* and *GameBots* allows any Java-based software to directly interact with UT2003. Additionally, *JavaBots* contains an API that allows a human user to connect and interact with UT2003 and its Bots therein.

The *JavaBots* API, which our development team has modified, provides two modes with which a human user can utilize for Bot control. In one mode, a human user has the ability to "command" a Bot to become a specific role (described in Section 3) by clicking an appropriate order related button. In the other mode, a user can control specific tasks (RunTo, Jump, Shoot, etc.) for a specific Bot. Selecting a task, providing appropriate data into an input box, and clicking the Send button will cause a Bot to immediately perform the selection. A Bot communicates with the human user in two ways. First, a Bot textually communicates with a human user through the UT2003 log window. Second, a Bot vocally communicates with a human user through the UT2003 game environment window. The purpose of a Bot communicating with a human user is to provide feedback on what a Bot is thinking or doing. Figure 4 shows the Java Client Window where the human user and Bots communicate each other.

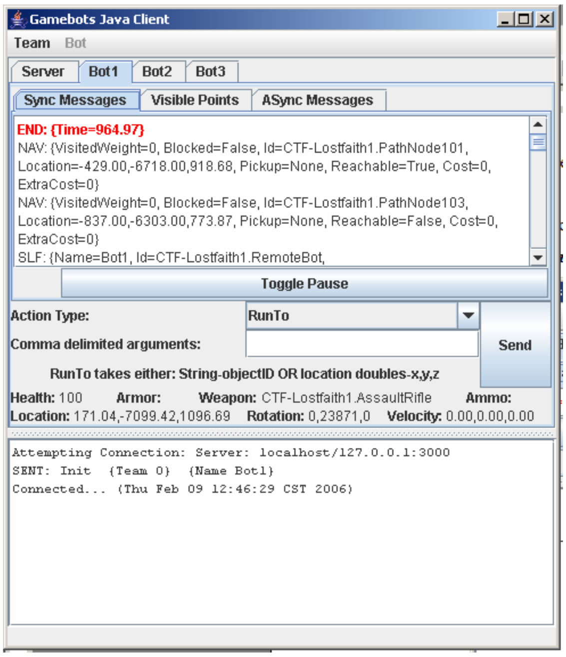


Figure 4. Java Client Communication Window

Finally, a *Capture the Flag* game is graphically animated in the UT2003 game environment window. By clicking the UT2003 window, a human user is able to cycle through each generated Bot's view. This allows a human user to visually see what a Bot is doing at a particular time.

3.2. Bot Autonomy

There are three levels of Bot autonomy, approximately labeled 0, 50, and 100 percent.

- **0%** - A Bot has little to no independent reasoning and must constantly rely on a human commander what role to take (*CaptureEnemyFlag* is the default).

- **50%** - A Bot has moderate independent reasoning, but must occasionally ask a human commander what role to take.

- **100%** - A Bot is fully independent, capable of reasoning what role to take on its own.

Regardless of a created Bot's autonomy level, a human commander can directly command a Bot by issuing a command through the JavaBots API interface.

4. Presence in technologically mediated environments

The framework developed in the study can be used to investigate individual's perceived presence and its affecting factors in technologically mediated environment, especially in the context of adaptable multi-agent system environments. The four interposed factors that may hinder or facilitate presence will be described: control, sensory, distraction, and realism.

4.1. Control Factors

The greater the level of control a user has, regarding their actions in an environment, the higher the level of presence [2]. The Control factors are determined by the following aspects:

- **Degree of control:** It is generally believed that the control over environment increases the presence level of a human user.

- **Immediacy of control:** The delay in the response of virtual characters decreases the presence level [5]. For example, a virtual character (or avatar) should be highly responsive to reciprocate a high level of presence to the human user.

- **Anticipation of control:** Human users will experience a greater sense of presence if they are able to anticipate the next action of a character (or avatar), regardless of whether it is under their personal control [5].

- **Mode of control:** If the user interacts with a system via effective modes, then the presence a user experiences increases. Additionally, the more modes of control available to a user, the higher level of presence.

- **Physical environmental control:** The degree to which an environment's physical objects can be manipulated by a user affects the degree of presence experienced by a user.

In our framework, the Interface and Bot autonomy directly affect Control. Our Interface provides a *Degree of Control* that provides a user with the necessary control to directly influence a Bot's role and many specific movements, yet a user cannot directly control all aspects of a Bot's movement. The *Immediacy of Control* is a mere fraction of a second. Any user-issued order is immediately performed by a Bot. Because a user can view a Bot during a *Capture the Flag* game, there is a higher degree of presence experienced through the *Anticipation of Control*. There are two effective *Modes of Control* (described in Section 3.1) a user can utilize for Bot Control. UT2003 maps provide *Physical Environmental Control* for pickup items, flags, water, and other Bots within a *Capture the Flag* game.

A Bot's autonomy level can have a direct affect on presence levels. The more autonomous a Bot is, the less it needs to be controlled. However, if a user is commanding a group of Bots, then a user's focus shifts from controlling one Bot, to controlling a team. Thus, the level of presence can vary based on a user's self-perceived role. Regardless, a user does have the ability to directly influence any single Bot.

4.2. Sensory Factors

Sensory information is the information humans receive through their sensors: eyes, ears, touch, etc. A greater degree of sensory information will lead to a higher level of presence [2]. Sensory is defined by the following aspects:

- **Sensory modality:** Different sensory modalities influence the degree of presence experienced by the human user. Visual sensory provides the greatest degree of presence, while other sensory channels provide lesser degrees.

- **Environmental richness:** The amount of information transmitted by system is proportionate to the level of presence experienced by a human user [1]. A vast array of environmental stimulations generates a greater sense of presence.

- **Multimodal sensory:** Senses stimulated in tandem by a system increased the presence experienced by a user [5]. Nam & Chung (2006) also showed that thermal feedback can facilitate the user's perceived presence in virtual environments [6].

- **Consistency of multimodal information:** The information received through all modalities should describe the same objective world [6]. If information from one modality gives a message that differs from that experienced through a different modality, a user's presence level will decrease [5].

- **Degree of movement perception:** If the user perceives self-movement through the system, to the extent that objects appear to move relative to the character, then the presence experienced by a user increases.

- **Active search:** The degree of environmental sensory control given to the user increases a user's presence level [1]. For instance, the extent to which users can adjust their viewpoint to change what they see, increases a user's presence.

In our framework, there exist visual, auditory, and haptic s*ensory modalities*. A user receives visual feedback through all three interface windows (*JavaBots* API, UT2003 Log, UT2003 game), which can also be manipulated into different levels (e.g., desktop virtual environment Vs. Head-Mounted Display). A wide array of auditory feedback occurs through listening devices used by the user. Auditory feedback includes all auditory information relative to a viewed Bot within the UT2003 window (e.g. item pickup sound, gunfire, Bot footsteps, etc.), verbal communication relayed among Bots (e.g. "Cover Me", "I've got your back", etc.) as well as UT2003 in-game messages (e.g. "Flag has been captured", "Flag returned", "Flag dropped"). Using a joystick (e.g., Logitech force 3D pro joystick and a precision game controller) provides haptic feedback to a user while playing as well. These three modalities affect the *Environmental Richness* of UT2003. These modalities are used in tandem, increasing *Presence* through *multimodal sensory*. All information relayed through these modes describe the same objective world (mapped environment) in UT2003, upholding the *Consistency of Multimodal Information*. The *Degree of Movement Perception* is consistent through the use of *Third-Person Perspective* in relation to each individual Bot. A user can *Actively Search* through multiple onscreen characters by clicking the UT2003 game window.

4.3. Distraction Factors

Distraction factors are sensory impacts from the natural world, not the mediated environment [2]. Distraction Factors are defined by the following:

- **Isolation: S**ystems that isolate a user from their physical environment may increase the presence experienced by a user. For example, a head-mounted display that isolates users from the other disturbing factors may increase presence in the system in comparison to a standard two-dimensional, flat screen display.

- **Selective attention:** The observer's willingness or ability to focus on the stimuli as well as to ignore distractions that are external to the system directly affects the amount of presence experienced in that environment.

- **Interface awareness:** Unnatural interface devices interfere with the direct and effortless interpretation of (and interaction with) an environment [5]. Hence, an intuitive, natural interface will increase a user's presence.

4.4. Realism Factors

Realism factors represent the continuity and connectedness [in comparison to reality] of the user experience [2].

- **Graphical realism:** Graphical realism refers to the connectedness and continuity of the stimuli which is being experienced. *Presence* increases as a system's graphical realism more closely resembles photorealism. Examples of rendered graphical elements include: graphical content, texture, resolution, light sources, field of view (FOV), and dimensionality [7].

- **Consistency of information with the objective world:** The more consistent a system conveys information resembling real-world experiences, the more a user will experience presence.

- **Meaningfulness of experience:** Meaningfulness pertains to user motivation, task saliency, and previous experience. A more meaningful situation will increase user presence.

- **Separation anxiety/disorientation:** System users may experience disorientation or anxiety when returning from the system to the real world. The amount user disorientation is proportionate to the presence a user experiences in a system.

UT2003 directly affects Realism. The UT2003 game engine is solely responsible for generating graphical information in the UT2003 game window. Thus, UT2003 determines the *Graphical Realism* of our system. The UT2003's physics engine is responsible for the *Consistency of Information with the Objective World. Meaningfulness of Experience* and *Separation Anxiety/Disorientation* are determined by the *Presence* level of the individual user.

Conclusions

This paper described a research framework that can serve as a test-bed to presence research, while giving a basis for more flexible and systematic study on presence in technologically mediated environments. Therefore, the framework will allow presence (and even copresence) researchers to manipulate and investigate system configurations and parameters (e.g., dynamically changing agent roles, autonomy level, etc.), as well as psychological (e.g., personality, trust, etc.) and social factors (e.g., collective efficacy, etc.), which may hinder or facilitate the user's perceived presence in multi-agent contexts.

As it is still undergoing constant development, future work can be done in two directions. One is the implementation of more sophisticated development strategies for intelligent agents to be able to change their roles (e.g., as a team member, as an individual helper, or as a team helper) and autonomy levels (e.g., 0%, 50%, or 100%) according to the situations agents made sense. Another future work includes a series of empirical experiments to investigate how these technical factors and other variables (e.g., psychological and social variables, or interface types) affect the perceived presence of a human user(s) interacting with technologically mediated environments such as adaptable multi-agent systems.

Acknowledgements

This research was in part supported by the Air Force Office of Scientific Research (AFOSR) under Grant No. FA9550-05-1-0182. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily

reflect the views of the Air Force Office of Scientific Research (AFOSR).

References

[1] T.B. Sheridan. Musings on telepresence and virtual presence. *Presence*. 1992, 1(1), 120-126.

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

[3] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*. 1992, 42(4), 73-93.

[4] D.W. Schloerb. A quantitative measure of telepresence. *Presence.* 1995, 4(1), 64-80.

[5] R. Held, N.I. Durlach. Telepresence. *Presence: Teleoperators and Virtual Environments*. 1992, 1, 109-112.

[6] C.S. Nam, D. Chung. The influences of haptic thermal feedback on object identification and perceived presence. In *Proceedings of the 16th World Congress of the International Ergonomics Association (IEA)*. 2006. Maastricht, Netherlands: IEA Press (to appear).

[7] W.A. IJsselsteijn, H. de Ridder, J. Freeman, S.E. Avons. Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. *Presence*, 2000, 10, 298-311.

Defining Presence

Matthew Lombard and Matthew T. Jones
Temple University
{lombard@temple.edu, matjones@temple.edu}

Abstract

This poster presents a framework to examine divergent and overlapping definitions of presence and promote a standardized terminology for discussing presence phenomena; conference attendees are invited to assist in refining the framework by placing their favored definition(s) within it. The benefits and dangers of the endeavor are discussed, followed by an overview of the framework, and recommendations for its use.

1. Introduction

In the last half century, and especially during and since the 1990s, many scholars have advanced a wide variety of unidimensional and multidimensional conceptualizations, and corresponding terminology, for presence. While individually useful, many of the definitions overlap and contradict one another. And while it's a sign of the growing sophistication in presence scholarship, the identification of many new dimensions of presence has led to a glut of composite terms (e.g., spatial, social, mediated, virtual, immersive, perceived, objective, subjective, environmental, inverse, backward, forward, physical, and corporeal presence). As Waterworth and Waterworth [1] note, "Presence is still a vague concept; researchers in the area agree that there is something important conveyed by the term, but differ widely on exactly what that something is" (Conclusions).

In this poster we present a framework for untangling the many conceptualizations of presence. The diversity of definitions is the result of necessary conceptual 'brainstorming,' but if scholars are to constructively collaborate and ultimately better understand presence, we need a common framework and terminology.

2. Benefits and dangers

To build knowledge, researchers and theorists must have a common understanding of the meanings of the words they use. White, Maltais, and Nebert [2] note that "It is essential to the process of communication that all individuals and groups concerned either use the same term for a particular object or concept, or at least have the ability to translate between different terms," and Heilbron [3], echoing Francis Bacon [4], observes that "Among the obstacles to the steady advance of science are the words invented to denote its conquests" (p. 585). Adopting a common framework for definitions and terminology of presence will allow us to communicate and collaborate more effectively, compare theoretical propositions and empirical results within and across disciplines, and ultimately build knowledge in this area. The availability of common and generally accepted definitions means that scholars don't have to continually construct new definitions that are similar to those already in use. Although they don't insure more consistent and comparable measurements of presence, standardized definitions are a prerequisite for standardized measurements. And such a framework would eventually allow us to more accurately characterize acquired knowledge about presence phenomena via meta-analysis.

Despite the need for such a framework, there are reasons to be cautious. An inflexible, prescribed set of definitions and labels could constrain creativity and limit the development of innovative approaches and therefore academic progress. What is needed is a categorization of the important definitional work that has been done in a format that won't restrict, and will even encourage, the evolution of that work in the future.

3. A framework for presence definitions

In the poster we present a framework that organizes most scholarly definitions of presence and variants of presence in the literatures of diverse disciplines. A more detailed version can be found online [5]. The framework is designed to characterize and organize existing definitions and guide the evolution of current, and the development of new, conceptualizations.

The left-most column contains questions that organize the definitions based on their fundamental characteristics. The definitions at the top of the figure are the most general or broad, and those at the bottom are the most specific or narrow. The organizing questions are discussed below.

2.1. Is technology involved in the phenomenon?

The first and most basic distinction among definitions of presence concerns the issue of technology. Some definitions focus on objective properties of communication that explicitly exclude technology. Other definitions explicitly involve the use of technology, "a machine, device, or other application of human industrial arts including television, radio, film, the telephone, computers, virtual reality, and simulation rides; traditional print media such as newspapers, books, and magazines; and traditional arts such as painting and sculpture" [6]. And some definitions can apply in either context, when technology is involved or not.

2.2. What is the phenomenon a property of?

A second key distinction concerns whether the phenomenon being defined is an objective property of a mode of communication, person, object or entity, or a subjective property of a person.

2.3. What is the source of the stimuli?

For those definitions of presence that involve a subjective property of an individual, the source of the experience or perception can be external – i.e., outside the body, in the 'real' world, or it can be internal – i.e., inside the body (specifically the brain). External sources are basically all impingements on our senses from the physical world around us, while internal sources are controlled or automatic mental processes that result in remembering a vivid experience, dreaming, daydreaming, and the like.

Some definitions explicitly or implicitly apply only to our experiences of the external world (via technology or not), while other definitions are more inclusive, with either external or internal stimuli generating presence.

2.4. How is technology perceived?

The fourth distinction in presence definitions concerns the perception of technology in an experience. There are four logical possibilities: When technology is not involved in an experience, as in "face to face, body to body" communication [7], the fact that technology plays no role in the experience can be either accurately or inaccurately perceived, and when technology is involved, as when a person uses virtual reality or other media, the role of technology in the experience can be either accurately or inaccurately perceived.

The two most common types of definitions describe the accurate perception that there is no technology involved, and the inaccurate perception that technology is not involved when in fact it is (e.g., "the perceptual illusion of nonmediation" [8]. In the first of these scenarios, a 'natural' or 'direct' or 'non-technology-based' experience is accurately perceived as such, and in the second, a person automatically or willfully overlooks the 'artificial' or 'indirect' or 'technology-based' nature of an experience created or modified by technology.

2.5. What aspect of the phenomenon is of interest?

The fifth and last distinction among presence definitions in the framework concerns the different aspects of the phenomenon. These definitions typically denote distinct but overlapping dimensions or types of presence, including spatial or environmental presence, social presence, psychological engagement, perceptual and social realism, cultural presence and para (or logically impossible) presence.

3. Recommendations

For presence scholarship to advance, those who study it need to all be "on the same page." Because there are so many different definitions, and because it's often not clear which definition scholars have in mind, "when people talk about *presence* they are often not talking about the same underlying concept at all" [9].

Rather than attempting to build consensus around a single, ideal definition of presence, we urge scholars to make very explicit the definition(s) that they are using in their work. A logical way to do this is to answer for readers and listeners the five key questions that organize the framework presented here (i.e., locate the definition in the framework of definitions). We invite conference attendees and online visitors [5] to help refine the framework by doing this. Our collective work will also advance more quickly if we use existing terms and definitions whenever possible. We'll update the online framework [5] as our collective understanding of presence evolves.

References

[1] Waterworth, J. A. & Waterworth, E. L. (2003, July). The core of presence: Presence as perceptual illusion. *Presence-Connect, 3*. Retrieved on February 15, 2005 from http://presence-connect.com

[2] White, J. A., Maltais, L. J., & Nebert, D. W. (1998). An increasingly urgent need for standardized gene nomenclature. *Nature Genetics*. Retrieved on January 2, 2006 from http://www.nature.com/ng/web_specials/nomen/nomen_article.html via Google Scholar.

[3] Heilbron, J. L. (2002). Coming to terms: Caloric, cathode, curium and quark — coinage from the mint of science. *Nature, 415*, p. 585.

[4] Bacon, F. (1889/1620). *Novum Organum.* (T. Fowler, Trans.) Oxford, UK: The Clarendon Press.

[5] Lombard, M. & Jones, M. T. (2006). Defining presence: A framework. Temple University. Available: http://astro.temple.edu/~lombard/presence_definitions.html

[6] International Society for Presence Research (2000). An explication of presence. Retrieved on January 13, 2006 from http://ispr.info

[7] Zhao, S. Y. (2003). Toward a taxonomy of copresence. *Presence: Teleoperators and Virtual Environments, 12*(5), 445-455.

[8] Lombard, M., & Ditton, T.B. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2). Retrieved on February 15, 2005 from http://jcmc.indiana.edu/vol3/issue2/lombard.html

[9] Slater, M. (2003, January). A note on presence terminology. *Presence-Connect, 3*. Retrieved on February 15, 2005 from http://presence-connect.com

The Effect of Static Anthropomorphic Images on Emotion Perceptions in Mobile-Phone Communication

Sin-Hwa Kang, James H. Watt and Katherine Isbister
Rensselaer Polytechnic Institute
{kangs@rpi.edu, wattj@rpi.edu, isbisk@rpi.edu}

Abstract

We describe how static anthropomorphic images of avatars affected users' emotion perceptions of interactants and medium when they engaged in a conversation on a mobile telephone. We utilized measurements of social presence, specifically interactant satisfaction and social richness of medium. The items of these two measurements were modified to include new items of emotional credibility. The results showed no statistically significant effect of static anthropomorphic images on the perceptions of interactant satisfaction and social richness of medium. However, mobile telephone users who saw no images reported significantly higher interactant satisfaction.

1. Introduction

People are likely to use a mobile phone to make social ties stronger, even between strangers. The technological development of mobile phone services means that the small screen of a mobile phone is now called upon to display various computer-generated graphics such as game characters, social agents for online shopping or education, etc. Earlier studies [1, 3, 4, 5, 8] have reported the impact of visual images of both human and nonhuman beings in virtual environments, but on computer systems having large visual displays, as compared to the very small mobile phone screen. Further, a different type of measurement has not been used to evaluate users' interactant satisfaction in the emotionally communicative situation typical of mobile phone use. Both these limitations imply the need for research into the impact of anthropomorphic images in communications whose utility is strongly or primarily emotional, via technology which permits only very limited visual displays.

2. Research Problems and Questions

Walther's "hyperpersonal communication" [9] theory supports the idea that people feel more affection and emotion with their interactants when they have fewer cues. This leads to the prediction that "lean" media may actually provide a stronger perception of emotional

social presence than "rich" media in some instances. Introducing the impact of anthropomorphic images into this speculation, it may be that people's responses to a highly anthropomorphic image in a rich medium might be similar to their responses in a situation in which no image is displayed in a lean medium like the limited display of a mobile phone [2]. Moreover, it has been found that simply adding human faces to interfaces is not necessarily preferable in all cases [1]. Therefore, we aimed at testing the following hypotheses.

H1: People will report higher satisfaction with their interaction partners when they are represented by no images, rather than by high or low anthropomorphic static images in emotionally engaged and synchronous mobile-phone communication.

H2: People will report higher social richness of medium when their interactants are represented by no images, rather than by high or low anthropomorphic static images in emotionally engaged and synchronous mobile-phone communication.

3. Experiment Design

The basic research design is a 3x3 factorial between-subjects experiment involving two factors: three levels of static anthropomorphic images: high anthropomorphic, low anthropomorphic, and no image; crossed with three gender combinations of dyadic partners: male-male, male-female, and female-female.

3.1. Participants

Ninety-eight university students (45 males and 53 females) were recruited at two northeastern universities. Each participant was paired with someone whom they were unlikely to know beforehand. The participants were randomly assigned to one of nine experimental conditions in the 3 x 3 design.

3.2. Procedure

The interaction took place in two separate rooms to avoid any initial face-to-face interaction. Participants were asked to complete a pre-test questionnaire measuring demographics and mobile phone use. They were then given a hypothetical conversational scenario which asked them to take on the role of a student seeking to find out if the other person was a suitable match with whom to share an apartment. Participants were given a mobile phone and a hands-free headset which allowed participants to look at the static anthropomorphic image on the screen during the conversation. The typical conversation lasted about 10 minutes. When the participants finished the conversation, they were asked to complete a post-test questionnaire.

3.3. Measurement Instruments

3.3.1. Response Variables

Interactant Satisfaction. A 15 item Likert-type scale was measured with a 7 point metric. Among 15 items, 6 items were adopted from the items of social attraction used in Nowak's study [3]. The other 9 items, called emotional credibility, were newly created to measure interactants' emotional perceptions of their interaction partners. The new items originate in the studies of Mayer and Salovey [6], which Smith applied to describe the results of the Emotional Intelligence test [7].

Social Richness of Medium. A 5 item Likert-type scale was measured with a 7 point metric. These 5 items were modified from the items of "subjective social richness of the medium" used in Nowak and Biocca's study [4] and worded to be applicable to a mobile-phone communication.

3.3.2. Control Variable In the analysis of this study, gender combinations of interactants were considered a measured control variable. Participant dyads were almost evenly distributed across male-male (15), male-female (15), and female-female (19) pairs.

4. Results

Reliability tests showed good internal consistency among the 15 items of Interactant Satisfaction when they were combined into a summed scale (Cronbach's alpha = .85), as well as the 5 items of Social Richness of Medium (Cronbach's alpha = .82). Overall ANOVA results did not reveal a statistically significant main effect for the degree of anthropomorphism on either Interactant Satisfaction [$F(2, 95)=2.4$, $p=.10$] or Social Richness of Medium [$F(2, 95)=.34$, $p=.71$]. There was no main effect for the gender mixture of the dyads on Interactant Satisfaction [$F(1, 96)=.80$, $p=.37$] or Social Richness of Medium [$F(1, 96)=.09$, $p=.77$], nor for the interaction of avatar anthropomorphism and dyad gender type on Interactant Satisfaction [$F(2, 92)=.45$, $p=.64$], or on Social Richness of Medium [$F(2, 92)=.18$, $p=.83$].

However, a comparison of the no-image condition to all conditions in which an avatar was present showed a significant difference in users' Interactant Satisfaction [$F(1, 96)=4.8$, $p=.03$], with users who saw avatars reporting lower Interactant Satisfaction ($M=4.84$, $SD=.70$ versus $M=5.15$, $SD=.57$ for those who saw no image). The eta square value for this difference was .05 which could be considered a medium effect size in Cohen's (1988) terms. This is partial support for Hypothesis 1. The same comparison of avatar versus no avatar conditions did not result in a significant difference in users' reports of the Social Richness of Medium [$F(1, 96)=.04$, $p=.84$]. There is thus no support for Hypothesis 2.

5. Conclusions and Discussion

In this study, participants reported feeling more engaged in their communication when they did not use any anthropomorphic images representing their partners on the small screen of a cell phone. In an open-ended question in the post-test questionnaire, there is an indication that participants' expectations of the anthropomorphic image quality and behavior may have affected their evaluations of the interaction partner and the medium. This finding may mean that dynamic images might provide greater emotional connection between conversational partners, reversing the negative effect of static images found in this study. That possibility remains a topic for future research. Furthermore, there was no evidence of gender effects on Interactant Satisfaction in this study. Conventional wisdom would predict these effects, particularly in mixed-gender dyads. We propose to further explore this subject in further studies.

Acknowledgements

We gratefully acknowledge the VECG group at UCL, Dr. Teresa Harrison at SUNY at Albany, Dr. Junho Choi at RPI and their colleagues for allowing us to design and conduct our experiment using their resources.

References

[1] Koda, T. Agents with faces: The effects of personification. In *Proceedings HCI 1996*. 1996.

[2] Mosmondor, M., Kosutic, T., and Pandzic, I. LiveMail: Personalized avatars for mobile entertainment. In *Proceedings 3rd International Conference on Mobile Systems, Applications, and Services*. 2005.

[3] Nowak, K. L. The Influence of Anthropomorphism and Agency on Social Judgment in Virtual Environments. *Journal of Computer Mediated Communication, 9*(2), 2004.

[4] Nowak, K. and Biocca, F. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, Vol. 12, No. 5, 2003.

[5] Nowak, K. and Rauh, C. The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication, 11*(1), 2005.

[6] Salovey, P., Brackett, M., and Mayer, J. *Emotional Intelligence: Key readings on the Mayer and Salovey Model*. Dude Publishing. 2004.

[7] Smith, J. *Emotional Intelligence Report: EIQ16 Questionnaire*. (2004) http://www.myskillsprofile.com/eiq16_sample_report.pdf

[8] Takeuchi, A. and Naito, T. Situated facial displays: Towards social interaction. In *Proceedings CHI'95 Human Factors in Computing Systems*, Addison Wesley. 1995.

[9] Walther, J. B. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 23,* 3–43, 1996.

Presence and video games: The impact of image quality and skill level

Cheryl Campanella Bracken
Cleveland State University
{c.bracken@csuohio.edu}

Paul Skalski
University of Minnesota Duluth
{pskalski@d.umn.edu}

Abstract

This study investigates the impact of image quality and skill level on presence-related reactions to video games. Past research has demonstrated positive associations between image quality and presence and video game technology and presence. No study to date, however, has examined the presence effects of video games played in enhanced or high definition. This paper reports the results of two experimental studies. In the first study, a pilot investigation, 22 college students played a video game either in enhanced definition (higher image quality) or standard definition/NTSC (lower image quality). In the second study, 50 college students played a video game in either high definition (highest image quality) or standard definition/NTSC (lower image quality). Following exposure in both studies, participants completed several measures of presence and a measure of video game skill. Skill was included as a second independent variable in the study, and both image quality and skill were expected to relate positively to presence dimensions. The results of the study provide some support for both image quality and skill affecting presence, though not all results were in the direction expected.

1. Introduction

Video games have become one of the most popular forms of media in the United States and abroad. Global sales in the industry were projected to exceed $30 billion dollars by the end of 2002 [1], and half of Americans age six and older are currently estimated be playing [2]. The popularity of games has been fueled in part by advancements in gaming technology, a trend that has persisted through the earliest days of the medium [3]. Over time, games have evolved considerably in graphic richness and realism. The simplistic character representations in games like *Pac-Man*, for example, have now been replaced with the realistic human figures and environments in popular new titles like *Grand Theft Auto: San Andreas* and *Halo 2.* These and other advances in game technology have important consequences for how games are experienced. In particular, they are expected to contribute to the sense of *presence*, or "perceptual illusion of nonmediation" [4], felt by users. Presence has recently been identified as a potentially important variable in video game research that may affect use and a variety of outcomes of exposure, ranging from enjoyment to aggression [5-6]. Few studies, however, have examined the relationship between exposure to game technology and presence. Tamborini et al. [7] found that playing a game created a stronger sense of presence than observing a game, presumably due to the addition of interactivity. Though many technological features of video games are expected to contribute to the sensation of presence, one that has received no attention to date is image quality. Definition Television (HDTV), for example, sharply improves the quality of TV images and, with TV and movie clips, has been shown to relate positively to the experience of presence [8]. But what about video games, which add the crucial feature of interactivity to HDTV and other high-quality media images?

Based on this and evidence presented earlier in the paper, the following hypotheses are posited:

Hypothesis 1: Participants who play the enhanced definition or high definition version of a video game will experience a higher level of *presence* than those who play the NTSC version of the video game**.**
Hypothesis 2: Participants who score higher on a video game skill scale with report higher levels of presence than those who score lower on the video game skill scale.

2. Methodology

Two studies were conducted using the same procedures but with different games and levels of image quality. Study 1 had 22 participants, who played a video game in either lined doubled progressive scan/enhanced definition (480p lines, component video) or standard definition (NTSC, 480i lines, composite video). In study 2, 50 participants played a game in HDTV (1080i) or standard definition (NTSC, 480i lines). In both studies, the independent variables were image quality (enhanced/HDTV versus standard television) and player skill (high versus low).

3. Analysis and Results

A series of 2-way analyses of variance with the independent variables image quality (enhanced [ED] versus low image quality [NTSC]) and skill level (high versus low) were used to test the hypotheses and research questions.

Hypothesis 1, predicting that participants who played video games with higher image quality would experience a higher level of spatial presence (being there in the video game) than those who played in lower image quality, was not supported. A main effect approached significance for immersion, with those participants who played the video game in higher image quality reporting higher levels of immersion (*M* = 4.37, *SD* = 1.27) than those who played the game in NTSC (*M* = 3.48, *SD* = 1.51). Further support for improved image quality leading to higher levels of presence was found in study 2, in which participants who played the video game in the highest image quality reported higher levels of immersion (*M* = 4.80, *SD* = 1.44) than those who played the game in NTSC (*M* = 3.65, *SD* = 1.24).

Hypothesis 2, predicting that skill level will impact players' perceptions of presence, was not supported. In both studies, in fact, players who scored lower on the video game skill scale reported higher levels of presence. In study 1, lower skill players reported higher levels of immersion (*M* = 4.43, *SD* =1.13) than players with higher skill scores (*M* = 3.42, *SD* = 1.56). In study 2, the same pattern was found for immersion, with players with a lower game skill score (*M* = 4.79, *SD* = 1.47) reporting higher levels of immersion than players with higher game skill scores (*M*= 3.70, *SD* = 1.27).

4. Discussion

The results of these studies provide some evidence that image quality impacts both the level and types of presence dimensions experienced by video game players. The results strengthen the claim that image quality influences sensations of presence (4, 8]. Further, the results support previous work with video games and presence [5,7,9] In doing so, they add to the growing body of literature on video games, image quality, and presence and begin the process of synthesizing these important bodies of research.

The current study also introduced the use of measuring players' prior video gaming skill as a variable that may impact their playing experience, including level of presence. While some evidence suggests that this is a variable worth future consideration, the results were not in the direction expected. In both studies, skill negatively impacted the immersion dimension of presence, with low skill players reporting higher levels of immersion. It may be that unskilled players are not as familiar with today's game technology, leading them to be more enveloped by its realism. Or perhaps more likely, unskilled players may need to focus more to succeed and "survive" in fast-paced game environments like the ones used in this research, given their lack of playing ability. For skilled participants, the early game levels in this research may have been easy, and this could have negatively impacted immersion by not challenging their skills enough and instead calling their attention elsewhere. For unskilled players, the game may have been a considerable challenge, leading them to pay more attention and become more immersed in the experience. As the "flow" literature suggests, the potential for immersion is maximized when there is a match between challenge and skill [10]. Future work on video games should attempt to match player skill levels with game difficulty, to maximize the potential for presence to occur.

Conclusion

The current study provides a basis for the inclusion of a gaming skill scale in future studies. The results also provide evidence that image quality in video games has an effect on participants' sensations of some dimensions of presence. Though more work in this area is needed, this investigation serves as important first step toward better understanding the impact of video game image quality and skill on presence.

References

[1] Detroit Free Press. (2002, July 1). *It's a fun time for video game industry.* Retrieved October 30, 2004 from http://www.freep.com/entertainment/videogames/hedtip1_20020701.htm.

[2] Entertainment Software Association. (2002, May). *Top ten industry facts*. Retrieved October 30, 2004 from http://www.theesa.com/pressroom.html.

[3] Skalski, P. (2004) *The quest for presence in video game entertainment.* Presented as part of presence panel at the Central States Communication Association Annual Conference, Cleveland, OH.

[4] Lombard, M., & Ditton, T. B. (1997) At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2), Retrieved April 30, 2002 from http://www.ascusc.org/jcmc/vol3/issue2/lombard.html

[5] Tamborini, R. & Skalski, P. (2006). The role of presence. In Vorderer, P. & Bryant, J. (Eds.), *Playing computer games: Motives, responses, and consequences.* Mahwah, NJ: Lawrence Erlbaum Associates**.**

[6] Peng, W., Lee, K. M. (2004). *What do we know about computer and video games? A comprehensive review of literature*. Paper presented at the Annual Conference of the International Communication Association (ICA), New Orleans, LA, May, 2004.

[7] Tamborini, R., Eastin, M., Skalski, P., Lachlan, K., Fediuk, T. & Brady, R. (2004). Violent virtual video games. *Journal of Broadcasting and Electronic Media,* 48(3), 335-357.

[8] Bracken, C. C. (2005). Presence and image quality: The case of high-definition television. *Media Psychology*, *7,* 191-205.

[9] Schneider, E. F., Lang, A., Shin, M., Bradley, S. D. (2004). Death with a store: How story impacts emotional, motivational, and physiological responses to first-person shooter video games. *Human Communication Research, 30*, 361-375.

[10] Sherry, J. (2004). Media enjoyment and flow. *Communication Theory*, *14*(4), 328-347.

Applying telepresence robot to interpersonal communication: implications and applications

Tzung-Cheng Tsai, Yeh-Liang Hsu, Chang-Huei Wu
Department of Mechanical Engineering, Yuan Ze University
135, Yuan-Tung Rd., Chung-Li, Taiwan, ROC
{s919201@mail.yzu.edu.tw, mehsu@saturn.yzu.edu.tw, mechwu@saturn.yzu.edu.tw}

Abstract

This paper discusses applying telepresence robot to interpersonal communication. First the "user" and "people" connected by a telepresence robot in the interpersonal communication are defined. This paper then describes telepresence research literature from the framework of projection-immersion (from the user's perspective) and observer-dialogist (from the people's perspective). The design elements in telepresence systems in research literature are identified and classified.

The development of a telepresence robot for interpersonal communication with the elderly in the Gerontechnology Research Center, Yuan Ze University is presented. The design elements included in this telepresence robot are discussed. In particular, how to implement proper autonomous behaviors in the telepresence robot to increases the user's capability of projection to operate the telepresence robot and to increase the interactive capability of the people as a dialogist is emphasized. The concept of transferring the control authority of the telepresence robot from the user to the environment, and further to the people interacting with the telepresence robot is discussed.

Keywords--- **Telepresence, interpersonal communication, autonomous behavior.**

1. Introduction

"Telepresence" is an interesting field that includes virtual reality implementations with human-system interface, communication technologies and robotics. Telepresence provides a connection between a user and distant people or environment (real world or computer generated world), to perform social interactions (user-telepresence-people) or specific tasks (user-telepresence-environment). This paper focuses on the application of telepresence in communication and interaction between the "user" and "people." In such applications, "people" are not only "observers", but also participants or "dialogists" in the interpersonal communication.

There are two views in telepresence application in interpersonal communication: the user's view and the people's view. From the user's view, telepresence enables the user to project herself/himself to another place by controlling the telepresence robot or system. In the meantime, the user perceives immersion from the sensory feedback from the remote environment created by telepresence. As discussed above, "people" may have two roles in telepresence application in interpersonal communication: observers or dialogists. From the people's view, telepresence provides necessary elements to the teleoperator, and the telepresence robot, so that people recognize it as a representation of the user. Telepresence also enables dialogue between people and the user by sending audio, video, gestures, and other environmental information helpful for effective communication.

This paper tries to describe telepresence research literature from this framework of projection-immersion and observer-dialogist. The design elements in telepresence systems in research literature are identified and classified. Then an on-going project conducted by the Gerontechnology Research Center of Yuan Ze University (Taiwan) for developing a telepresence robot is described. This robot is for interpersonal communication with the elderly in a home environment. The development road map and current design elements are reviewed. Finally, the necessary design elements for such telepresence application in interpersonal communication are concluded, and future development of this telepresence robot is outlined.

Table 1. Design elements and related technology details for telepresence

Design elements	Related technology details
data transmission	RF, Internet, time-delay algorithm
modify environment	teleoperation, master-slave, simultaneous operation, robotic embodiment
supersensory	dexterity, maneuverability, scaled motion, zoom
anthropomorphic elements	facial expression, body motion, humanoid
stereoscopic	binocular, panoramic, autostereoscopic, image processing
stereophonic	HRTF, stereo audio
eye contact	camera and screen with specific placement
autonomous behaviors	navigation, mapping, localization

2. Design elements in telepresence systems in research literature

This paper surveys the application-oriented telepresence literature which describes the development of a telepresence system. The design elements emphasized in these studies are extracted and summarized in Table 1. These design elements are fitted into the framework of projection-immersion and observer-dialogist described in the previous section. A discussion of the eight design elements in Table 1 is given below.

3. Developing a telepresence robot for interpersonal communication with the elderly

The Gerontechnology Research Center of Yuan Ze University in Taiwan has been developing a telepresence robot for interpersonal communication with the elderly in a home environment. Figure 1 shows the development road map.

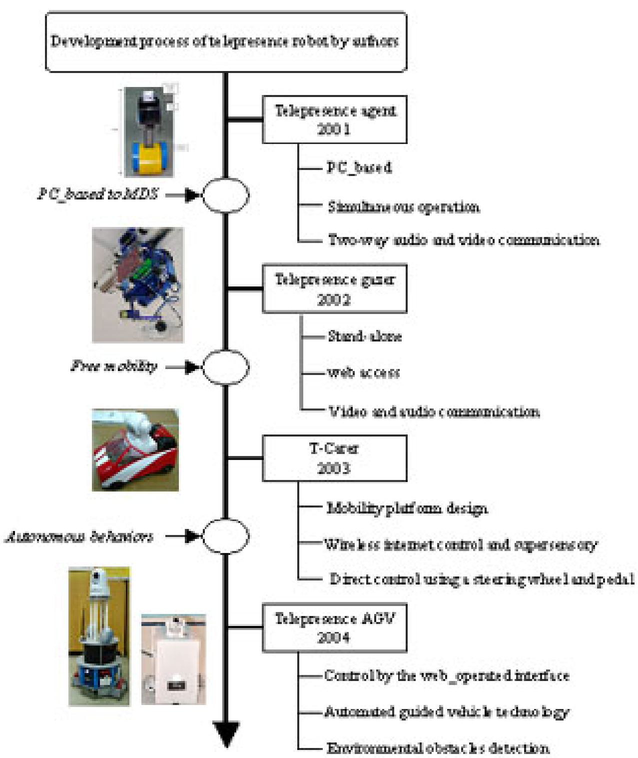


Figure 1 Development process of telepresence robot

Discussion and future work

The fifth generation telepresence robot, the "Telepresence InterComm", for interpersonal communication with the elderly is being developed. Telepresence InterComm aims to be a low-cost home robot. The total material cost of telepresence InterComm (not including the IP cam) is estimated to be about USD 400.

Table 2 shows the design elements already implemented or planned. As in the previous generations, the core of Telepresence InterComm for control and data transmission is the MDS. The design element "to modify the remote environment" is not essential for the communication telepresence robot. Only the fundamental mobility control is implemented. Supersensory is reflected in the zooming of IP cam and various sensors for environment detection installed on Telepresence InterComm. With the limited processing ability of the MDS, the user's face cannot be displayed (so eye contact is not possible on Telepresence InterComm). Instead, mechanical facial expressions are planned to include an anthropomorphic element. Sophisticated stereoscopic and stereophonic elements are not planned either.

Table 2. Design elements included

Design elements	Related technology details
data transmission	MDS
modify environment	mobility control
supersensory	zooming of IP cam, various sensors for environment detection
anthropomorphic elements	mechanical facial expression
stereoscopic	not included
stereophonic	not included
eye contact	not available
autonomous behaviors	major emphasis

The major emphasis of this research will be on how to implement proper autonomous behaviors in the telepresence robot to increases the user's capability of projection to operate the telepresence robot and to increase the interactive capability of the people as dialogists.

As discussed earlier, in principle, a telepresence robot is operated by a remote user. The user possesses the complete control authority. From the user's view, adding autonomous behavior to the telepresence robot shares the control authority with the environment. This autonomous behaviour will allow the AGV to identify an obstacle in the environment, so that the user can operate the telepresence robot safely and reliably in dynamic environments.

From the people's view, to increase the interactive capability with the people as dialogists, the control authority of the telepresence robot has to be shared with the people it is interacting with. For example, the telepresence robot will turn to the person/people who are speaking and turn to the direction the people are pointing so that the user and people share the same view. In other interpersonal communication scenarios, the telepresence robot may follow the people's guide to move around. The detailed development of Telepresence InterComm emphasizes sharing the control authority with the people in various interpersonal communication scenarios.

Projecting Presence: A Mimetic Approach to the Creation of Presence

Benjamin Unterman
Université du Québec à Montréal
{unterman.benjamin@courrier.uqam.ca}

Abstract

The nature of computer mediated presence and its relationship to our everyday reality are complex, and the study of its nature is relatively new. Theatrical presence, on the other hand has been the subject of analysis for thousands of years. Looking at these two manifestations of presence, the theatrical concept of mimesis is used to better understand their relationship. In recontextualizing presence as the final goal of creation rather than an internal process, this paper aims to expand our definitions of and perspectives on presence.

> "There is nothing more illusory in performance than the illusion of the unmediated. It can be a very powerful illusion in the theater, but it *is* theater and it is *theater*, the truth of illusion, which haunts all performance whether or not it occurs in the theater, where it is more than doubled over."
> – Herbert Blau ([1], 164-165)

1. Introduction

As Brenda Laurel pointed out in *Computers as Theatre* in 1992, theatrical metaphors are particularly rich when it comes to exploring our interaction with computers [2]. This paper extends her examination of the interface into the realm of presence, using theatrical metaphors and methods to help expand our definitions of and perspectives on computer-mediated presence.

A distinction is made repeatedly in this article is between theatrical presence and telepresence. Theatrical presence is, quite simply, the presence felt during a theatrical production. This can refer both to the audience's sensation of the actors' presence, and the actors' perceptions of the audience's presence. Both are current in our use of presence in theatre, and demonstrate the complexity of the mediated environment which is the theatre. Telepresence is used to describe the presence a user experiences while using a computer-mediated environments or other computer interfaces.

Though these forms of presence are distinct, they both have a mimetic relationship with everyday presence. This shared relationship form the basis of this analysis which strives to enrich our understanding of presence, regardless of its medium.

2. The Mimesis of Presence

The discussion of presence is one of the fundamental issues in theatre, treated as early as the 4th century BCE in Aristotle's *Poetics*. For Aristotle, the audience's relationship with the hero, their presence throughout the performance, is enabled by the use of *mimesis* on the part of the actor ([3], 1449-1453). Mimesis is usually translated in English as "representation" or "imitation," but is much more complex than that. The artist's intent to create a representation is also key, and the act of creation is in and of itself a mimetic act. As such, mimesis comprises a moral judgment that the artist makes on the world, both in his choices of how to represent the original object, but also in his selection of what object or traits of that object are to be represented. The artist is thus able to affect our perceptions of the original object by subtly altering or masking certain elements.

I want to suggest a shift in our usage of mimesis, applying it to the abstract model of presence, in order to better understand the nature of mediated presence, whether theatrical or digital. This assumes, of course, that we have an original model, an unmediated form of presence. While our everyday experiences can be considered as mediated, for the sake of this paper, our reference point is direct face-to-face communication between two or more people. The mimesis of presence can be applied to presence in the mediated environments of the theatre as well as computer-mediated environments.

In theatre, the separation of the audience from the stage creates a hierarchy. This separation is then reinforced by conventions, the darkening of the audience part of the room, for instance. The actors, who normally follow a script, present the action to the audience who, while they can subtly affect the performance, have no say in the outcome of any given play. The attention of the audience is focused instead on the concentrated narrative flow of the play they are watching.

Similarly, in virtual environments, the user is faced with an environment which, while obviously artificial, represents a potentially vast field of interactivity. As Lev Manovich suggests, such interactivity is essential for any computer-mediated experience ([4] p. 56-57). The focus then shifts somewhat away from the idea of a fixed narrative towards a model, as that suggested by Janet Murray, which puts greater emphasis on the agency of each participant ([5], p. 10).

Both theatre and computer-mediated environments depend on presence to capture our interest and drive our interactions with them. But both also make use of forms of presence that are similar to everyday presence, but differ in important ways. In short this purposely created presence, this mimesis of presence, is the backbone of both theatrical and computer-mediated creation.

3. Presence as the goal of creation

The movement towards a mimetic model of presence requires a reevaluation of the ways in which we treat presence, both for theatre and computer-mediated experience design. In both of these domains, presence has traditionally been understood as a method for achieving the ends of each particular performance or piece of software. Once we understand that presence itself is created, a new possibility emerges: the consideration of presence as the final creative product and guiding principle of the work.

Indeed, once we admit the possibility of presence as a purposefully created end product, it is easy to see how any art form can be reinterpreted in terms of the ways in which it creates presence. In 1936, Walter Benjamin analysed the subject in depth, most famously in his essay "The Work of Art in the Age of Mechanical Reproduction." In this essay, he associates presence (which he calls aura) not only with great works of art, but with the viewer's physical or conceptual proximity to the original work [6]. The goal of artwork is thus to create a sense of presence in its audience.

When we understand the process of using presence as being mimetic, there is a subtle but powerful shift. What is experienced when we are affected by a work of art is not presence itself, but an artificial and purposely created sense of what the presence might be. The role of the environment, then, is to facilitate the (re)production of presence in order to enrich the audience's experience of the work.

Which brings us to the consideration of telepresence. To use a mimetic model of presence, which is essential for any consideration of the artistic nature of any creation, telepresence must be treated as its own end, rather than as an enabler of other more tangible goals. In short, a recognition of presence in computer-mediated environments as being mimetically created enables the growth of a new artistic concept of presence which is separate and apart from the existing scientifically-oriented model.

4. Understanding the impact of mimetic presence

If we consider computer-mediated design as art, there are lessons to be learned from the wide ranging artistic experiments of the last 150 years. Our current tendency towards photorealistic graphics, for instance, is brought into question by this analysis. The effects of the adoption of a mimetic model of presence are more wide-reaching than this, but the question of graphical realism provides us with one of the more controversial consequences.

In theatre, there was a brief but influential movement towards realism at the end of the 19^{th} century. However, those plays which did not make use of stylized language and a wide range of symbolic or poetic meaning did not last. Realism was equally problematic in other art forms, such as painting (where incongruous elements are often presented together in a realistic fashion) or cinema (where time is segmented and distorted to control the narrative flow). In fact, most often it is the departures from reality which prove to be the most interesting, and presence-inducing, aspect of any work of art.

A distancing of computer-mediated creation from the realistic paradigm is also supported in some research studies done regarding telepresence. One such study, by Kristine Nowak and Frank Biocca, found that users reported increased levels of presence when faced with less realistic avatars [7]. While this study may not be conclusive in this regard, it definitely suggests that, when it comes to presence, a more symbolic approach to the creation of computer-mediated environments may provide a functional advantage over realistic ones. Video game designers have also discovered that players want to be able to recognize reality but don't want its restraints, and it has been suggested that a certain amount of "virtual unreality" is required to maintain the user's attention ([8], 59-60). This tendency towards the symbolic or unreal can be understood as a simple consequence of the mimetic nature of telepresence. While this may be surprising in some models of telepresence, it is the natural consequence of recognizing the mimesis of presence.

Conclusion: Presence as Artistic Product

Treating presence in different media as a mimetic subject opens up an entire range of reflection which would be otherwise impossible. It allows us to better understand and analyze the creative elements of computer-mediated design, without having to understand presence within a positivist or scientific framework. New models of interpersonal mediated communication also become possible using mimesis as the central conceptual tool. By shifting to a mimetic model of presence, we can begin to push the limits of both the expressive limits of computer interfaces and our ways of understanding our reactions to them.

References

[1] H. Blau. *The Eye of Prey: Subversions of the Postmodern*. Indiana UP. 1987.

[2] B. Laurel. *Computers as Theatre*. Addison-Wesley. 1993.

[3] Aristotle. *Poetics*. Perseus Digital Library <http://www.perseus.tufts.edu/>. 4^{th} century BCE.

[4] L. Manovich. *The Language of New Media*. MIT Press. 2002.

[5] Murray, Janet. "From Game-Story to Cyberdrama," *First Person*, N. Wardrip-Fruin & P. Harrigan, eds. MIT Press. 2004. pp. 2-11.

[6] W. Benjamin.Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction," *Video Culture: A Critical Investigation*. Visual Studies Workshop Press. 1986.

[7] K. Nowak, F. Biocca. "The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments." In *Presence: Teleoperators & Virtual Environments* 12/5, 481 – 494. 2003.

[8] Pietro, Guardini. "Virtual Unreality of Videogames." *PsychNology Journal* 1/1, 57-70. 2002.

From Film to the Web: Presence and the Medium

Kimberly A. Neuendorf
Cleveland State University
{k.neuendorf@csuohio.edu}

Abstract

This presentation lays a foundation for the development of a model of presence that considers the form and content characteristics of a medium, the various types of presence, and the range of outcomes of presence.This model development process relies on five assumptions:

1. Media form has influenced presence potential, and differentially across the types of presence (e.g., Lombard & Ditton et al.'s six types [1]; or, Biocca, Harms, & Burgoon's work toward three "orders" of social presence in mediated communication [2]).

Film has been called "the original immersive medium" [3]. It includes many unique presence-inducing characteristics, especially when experienced in the darkened environment of a movie theater [4]. Other media to follow have featured sensory channel capabilities that have enabled or limited presence potential. Radio, recordings, and pod-casting have capitalized on the auditory capacity, while print media have emphasized the pictorial. The evolving online environment has introduced new modes of presence induction, such as the immediacy of both visual and auditory cues. Gaming has adopted much of the "language" of film to ensure a level of familiarity that may be presence-invoking [5].

2. While the importance of considering content as well as form has been acknowledged on occasion (e.g., Bracken & Botta's inclusion of TV genre type [6]; [7]), the majority of presence studies have emphasized form over content. Further, although content factors are ostensibly those that may transcend medium, there are clearly medium/content patterns that have influenced presence potential.

3. Possible interactions between form and content should also be considered. For example, a study of reactions to presidential candidates during the 2000 debates indicated that large-screen presentations may lead to decreased presence evaluations [8]. Here, the content (i.e., faces of politicians) and form (i.e., large-screen closeups) may have interacted; large-screen closeups of other content types (e.g., sports footage, nature scenes) have tended to result in higher presence outcomes.

4. Existing theories from media and communication literatures need to be accessed to help identify "critical variables" [9] in the study of presence. While some attempts to incorporate existing theories may be found (e.g., Lombard & Ditton et al.'s invocation of parasocial interaction with TV characters [1]), such integration attempts have been limited.

An examination of classic film literature is illustrative for the investigation of form attributes of film, including Hugo Munsterberg's [10] analysis of film form as mirroring mental activity; Sergei Eisenstein's work during the 1920's that considered precise filmic techniques that generate cross-modal (synesthetic) sensations; Pudovkin's contention that sound reproduction has greater veracity than pictorial reproduction. Other media and communication theories may be accessed—e.g., genre theory, expectancy theory, and uses and gratifications may aid in an understanding of the contribution of content elements in the cases of film, radio, television, gaming, and recorded music.

5. Efforts should be made to develop a more comprehensive approach to modeling presence types, media form and content, and theory and effects. Seminal attempts to typologize media by presence potential have maintained a narrow focus on particular presence types (e.g., Short, Williams, & Christie's treatment of social presence [11]) or on specific applied contexts (e.g., Rice's exploration of media in organizational task settings [12]). Thus, a model of (a) both the main effects and interactions of (b) form and content in (c) producing various types of presence, and (d) leading to presence profiles that moderate communication effects in different ways, ought to be considered.

References

[1] M. Lombard, T.B. Ditton, with D. Crane, B. Davis, G. Gil-Egui, K. Horvath, J. Rossman, S. Park. Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. Paper presented at the 3rd International Workshop on Presence, Delft, The Netherlands. 2000.

[2] F. Biocca, C. Harms, J. Burgoon. Guide to the Networked Minds Social Presence Inventory (Version 1.2): Measures of co-presence, social presence, subjective symmetry, and intersubjective symmetry. Retrieved October 26, 2005 from *http://mindlab.msu.edu/networkedminds*.

[3] K.A. Neuendorf. Presence and film: The original immersive medium. Presentation to the Central States Communication Association, Cleveland, OH. April 2004.

[4] J. Steuer. Defining virtual reality: Dimensions determining telepresence. In F. Biocca, M.R. Levy (Eds.), *Communication in the Age of Virtual Reality,* 33-56. Lawrence Erlbaum Associates. 1995.

[5] L. Manovich. The Language of New Media. MIT Press. 2001.

[6] C.C. Bracken, R. A. Botta. Presence and television: Form versus content. Paper presented to the 5th International Workshop on Presence, Porto, Portugal. 2002.

[7] M. Lombard, T.B. Ditton. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 13 (3).* 1997.

[8] C.C. Bracken, K.A. Neuendorf, L.W. Jeffres. Source credibility, presence, and screen size. Paper presented to the Visual Communication Division of the Association for Education in Journalism and Mass Communication, Kansas City, MO. July 2003.

[9] K.A. Neuendorf. The Content Analysis Guidebook. Sage Publications. 2002.

[10] H. Munsterberg. The Film: A Psychological Study. D. Appleton & Co. 1916.

[11] J. Short, E. Williams, B. Christie. The Social Psychology of Telecommunications. John Wiley & Sons, Ltd. 1976.

[12] R.E. Rice. Task analyzability, use of new media, and effectiveness: a multi-site exploration of media richness. *Organization Science*, *3*(4), 475-500. 1992.

Relational Presence in Distanced Interdependent Relationships

Katheryn C. Maguire, Stacey L. Connaughton
Cleveland State University, Purdue University
{k.maguire77@csuohio.edu, staceyc@cla.purdue.edu }

Abstract

This paper synthesizes research from the presence literature (i.e., social presence, co-presence) as well as organizational and relational communication literatures to introduce the concept of relational presence: the perceived salience of human-to-human connection in mediated contexts. Relational presence ranges from "fully present" to "absent", and involves 4 dimensions: degree of mutuality, proximality, cognitive orientation, and tangibility. Understanding relational presence may be the key to explaining the difference between successful and problematic distanced interdependent relationships.

Keywords--- **social presence, distanced relationships**.

Distanced interdependent relationships (DIRs), across several contexts, are an ever-increasing occurrence. Up to one million marriages and as many as one-third of college premarital relationships experience long-term separations annually [1]. Distanced relationships are also common in organizational contexts. Indeed, distributed organizations and teams are an integral part of governmental, political, business and military organizations. Although some argue that DIRs are fundamentally different from proximal relationships, we argue that they are actually quite similar in that unit members in both situations must depend on one another to achieve their goals. The difference, we suggest, lies in the preferred and/or available means of interaction. Furthermore, although scholars have identified several disadvantages of DIRs, advantages have been noted as well [2-3]. A central question posed in much of this research is: What makes distance "work" in some relationships and not in others? In response, we introduce the concept of *relational presence* (i.e., the perceived salience of human-to-human connection in mediated contexts) to explain the difference between successful and problematic DIRs.

On the surface, relational presence might seem similar to the concept of *social presence*, which has been defined as the "sense of being with another" [4]. Indeed, we view relational presence as a specific type of social presence that is limited to human-to-human connection (as opposed to connections with artificial intelligence) and facilitated by communication technologies within the DIR context. Relational presence also seems similar to *co-presence*, a person's sense of the salience and accessibility of another [4]. Whereas co-presence implies that social actors have a sensory awareness of one another, relational presence suggests a cognitive orientation to the connection itself, above and beyond the unit members, which can happen with or without interactivity. As Sigman states, "relationships are 'larger' than the physical presence or interactional accessibility of the participants. Social relationships can therefore be said to be continuous, or to be oriented to and produced as such by relationship partners" [5]. Building on Sigman's argument, we suggest, that unit members can experience varying degrees of relational presence *independent* of other unit members.

Relational presence exists along a continuum ranging from fully present absent, and it consists of four dimensions: degree of mutuality, proximality, cognitive orientation, and tangibility. A fully present relationship is one in which unit members are physically proximate to and focused on one another; thus, the connection among unit members is at more apparent and is reinforced through direct sensory input as they work towards unit goals. An absent relationship, on the other hand, is one in which members are physically and psychologically distanced from one another; a connection is less apparent as unit members focus on their own goals. We hypothesize that unit members in successful DIRs are able to cultivate the degree of relational presence needed to accomplish particular goals, some of which may require a higher degree of relational presence while others necessitate greater distance among unit members. Future research is needed to explicate *relational presence* and test the propositions offered in this presentation.

References

[1] Aylor, B. Maintaining long-distance relationships. In D. J. Canary & M. Dainton (Eds.), *Maintaining relationships through communication: Relational, contextual, and cultural variations,* 127-140. Mahwah, NJ: Lawrence Erlbaum. 2003.

[2] Connaughton, S.L., and Daly, J.A. Leadership in the new millennium: Communication beyound temporal, spatial, and geographical boundries. In P. Kalbfleisch (Ed.), *Communication Yearbook,* 29(pp. 187-213). Mahwah, NJ: Lawrence Erlbaum Associates. 2005.

[3] Stafford, L. *Maintaining long-distance and cross-residential relationships.* Mahwah, NJ: Lawrence Erlbaum. 2005.

[4] Biocca, F., Harms, C., & Burgoon, J. K. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence*, 12, 456-480. 2004.

[5] Sigman, S. J. Handling the discontinuous aspects of continuing social relationships: Toward research of the persistence of social forms. *Communication Theory, 1,* 106-127. 1991.

Social Presence in Virtual Teams

Guowei Jian, Joseph Amschlinger
Cleveland State University
{g.jian@csuohio.edu, stamschlinger@yahoo.com}

Abstract

This paper seeks to identify processes that construct the sense of social presence in virtual teams. A review of extant literature on virtual teams uncovers identification, structural interdependence, and leadership as three key processes in constructing and maintaining social presence. Research suggests that technologies function as an enabling, instead of determining, factor.

Keywords---**social presence, virtual teams**.

The accelerated growth of information and communication technologies (ICTs) in the past two decades has enabled organizations to adopt geographically distributed work teams, known as virtual teams [1]. The fundamental difference between virtual and traditional teams lies in their modes of social presence [2]. Electronic proximity, specifically, corporeal telepresence [2] in virtual teams replaces physical proximity that characterizes traditional work teams. The sense of social presence [2], we argue, is an important dimension of team cohesion, which has been shown to have strong reciprocal relationship with team performance [3]. Our research question in this paper is, *What are the processes that construct and maintain the sense of social presence in virtual teams*? We approach this question by reviewing extant literature on virtual teams. Using several electronic databases, we identified more than forty articles on virtual teams published in twenty-nine peer reviewed scholarly journals across six disciplines from 1998 to 2006.

Our review identified three processes that are instrumental in constructing and maintaining the sense of social presence: identification, structural interdependence, and leadership. First, team identification is both a psychological and communicative process that helps create the sense of togetherness. A common team identity is a resource sustained through communication and unites team members through psychological attachment [4]. Second, structural interdependence is required at both organizational and team levels. For instance, at the team level, interdependence could be achieved through high quality of goal setting process and high task interdependence [5]. At the organizational level, reward and recognition systems should promote outcome interdependence. An interdependent structural design creates opportunities for interaction and establishes the salience of other team members. Finally, leadership has a prominent role in helping achieve the structural and psychological connectedness through communicative actions, such as role specification, feedback, and motivation. Research has noted the high communication demand on leaders of virtual teams [6, 7].

These three processes are not mutually exclusive. They interact and reinforce each other in creating the sense of social presence. For example, both identification and structural interdependence are contingent on the quality of leadership process in a team. Structural interdependence is conducive to creating a common identity.

Additionally, a common theme across the studies we reviewed is that technology is an enabler in building the sense of social presence but not the solution. Technological choices should be determined by task characteristics, relationship needs, and social contexts [8]. Attention to social contexts is especially important considering the diverse cultural backgrounds that virtual team members tend to have. Therefore, it is not the bandwidth of an ICT but the fit among task, relationship, context, and technology that helps bring out the sense of social presence in virtual teams.

References

[1] Lipnack, J., & Stamps, J. *Virtual teams: People working across boundaries with technology* (2^{nd} Ed.). New York: John Wiley & Sons. 2000.

[2] Zhao, S. Toward a taxonomy of copresence. *Presence, 12*, 445-455. 2004.

[3] Keyton, J. Relational communication in groups. In L.R. Frey, D.S. Gouran, & M. S. Poole (Eds.), *The handbook of group communication theory and research* (pp. 192-222). Thousand Oaks, CA: Sage. 1999.

[4] Fiol, C.M., & O'Connor, E.J. Identification in fact-to-face, hybrid, and pure virtual teams: Untangling the contradictions. *Organization Science, 16*, 19-32. 2005

[5] Hertel, G., Konradt, U., Olikowski, B. Managing distance by interdependence: goal setting, task interdependence, and team-based rewards by virtual teams. *European Journal of Work and Organizational Psychology, 13*, 1-28. 2004.

[6] Connaughton, S.L., and Daly, J.A. Leadership in the new millennium: Communication beyound temporal, spatial, and geographical boundries. In P. Kalbfleisch (Ed.), *Communication Yearbook,* 29(pp. 187-213). Mahwah, NJ: Lawrence Erlbaum Associates. 2005.

[7] Kayworth, T.R., & Leidner, D.E. Leadership effectiveness in global virtual teams. *Journal of Management Information Systems. 18*, 7-40. 2002.

[8] Maruping, L. M., & Agarwal, R. Managing team interpersonal processes through technology: A task-technology fit perspective. *Journal of Applied Psychology, 89*, 975-990. 2004.

Is Bigger Really Better? An Experimental Study of Presence and Online Political Advertising

Edward M. Horowitz, Ph.D.
Cleveland State University
{e.horowitz1@csuohio.edu}

Abstract

A pilot study was conducted to explore the relationship between presence and online political advertising. Ten college students watched online political advertisements and ten watched political advertisements on a 24-inch TV. Results indicated that viewers of online political ads experienced a greater sense of presence than those watching the ads on TV and greater political affect than viewers of TV ads. Findings suggest political candidates should increase usage of online political ads.

Keywords--- **presence, political advertising**.

Political communication researchers have long identified the importance of an emotional connection that must be made between candidates and voters to gain political support. Ever since the first televised presidential debate between Kennedy and Nixon in 1960 the "winning" candidate has usually been the one that has been able to make a better emotional connection with the viewing audience [1]. This suggests that the media are the conduit for a para-social relationship that exists between political candidates and the public. Beyond debates, candidates focus particular efforts on reaching voters through commercials. Voters may develop strong affect (positive or negative) for political candidates, despite the fact that most members of the public rarely meet or personally get to know a candidate in any way other than via a mediated experience.

Communication scholars within the growing field of presence research have documented the mediated perceptions and the resulting psychological processes in which a para-social relationship is formed even when such a relationship is rationally inappropriate or illogical. Research has found that individuals experiencing such para-social relationships are affected by the communication styles of social actors in mass media and by interpersonal distance cues [2].

This study attempts to explore the para-social relationship and experience of presence within online political advertising. Sitting in front of a 17 inch computer monitor, individuals may experience a sense of presence as they watch online political advertisements. Researchers studying the phenomenon of presence have found in a variety of contexts that size and quality of visual image is directly correlated to the intensity of viewers' presence-related responses [3]. In addition, experiments suggest that the psychological immersion experienced through computer interactivity is equal in its effects on the brain as real-life human interaction [4]. While a 17 inch computer monitor may not be as large as a big screen TV, looking at the screen from the seat of one's desk chair can place that political advertisement fully within the viewer's visual frame.

During the 2004 presidential campaign more than 40 percent of online users used the Internet to find political material about the election—more than 50 percent higher than in 2004. Young adults, whose political socialization has been directly impacted by the Internet, are particularly heavy users of the Internet for politics [5].

Twenty college students participated in a pilot study to test these two research questions. RQ1: Do viewers of online political advertisements experience a greater sense of presence when watching political advertisements on a desktop computer monitor? RQ2: Does candidate affect increase after watching online political advertisements in which viewers experience a greater sense of presence? Half of the participants watched online political advertisements while the other half watched televised political advertisements (24 inch). Results indicate that men experience a greater sense of presence than female participants when viewing online political advertisements. Candidate affect was also greater for respondents watching online political advertisements, although this was only partially supported. Overall, results suggest that candidates should consider increasing their use of online political advertising.

References

[1] Schroeder, A. (2001). *Presidential debates*. New York: Columbia Univ. Press.

[2] Lombard, M. (1995). Direct responses to people on the screen: Television and personal space. *Communication Research, 22*(3), 288-324.

[3] Reeves, B., Detenber, B., & Steuer, J. (1993, May). *New televisions: The effects of big pictures and big sound on viewer responses to the screen.* Paper presented to the Information Systems Division of the International Communication Association, Washington, D.C.

[4] Reeves, B., and Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places.* Stanford, CA: CSLI Publications.

[5] Pew Internet & American Life Project. *The Internet and Democratic Debate*. October 27, 2004, page 1. Available online at: http://www.pewinternet.org/ PPF/r/141/ report_display.asp

Content knowledge and thematic inertia predict virtual presence

David Nunez, Edwin Blake
University of Cape Town
{dnunez@humanities.uct.ac.za, edwin@cs.uct.ac.za}

Abstract

This paper informs the debate between the impact of content and form factors on presence. From cognitive principles, we predict that the content of a VE will affect presence by interacting with expectations held by the user. Furthermore, a particular cognitive tendency (thematic inertia), should facilitate the effect of the expectations. A sample of 461 users of desktop based flight simulations was measured on ten predictors, including degree of simulation related content knowledge (generalized and specific knowledge), thematic inertia, as well as controls for age and immersion/display factors. The ITC-SOPI was the dependent variable. The data suggest that content factors explain almost as much presence variance as form (immersion) factors. As predicted, thematic inertia is a reliable predictor. Also, the degree of generality of content knowledge predicts presence (with knowledge of the specific content being an inverse predictor). This strongly suggests that the degree to which a simulation is able to match the expectations of its users is an important element of the presence experience.

Keywords--- Presence, Content, Cognition, Theory.

1. Introduction

A large body of literature, both theoretical and empirical, exists to support the notion that presence is a function of display and immersion related variables (see [1] for an extensive review). More recent examinations into the role of content related factors on presence, however, remain controversial. The debate, which could be named the content-form debate, was most explicitly delineated by Slater [2]. In that paper, Slater makes a strong distinction between *spatial presence* (the conventional concept that a virtual place is experienced as if it were not mediated), and a host of concepts, such as engagement and involvement, which have become incorporated into definitions of presence by others (such as [3]). By analogy, he argues that presence comes about by the form in which information is presented to the subject; interest, involvement, and so on are brought about by the subject's relationship to the content – therefore, content related factors are considered as not determining presence. The argument that form and content are separate in media has existed for some time;

for instance, during the 1960s discussions existed about whether television (a new medium at the time) was able to deliver novel types of content or not [4]. In general, media theorists consider form and content to be theoretically separate concepts [5]. Logically, they are neither equivalent nor necessarily related. However, there is not much empirical evidence to suggest that, from a user perspective, they might not be related, causally or otherwise. In the presence literature, the role of content has generally been discussed theoretically only (for instance, [6, 7, 8]). The consensus seems to be that for presence to occur, the environment must make some sense or contain some meaning for the subject; and it is the content that provides that meaning. For example, in a factor analysis of eight presence measures, [9] found that the factor *drama* (the degree to which the virtual environment presents a story in which events unfold in a meaningful, predictable way) ranked 4th out of 8 extracted factors, and had an eigenvalue greater than 3 [9]. Such findings are quite suggestive that the content of a VE does have a role to play in the presence experience. Given that little empirical work has been done on this problem, and given that it is a focus of attention in presence theory, it is worth making a detailed examination of the extent to which content factors play a role in presence.

It is difficult to understand how content may affect presence without some theoretical framework. A recent development in presence theory, expressed in [7] and [10], is that presence does not occur from perceptual data alone (as an illusion such as vection does), but rather is constructed from both perceptual and conceptual data by the subject. In this view, the percepts can only be constructed into an experience of space in the correct cognitive context [10]. This context is likely to be affected by two factors: temporary effects such as priming [11, 12], and more permanent effects associated with knowledge of a particular theme [13]. The impact of these temporary effects on presence have been empirically investigated to a limited degree [11], but the long term effects remain in the domain of theory.

One reason why content may not have attracted a great deal of research attention is the difficulty of working with it as an operationalized variable. Are there factors along which different contents can be compared? Is there a way to measure the impact that a particular content area will have on subjects? We believe that for the purpose of examining its effect on presence, the particular content itself is not central. Rather, it is the examination of how the content is integrated and processed cognitively that might lead to viable research designs [10]. To quantify this degree of integration, one can, according to a schemata based theory, measure a participant's knowledge of the content [14, 10]. How this existing knowledge then interacts with the VE display during the mediated experience can be usefully modeled by thinking in terms of user expectations. As the user begins the experience, the VE content cues a slight activation of particular schemata. If the content is well integrated, then that activation will spread efficiently though the user's semantic knowledge networks, leading to expectations of subsequent experiences in the VE [15, 13]. If the VE matches these expectations, one can expect a coherent cognitive construction of the environment [10]. If the VE fails to match those expectations, then an impoverished presence experience will result [16, 10]. If one accepts this explanation, then one can categorize the user's knowledge in terms of the types of expectations it will lead to – detailed knowledge will lead to highly specific expectations, which will be hard for the VE to match; generalized knowledge will lead diffuse expectations which should be easier for a VE system to match [16]. Therefore, one can expect an expert in a particular content area to find simulations of that content to be largely unsatisfactory, unless the simulation's content has been designed to a high degree of fidelity. They would constantly notice errors in the simulation, and would therefore have reduced presence experiences. On the other hand, a novice in that content area, with less content knowledge would find the same simulation satisfactory due to having only very general expectations for that content, and might therefore experience more presence. This is analogous to the well-known "uncanny valley" phenomenon found in simulations of humans [17]. Almost all people have extremely detailed knowledge of the human form (albeit largely implicit), which leads to very specific expectations. A simulation must be of an extremely high degree of fidelity to match such an expectation; indeed, most contemporary systems fail at this task, leaving users largely unsatisfied by the simulation. Our model of knowledge, expectation, and simulation matching of these expectations can essentially be understood as a general explanation of the uncanny valley phenomenon, from a cognitive perspective.

That content could have an effect on the user's experience in this way seems plausible; however the questions of whether this effect is on spatial presence or other related factors, as argued in [2], and of the degree of impact that content factors have on presence are ones which must be addressed empirically. From this discussion, we can define three broad aims for this study:

1. To examine the role of content knowledge on the presence experience (both in terms of spatial presence and other related factors such as engagement). We will also examine if a difference of effect exists between general and specific knowledge.

2. To examine the relative importance of temporary content effects. This can be done by examining thematic inertia, a measure of the degree of cognitive integration of a particular semantic content area [12] and priming, the phenomenon where users engage in behaviors to prepare themselves cognitively for the VE experience [12]. However, we will examine these effects on a standardized presence measure, rather than using the evolutionary approach used in [12].

3. To provide some sense of the relative contributions of form related and content related factors, so as to shed some light on the content-form debate.

2. Method

To achieve these aims, we followed a relational design, preferring to collect a large sample of habitual VE users in naturalistic conditions over a controlled experiment using a smaller sample. We therefore used a large scale online survey of computer game players who play flight simulation games. We measured our sample on a number of content, cognitive and form/immersion variables, and used these to predict their presence during their last simulator session using the ITC-Sense of presence inventory (ITC-SOPI) [3].

2.1 Procedure

The study was advertised as a 'flight simulator gaming habits' study, and posted as an on-line survey. A number of web-sites were selected to advertise the study: These were either web-portals to the flight simulation gaming community (presenting news, downloads, etc. relating to the hobby), or web forum sites whose primary purpose is the discussion of flight simulation related topics. The site administrators of 10 such sites were contacted and asked to post a link to the study on their site. Of these, 7 responded (70% response rate). In order to provide an incentive for participation, Flight1.com, an on-line retailer of flight simulation products, was recruited as a sponsor of the study. They provided three popular flight simulator products as prizes for a random draw of subjects.

The subjects were provided with a URL for the study website. On entering the site, they were provided with information about the study (enough to meet the informed consent ethical requirement while reducing possible expectancy effects), and if they agreed to continue, were presented with instructions, and then the questionnaires. The subjects were first presented with our ten content, cognitive and control factors, and then were asked to report the title of the last flight simulator they played, and the number of days since that last session (the experience-measure delay). They were then asked to complete the ITC-SOPI with regard to that session. With the exception of not administering the ITC-SOPI immediately after the VE experience, we followed all the administration guidelines given by the authors of the ITC-SOPI. Once all the items were completed, the subjects were asked to fill in their email address for entry into the random draw (this information was not stored together with their actual data).

2.2 Sample

A practical sampling problem arises in content related research – what population has varying degrees of knowledge of one well-defined content area, for which there exists a virtual environment which implements that content area? We decided to turn to the population of computer game players who use flight simulation games. For this case, the content is well defined (aviation), and knowledge of it can be reasonably measured – by either asking subjects to report on their level of knowledge, or by examining their interest in other activities related to the content area (reading aviation books, visiting aviation web-pages, etc.) One can also determine if the subjects have generalized or specific knowledge with relative ease (see measures in 2.3 below). Also, the flight simulation playing population is large, and easily accessible; it therefore serves as a useful starting point for such an investigation.

A total of 503 responses were collected from flight simulation players (see section 2.1 above for a description of the recruitment procedure). Of these, 461 (91.6%) reported using *Microsoft Flight Simulator 2004: A Century of Flight* during their last simulator session. These were selected as the sample for the study. This was done to control for software platform cross users. In effect then, this is a self-selected, volunteer sample. The sample consisted of 100% men. One may be forgiven for assuming that this is a massive overrepresentation, but this gender distribution probably correctly represents this particular population; the flight simulation site AVSIM.com, in the 2003 edition of its yearly census of users, found only 2.6% of users to be women with a sample of 14,247 [18]. Of course, such a population precludes any investigation of gender effects. Due to this, we decided to exclude gender as a variable in this particular study. This decision has some justification - a recent review [19] which examined nine studies considering gender in presence, found a difference in only one of those studies. In terms of age, the sample was far more diverse; the mean age was 31.7 years, with a range of 12 to 65 (s = 13.07 years).

2.3 Measures:

The ITC sense of presence inventory (ITC-SOPI) [3] was used as the measure of presence. This questionnaire measures four factors of the presence experience: Spatial presence, engagement, naturalness and negative effects. These are defined as follows (from highest to lowest degree of variance explained):

Spatial presence: A sense of physical placement within the VE, and of interaction with the objects in the VE.

Engagement: A sense of psychological involvement and a tendency to enjoy the VE experience.

Naturalness: A sense that the VE is believable and lifelike, or realistic (this factor is also referred to in [3] as *ecological validity*).

Negative effects: Negative physiological reactions to the VE experience such as dizziness, eyestrain and headaches.

The ITC-SOPI is a particularly useful measure in that each of the factors provides a separate score for the experience. This effectively allows it to satisfy a number of presence concepts at once, effectively separating out spatial presence from the other factors. For instance, if one follows the presence concept presented by Slater [2], then one can simply consider the spatial presence factor of the ITC-SOPI. However, if one is sympathetic to the views of IJsselsteijn and colleagues (eg, [20]), then one can consider both the engagement and spatial presence factors. This flexibility, its suitability for use across any medium representing a VE, and its high degree of psychometric evaluation [3] makes it a particularly useful instrument for a large study involving home computer based flight simulation software.

Our second major measure is a questionnaire, developed for use in this study, containing ten content knowledge, cognitive and general control factors to be used as predictors (see table 1 below for a summary). The factors are defined as:

Thematic inertia: This is the same concept proposed in [12]. It is the tendency for a subject to engage in thematically related activities (e.g. reading about aviation, as well as playing aviation related games). In this study, we considered situations where non-simulation activities (reading a book, taking a real flight) led to either a desire to play a flight simulator, or the actual playing of a flight simulator.

Priming: This refers to subjects engaging in particular behaviors before playing flight simulations, so as to set a cognitive context for the simulation experience in some way [21]. We considered cases where subjects read aviation books, manuals, aeronautical charts, or engaged in similar activities immediately before a flight simulation session. We hypothesize that priming and thematic inertia are closely related, although this has not been specifically investigated. Thematic inertia is likely a tendency or cognitive style, while priming is one of the behaviors which expresses that tendency.

Factor with Number of items (Cronbach's alpha in brackets)	Sample item
Thematic inertia 5 (0.79)	"Reading about real world aviation or flight in a book, magazine or web-page makes me want to play a flight simulator."
Priming 5 (0.76)	"Before I play a flight simulator, I usually read an aviation/flight book, magazine, or web page."
Content knowledge 8 (0.63)	"I prefer to fly virtual flights around places which I have been to in real life."
Hobby cluster 7 (0.55)	"How many model aircraft have you built in the past year (scale models or radio-controlled)?"
Simulator mechanics knowledge 9 (0.76)	"Have you ever created an aircraft (exterior model, flight model, etc.) for any flight simulator?"
Presence management 8 (0.69)	"What size of screen/display do you usually play simulators with?"
Evaluation of simulator realism 6 (0.76)	"The experience provided by current commercial flight simulators is like the real thing."
Enjoyment 6 (0.68)	"I normally find playing commercial flight simulators to be a fun experience."
Experience-measure delay 1 (-)	"How many days ago was this last session?"
Age 1 (-)	"What is your age?"

TABLE 1: The ten content, cognitive and control factors used to predict ITC-SOPI scores

Content knowledge: This refers to knowledge of the actual content being simulated; that is, specific knowledge of the real places and aircraft being simulated. This factor allows the measurement of the influence of the fit between the simulation display and a specific expectation of the scene. This follows our model that the amount of information about the content held by the subject is likely to correlate with how specific the subject's expectations are (see the discussion in 1 above).

Hobby cluster: This is a measure of the degree to which the subject engages in other activities which are related to aviation, such as building model aircraft or reading aviation publications. It is of interest in this study because it represents generalized knowledge of the content being simulated, in contrast to the *content knowledge* factor, which measures specific knowledge.

Simulator mechanics knowledge: This factor considers the subject's knowledge of how simulation software works. Apart from measuring this directly as done by [12] (a method which may give rise to self-report biases), we further estimated it by using the number of modifications, add-ons or simulation content created by the subject - we assume that being able to create simulator content requires knowledge of how the simulation works. This factor can be used to control for information relevance, by contrasting its effect with the content relevant *content knowledge* and *hobby cluster* factors.

Presence management: This is the same factor defined in [12]. It represents measures taken to improve the immersion of the hardware platform, and of the user to reduce attention distracters. We expanded this factor to include the use of consumer grade simulation input devices (joysticks, control yokes, rudder pedals, throttle quadrants, etc.) which are widely available. These not only provide improved control for the user, but also act as passive haptic devices [22], as they mimic the shape of real aircraft controls. This factor represents our notion of a display and attention related factors. It also expresses some consensus of immersion and display factors identified as important in the literature (as discussed in 1 above).

Evaluation of simulator realism: This is a measure of how realistic the subject considers flight simulations to be, in general terms. Notice that we do not use this as a measure of the realism of the system, but of the *perceived* realism. This cognitive factor represents arguably the most abstract level of expectation. Subjects who rate a simulation as realistic are presenting an interpretation bias; we can thus infer, according to the constructionist concepts of [10] and [14] that subjects who score high on this factor are less likely to interpret simulation artifacts as detracting from the experience.

Enjoyment: This factor estimates how fun or enjoyable the subject finds simulations in general. This is an important control, as there is evidence to suggest that presence varies with enjoyment of the experience [23, 24]. Given that the subjects in this study use flight simulations for recreation, it is likely to be a factor. It is also possible that subjects who find the experience enjoyable would have a bias to overestimate their presence (the converse bias is also technically possible, but due to the self-selection of this sample, it is unlikely).

Experience-measure delay: As this study takes the unusual step of asking subjects to complete the ITC-SOPI with regard to their last flight simulation experience (see 2.1 above for the procedure), this factor was included to control for any possible memory or delay effects. The granularity of this measure was chosen as one day.

Age: This is a control for two factors: The possible natural covariance of age with cognitive factors (such as attention, spatial ability, etc.) due to aging, as well as for the possibility of a general correlation between age and presence (as reported in [25]).

2.4 Models and analysis strategy

In this type of research, it is usual to create a single model from a set of predictors, and then evaluate the usefulness of that model by examining its fit to the data (as was done in [12]). However, we propose to go one step further by comparing the fit of two models to each other: The first will be our model including three sets of factors: display and attention factors (form related); content and cognitive factors (content related) and general control factors. The second model will be a reduced, conservative model including only display and attention factors (form related) and the general control factors. This comparison will allow us to evaluate our data in terms of the content-form debate: in essence, the model including content and cognitive factors represents the content position of the argument, while the conservative model represents the form position. Although it is possible to estimate the contribution of content on presence without this comparison (by examining the partial correlations, for example), making this comparison allows one to link the data to the theoretical debate far more strongly. Statistically, such a comparison is simple – one can perform a significance test on the difference between the error variances of two models [26], which in effect compares the models in terms of their R^2 values.

Although the idea of such a model comparison sounds straightforward in principle, it is far from simple to find a set of form related factors which satisfactorily expresses a consensus of published research. The list of factors we used was derived from the compilation used in [12] but was expanded slightly. These are immersion related factors such as display size and passive haptics used, and attention management strategies (keeping the room dark, preventing interruptions during the experience, etc.). The general control factors include age, amount of time since the subject last played a flight simulation, and their enjoyment of flight simulations (these factors are defined and justified in the measures section in 2.3 above). By comparing the fit of these two models, it should be possible to gain an insight into the relative contributions of content and cognitive factors to particular aspects of the presence experience.

3. Results

The data were analyzed using a set of four multiple regression analyses, one for each of the four ITC-SOPI factors. In each case, the ten predictors listed in table 1 were used. In order to more clearly show the contribution of the content and cognitive factors to the variance of the ITC-SOPI factors, we tested the difference in model fit between the full ten factor model and the reduced four factor conservative model. As discussed in 1 and 2.4 above, this conservative model poses *presence management* as the major predictor (it includes measures of well investigated variables in presence such as display size [27], passive haptics [22], and focusing of attention on the VE [12]). The conservative model also includes *age*, *enjoyment*, and *experience-measure delay* factors as general controls.

3.1 Spatial factor

For this factor, the ten predictor model is significant ($F = 17.41$, $p < 0.00001$) and gives $R^2 = 0.28$. The significant predictors (at the 0.01 level) are *thematic inertia*, *evaluation of realism*, *content knowledge* (as a negative factor), *presence management* and *age* (see table 2 below for corresponding partial correlations). The difference between the fit of this model and that of the reduced conservative model (which has $R^2 = 0.15$) is significant ($F = 12.67$, $p < 0.00001$) – see table 6.

Factor	Partial correlation
Thematic inertia	0.28
Evaluation of realism	0.18
Content knowledge	-0.12
Presence management	0.19
Age	0.14

TABLE 2: Significant predictors for the* spatial *factor, with partial correlations.

3.2 Engagement factor

Again, the ten predictor model is significant ($F = 30.77$, $p < 0.00001$) with $R^2 = 0.40$. The significant predictors (at the 0.01 level – see table 3 below) were the same as for spatial presence. These were: *thematic inertia*, *evaluation of realism*, *content knowledge* (as a negative factor) *presence management* and *age* (see table 3). The difference in fit between this model and the conservative one (whose $R^2 = 0.26$) is again significant ($F = 17.86$, $p < 0.00001$).

Factor	Partial correlation
Thematic inertia	0.33
Evaluation of realism	0.13
Content knowledge	-0.14
Presence management	0.33
Age	0.20

TABLE 3: Significant predictors for the engagement factor, with partial correlations.

3.2 Naturalness factor

Although the ten predictor model for this factor is significant ($F = 18.14$, $p < 0.00001$) and the fit is good ($R^2 = 0.29$), the significant predictors differ from the two previous models. As before, *thematic inertia*, *evaluation of realism*, *presence management* and *age* are significant predictors; however, *content knowledge* makes no contribution, and *priming* is a significant predictor (see table 4).

Factor	Partial correlation
Thematic inertia	0.24
Evaluation of realism	0.25
Presence management	0.16
Priming	0.10
Age	0.19

TABLE 4: Significant predictors for the naturalness factor, with partial correlations.

As with the other models, the difference in fit between this and the conservative model (with $R^2 = 0.13$) is significant ($F = 13.65$, $p < 0.00001$) – see table 6.

3.3 Negative effects factor

The ten predictor model is again significant ($F = 3.57$, $p < 0.00026$), as one would expect with such a large sample size; however, it shows very weak fit ($R^2 = 0.07$). The pattern of predictors is also quite different – only *thematic inertia* and *presence management* are significant predictors (see table 5). Again, this model explains more variance than the conservative model, which itself has a very weak fit ($R^2 = 0.04$), but the effect size of the difference is noticeably smaller – indeed, it does not reach significance at the 0.01 level ($F = 2.60$, $p < 0.02$) – see table 6.

Factor	Partial correlation
Thematic inertia	0.15
Presence management	0.13

TABLE 5: Significant predictors for the negative effects factor, with partial correlations.

3.4 Overall comparison of model fit

Table 6 below summarizes the differences in model fit (R^2) between the ten predictor model and the conservative model.

ITC-SOPI Factor	R^2 for ten predictor model	R^2 for conservative model
Spatial	0.28	0.15*
Engagement	0.40	0.26*
Naturalness	0.29	0.13*
Negative effects	0.07	0.04

TABLE 6: Summary of model fits for the ten predictor and conservative models (asterisk indicates p < 0.01 for the difference in fit between the models)

For all four ITC-SOPI factors, the difference between model fit is significant at the 0.05 level. At the 0.01 level however, the models for the negative effects factor do not show a significant difference in fit. In general, the ten predictor model explains substantially more presence variance than the conservative model.

4. Discussion

For the remainder of this paper, we will consider firstly some caveats and limitations of our design and sample, followed by a discussion of the theoretical importance of the significance and lack of significance of the ten predictors. We will end the paper considering the overall importance of content and cognitive factors in understanding presence. For the purposes of comparison during our discussion, table 7 below provides a summary of the analysis of the ten predictor model presented in section 2.4 above, showing the partial correlations of the significant predictors for each of the ITC-SOPI factors, as well as the model fit.

4.1 Implications of the design and sample

Possible effects of experience-measurement delay

In order to capture this large number of subjects, the design required violating the requirement that the ITC-SOPI be administered immediately after the experience. Although there exists no theoretical explanation as to why an experience-measurement delay should introduce systematic error into the presence measure (or indeed, a prediction of what such an effect would be), we controlled for this by asking for an estimate of the delay. Our analysis revealed no delay effects in any of the ITC-SOPI factors. One may conclude that with a quantification granularity of one day, no experience-measure delay effects are apparent. Although it is possible that an effect exists over a period less than one day, it would have to be a non-linear effect acting over a period of less than one day, and with no further effect after one day.

Positive bias due to enjoyment

As a large number of the participants in this study were likely to be habitual simulation users, it was necessary to control for a possible positive bias in their responses. This was achieved by use of the enjoyment predictor. The reported degree of enjoyment was, as expected, high – a mean score of 29.7 (on a scale ranging from 6 to 42), but it was not significantly skewed. However, it is important to note that enjoyment was not a significant predictor of any of the four ITC-SOPI factors. It is therefore possible to state in subsequent conclusions that the reports of presence given by our sample were not unduly inflated by their enthusiasm for the content.

Factor	Spatial Presence	Engage-ment	Natural-ness	Negative effects
Thematic Inertia	0.28	0.33	0.24	0.15
Evaluation of realism	0.18	0.13	0.25	-
Content knowledge	-0.12	-0.14	-	-
Presence management	0.19	0.33	0.16	0.13
Priming	-	-	0.10	-
Age	0.14	0.20	0.19	-
Overall R^2	0.28	0.40	0.29	0.07

TABLE 7: Comparison of partial correlations and model fit for the significant predictors of the ten predictor model on the 4 ITC-SOPI factors (a hyphen indicates the predictor was not significant at the 0.01 level)

Gender and age of participants

Central to any modeling study is a large sample which is able to represent its population correctly [28]. In terms of age, this is an extremely good sample. In fact, it provides a wider age range than any comparable presence study we have been able to find (compare for instance, with [12], [25] or [29]).

Although this study used a large, self selected sample, it contains no women respondents, which is a concern. The choice of population (flight simulation users) was made as it presents a population of habitual VE users who have knowledge at several levels of relevance of the content of the VE. It is unfortunate that a population which is so useful for our purpose should also have a massively imbalanced gender distribution. From a cognitive perspective, one of the most serious concerns is that of possible gender differences in spatial abilities, which would have a great impact on presence [7]. Currently, there is no definitive answer to whether such a gender difference exists; Some time ago evidence for a difference was clear, but thought to be diminishing over time [30]. Later meta-analyses revealed a more confused, inconclusive picture [31] [32]. It is therefore possible that some of the findings of this study may not generalize across the genders. However, many of the theoretically important findings of this paper (such as the contributions of content knowledge), probably rely more on semantic processes than spatial ones, so even if such differences exist, they may not negate the described effects.

Degree of control over display variables

A key feature of this design is the controlling of system variables to examine the role of content variables. This was achieved in by measuring presence management practices (which includes subjects' control over display and interface variables), as well as sampling only those who use a single software platform. It should be noted that although the software package was kept constant, it is still possible to have slight variations in terms of content as well as display. The software we selected (as with all desktop based simulations) allows users to trade display fidelity for simulation update rate, through adjusting a number of simulation and display parameters within a narrow range. Therefore, the exact degree of visual fidelity which any particular subject experienced during their last simulator session is not known. However, the range of such modifications allowed by the software is limited, so it is possible to understand this control as limiting such effects. Also, it is important to recognize that although this can be correctly understood as a threat to the internal validity of this design, our choice of a large sample of simulation users reporting on their experiences with their usual gaming situation gives this study an enviable degree of external validity; in an important sense, this is an example of the trading off of internal validity for external validity which is unavoidable in this type of work [28].

4.2 Thematic inertia and priming

Most interesting of the results obtained is perhaps the role of thematic inertia. It is a significant predictor of all four ITC-SOPI factors, and in all cases its contribution to the ITC-SOPI factor is either higher or only slightly less than that of *presence management*. We constructed the measure of presence management to include measures of display size and passive haptics, which have been established as important factors in presence [20, 22, 27]; however, the current data suggest that, if one holds the software platform relatively constant, then this cognitive factor is on average at least as important as display and attention related factors. One possible explanation for this phenomenon is that those with high thematic inertia gain more benefit from better displays, and thus have learned presence management strategies. However, this is unlikely, as the overall correlation between presence management and thematic inertia, although significant, is low ($r = 0.31$); also, if this were the case, we would not expect to see both of these factors appear as significant predictors in the multiple regression, due to a high degree of shared variance. A more theoretically driven explanation which is consistent with this data is that while presence management can be learnt, thematic inertia is probably part of a cognitive style and therefore cannot be learnt [12]. As discussed in section 1 above, thematic inertia likely contributes to presence through enabling the spreading of semantic activation; it is therefore probably relatively independent of perceptual factors (associated with the processing of the display). Subjects who are fortunate enough to have high thematic inertia and engage in presence management strategies would undoubtedly have the highest presence scores.

Intriguing is the lack of effect of priming on presence, given that priming has been found to be effective when manipulated experimentally [11]. It only seems to affect the *naturalness* factor, and then only to a slight degree (a partial correlation of 0.1). One possible explanation is that priming is effective, but subjects do not make use of it; however, this is at odds with the evolution of presence maximization strategies argument [12], which proposes that game players will evolve behaviors which maximize their presence experiences. Closer examination of the data reveals that if one re-computes the regressions after removing thematic inertia as a predictor, then priming becomes a significant predictor of all ITC-SOPI factors except negative effects. This suggests that priming has a higher covariance with thematic inertia than with the ITC-SOPI factors. It is probably correct to say that thematic inertia and priming both measure some more general cognitive factor. It seems reasonable to suggest that thematic inertia measures an automatic quality of cognition, where exposure to one type of stimuli associated with a content area (a book) automatically activates cognitions about related stimuli (a simulation) probably by means of the spreading of semantic activation. On the other hand, priming measures *active* engagement in behaviors. It seems reasonable that without the tendency measured by thematic inertia, priming would not be effective. Thus, subjects without the tendency would not have evolved priming behaviour. Also, not all those who have a high degree of thematic inertia would necessarily engage in priming behaviors for any number of practical reasons (limited time, lack of priming materials, etc.); priming would therefore have a much higher degree of error variance than thematic inertia. We can therefore expect thematic inertia to be a better predictor of presence than priming is.

4.3 The role of content knowledge and evaluations of realism

This study used measures of three types of content knowledge: Specific knowledge about the simulated content (*content knowledge* factor), general knowledge about the content (*hobby cluster* factor), and content irrelevant knowledge (*simulator mechanics knowledge* factor). For the spatial and engagement ITC-SOPI factors, specific knowledge of the VE content reduces the presence experience. The effect on the spatial factor can be understood in terms of the expectations for the system held by the user [16], as discussed in section 1 above. Therefore, the more specific content knowledge the user has, the more detailed and specific the expectation will be. Given that the simulation is giving a set degree of fidelity, users with more specific knowledge should notice more mismatches between their expectations and the display, leading to a reduction in presence. With relevant but non-specific knowledge of the simulated content, one would expect this effect to be significantly reduced, as generalized knowledge would not lead to specific expectations; and the data suggest that it has no effect at all. In the case of the engagement factor, we would suggest that subjects would be more engaged with a system that does not violate their expectations, by two mechanisms: firstly, their attention would not be diverted to the errors or omissions in the simulated content, and secondly, they would experience more interest, fun, and have their attention more focused on systems which present the content which they expect. Content irrelevant knowledge leads to no expectation of the interaction with the system, and it therefore produces no effect on presence on either the spatial or engagement factors.

In the case of the naturalness and negative effects factors, there are no content knowledge effects at all. It is simple to understand why the negative effects factors should not be affected by content knowledge, because simulator sickness is well understood to be due to a mismatch in information between the visual and vestibular systems [33], and therefore higher level cognition is unlikely to have an impact. However, how natural a VR system is experienced as, should be affected by one's expectations of the content being modeled, as is the case with the spatial and engagement factors. One might argue that although this is the case, the ITC-SOPI items of this factor are at an extremely non-specific level of interpretation (for example, "the content seemed believable to me"). It is possible that for the given population (flight simulation hobbyists), the lack of effect is due to a floor effect – they all have enough knowledge about aviation that the responses to such items would vary very little. This however is not borne out by the data – the naturalness factor has a mean and variance comparable to the other ITC-SOPI factors. Another possible explanation for this phenomenon is that the expectations associated with the naturalness factor are implicit rather than semantic – that is, they are expectations about how the simulation behaves and responds to input, rather than being explicit expectations about the shape or layout of the physical space and its meaning. Although we do not have any means of supporting this hypothesis with the current data, it is supported, we believe, by the theoretical distinction drawn between implicit and explicit cognition, which is for instance used in implicit memory research (such as [34]) and the mental models literature (for example, [35]).

A final interesting effect which appears in this data (or rather, fails to appear) is with regard to content irrelevant knowledge (the *simulator mechanics* *knowledge* factor). Even though we used a large sample, which would typically give significant results even for small effects, no effects were found. One can conclude that irrelevant information neither contributes nor interferes with presence. This, together with the contribution of relevant information presented above, suggests that the presence experience is constructed from selected subsets of perceptual and conceptual information as suggested in [10] (and to a lesser extent in [14]). It also supports the notion that attention and cognitive resources are allocated in terms of that construction, effectively expressing a bias for content relevant information, while excluding irrelevant information [10]. This suggests that presence is not due to a 'willing suspension of disbelief" (discussed in [36] and [37]), but rather that information which might lead to disbelief is filtered out, so that the only aspect of willingness in the presence experience may the decision to engage with the VR system. If this is true, then the research question becomes: under which conditions does such a filtering process engage, and how can one control the coarseness of the filter?

The question of the contribution of display realism and simulation fidelity to presence has been dealt with in the literature to a largely satisfactory degree [3] [38] [39]. We were interested in the *perception of realism* held by the users of simulations. As argued in [40], the fidelity or realism of a scene could be measured completely objectively by describing the various display parameters, and for content, one could measure fidelity in terms of variations between the simulated model and the actual phenomena being simulated. However, whether a subject finds a simulation realistic is a different matter; it returns to expectations about the content. Asking a subject to provide an assessment of the realism of a simulation is a measure of a very general, high level expectation of the experience the simulation will deliver. In this study, given that the simulation platform was kept relatively constant across all participants, any differences in their perceptions of realism can be more readily attributed to cognitive factors than to display factors; also the items in this factor asked subjects not about the perceived realism of the last simulator session (which would have been an actual measurement of the system), but of simulation software *in general*. We can assume that such an average evaluation of realism would have existed before the subjects played their last simulator session; and therefore it would have acted as an expectation for that session. As this expectation is extremely generalized, it is should be simple to satisfy even with a desktop simulation. This is borne out by the data – the evaluation of realism is a positive predictor (with high partial correlations) of all factors of the ITC-SOPI except negative effects.

4.4 The interaction of expectations and mediated content in presence

As discussed in 4.3 above, a high evaluation of realism was associated with an increase in presence, due to the generalized nature of the expectation associated with realism evaluations. Conversely, for content knowledge, a negative relationship existed with presence, due to the highly specific expectations associated with a large degree of content knowledge. At the same time, knowledge not associated with the simulation content (such as knowledge of the simulation mechanics), had no effect, as it creates no expectation for that particular content. During the mediated experience, the simulation provides a number of perceptual cues on several modalities. If these cues are interpreted as matching the expectations arising from the content knowledge, then the subject will have a coherent presence experience in the system; if not, the lack of match will lead to a reduced presence experience through the attracting of attention to perceived errors in the content, as well as a reduced sense of naturalness and reduced engagement with the material [10]. Such a mismatch is more likely to occur in the face of highly specific expectations, and less likely in the face of generalized expectations. We can summarize this discussion in a general principle: VE relevant knowledge creates a cognitive context in terms of expectations, with more knowledge leading to more specific expectations; and presence is more likely to occur when expectations are matched by the VE system.

This content expectation principle can be seen in the partial correlations of the content relevant factors – evaluation of realism, hobby clustering and content knowledge. If one orders them with respect to how specific an expectation we predict from each one, and plots them against their partial correlations (see figure 1), then the predicted pattern is discernible for all the ITC-SOPI factors except negative effects.

It should be noted that hobby clustering did not give a significant partial correlation with any of the ITC-SOPI factors; this means, in the strictest sense, that the population value of the partial correlation is zero. With regard to the hypothesis of generality of expectation, we can interpret this to mean that at the level of generality found at the hobby cluster level of knowledge, the contribution to presence is negligible; the size of its effect is so small as to be undetectable with our sample.

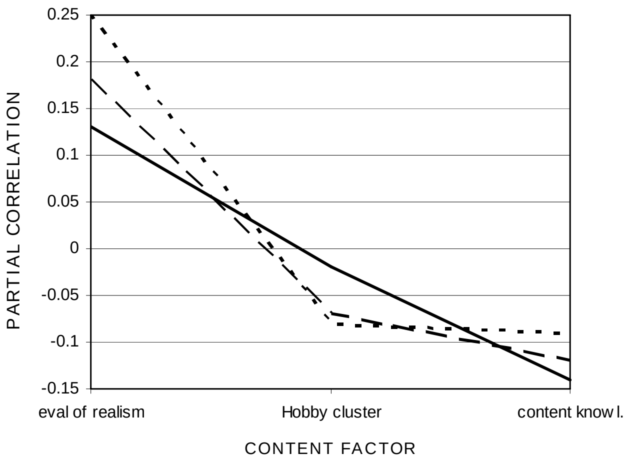


FIGURE 1: Plot of partial correlations of content factors, in decreasing level of generality of expectation (from left to right). The partial correlations with three ITC-SOPI factors are shown - *Spatial* is the dashed line; *Engagement* is the solid line; *Naturalness* is the dotted line.

We would like to suggest that its strength of effect represents a half-way point between the effects of content knowledge and evaluation of realism; we believe our expectations hypothesis is theoretically strong enough to support this idea. However, more data is required to categorically make this statement. A future study could overcome this weakness by refining the measure of generality of expectation so as to produce a continuous model of this effect.

4.5 Relative contribution of content and cognitive factors to presence

In general terms, the models we computed show that the addition of the content and cognitive factors add significant fit over the conservative model, for all ITC-SOPI factors except *negative effects* (summary in table 6 above). The conservative model replicates the large body of published work which argues for the importance of display related factors in presence (such as [22], [40, 2], [41], [42] and others); this validates the general procedure used in this study. However, the difference in fit between the models highlights the importance of considering content related factors when predicting presence. Although we have only included a limited set of display related factors in our *presence management* factor (as any study must), it is highly unlikely that the amount of increase in model fit brought about by adding the content related factors is simply an artifact of our selection of display factors, or of the method used. Furthermore, our use of the ITC-SOPI measure allows us to separate out the effect of these factors on several components of the presence experience. Therefore, if one argues that content factors are an important contributor to engagement but not spatial presence (as is done, for example, in [2]), our models would reply that this is incorrect. We can provide evidence that content factors affect both spatial presence, and the other factors – although it should be noted that it is still unknown if there is a causal flow between the ITC-SOPI factors, so that engagement might cause spatial presence, for example, or vice versa.

5. Conclusion

We have argued that these findings strongly support the idea that parallel to considering the role of the display and immersion in presence (form), it is beneficial to consider the role of VE content. To understand the interaction of content and form factors, it is important to consider the cognitive context in which the form in processed (both in terms of short term effects such as priming, and long term structures such as those measured by thematic inertia).

It should be noted that even our most powerful model (the ten factor model predicting engagement) explained less than half the variance in presence scores; a great deal of work clearly remains to identify further content, display and possibly other classes of variables which are important factors in the presence experience.

Acknowledgements

We would like to thank Jeff Smith of Flight1.com for his contribution of the recruitment incentive products for this study.

This material is based upon work financially supported by the National Research Foundation. Any opinion, findings and conclusions or recommendations expressed in this material are those of the author(s) and therefore the NRF does not accept any liability in regard thereto.

References

[1] M. J. Schuemie, P. van der Straaten, M. Krijn and C. A. P. G. van der Mast, *Research on presence in virtual reality: A survey.*, Cyberpsychology & Behaviour, 4 (2001), pp. 182-210.

[2] M. Slater, *A note on presence terminology.*, Presence Connect, 3 (2003).

[3] J. Lessiter, J. Freeman, E. Keogh and J. Davidoff, *A cross-media presence questionnaire: The ITC-Sense of Presence Inventory*, Presence: Teleoperators and Virtual Environments., 10 (2001), pp. 282-297.

[4] H. M. McLuhan, *Understanding media: The extensions of man*, McGraw Hill, New York, NY, 1964.

[5] S. L. Calvert, *The social impact of virtual reality*, in K. Stanney, ed., *Handbook of Virtual Environment Technology*, Erlbaum, Hillsdale, NJ, 2001.

[6] D. Nunez and E. H. Blake, *Cognitive presence as a unified concept of virtual reality effectiveness*, *1st International Conference on Computer Graphics, Virtual Reality, Visualization and Interaction in Africa (AFRIGRAPH 2001)*, ACM Press, Cape Town, South Africa, 2001, pp. 115-118.

[7] W. Wirth, T. Hartmann, S. Böcking, P. Vorderer, C. Klimmt, H. Schramm, T. Saari, J. Laarni, N. Ravaja, F. R. Gouveia, F. Biocca, A. Sacau, L. Jäncke, T. Baumgartner and P. Jäncke, *A Process Model of the Formation of Spatial Presence Experiences*, 2004.

[8] C. Fencott, *Presence and the Content of Virtual Environments*, *2nd International Workshop on Presence*, Essex, U.K., 1999.

[9] T. W. Schubert, F. Friedmann and H. T. Regenbrecht, *The experience of presence: Factor analytic insights.*, Presence: Teleoperators and Virtual Environments., 10 (2001), pp. 266-281.

[10] D. Nunez, *A constructionist cognitive model of presence.*, *7th International workshop on presence. (PRESENCE 2004)*, Valencia, Spain, 2004, pp. 313-316.

[11] D. Nunez and E. H. Blake, *Conceptual priming as a determinant of presence in virtual environments.*, *2nd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH 2003)*, Cape Town, South Africa, 2003, pp. 101-108.

[12] D. Nunez and E. H. Blake, *Learning, experience and cognitive factors in the presence experiences of gamers: An exploratory relational study.*, Presence: Teleoperators and Virtual Environments (In press).

[13] D. E. Rumelhart and A. Ortony, *The representation of knowledge in memory.*, in R. C. Anderson, R. J. Spiro and W. E. Montage, eds., *Schooling and the Acquisition of Knowledge*, Lawrence Erlbaum Associates, 1977.

[14] W. Wirth, P. Vorderer, T. Hartmann, C. Klimmt, H. Schramm and S. Böcking, *Constructing presence: Towards a two-level model of the formation of spatial presence experiences*, *6th Annual International Workshop on Presence*, Aalborg, Denmark, 2003.

[15] C. Martindale, *Cognitive Psychology: A Neural Network Approach*, Brooks/Cole, 1990.

[16] D. Nunez, *How is presence in non-immersive, non-realistic worlds possible? ACM AFRIGRAPH 2004*, Stellenbosch, South Africa, 2004, pp. 83-86.

[17] M. Mori, *Bukimi no tani (The uncanny valley)*, Energy, 7 (1970), pp. 33-35.

[18] avsim.com, *The AVSIM 2003 survey*, (2003).

[19] C. Youngblut, *Experience of presence in virtual environments*, Institute for defense analyses, 2003.

[20] W. A. IJsselsteijn, H. De Ridder, J. Freeman and S. E. Avons, *Presence: Concept, determinants and measurement.*, *Proceedings of the SPIE*, 2000, pp. 520-529.

[21] D. Nunez and E. H. Blake, *The thematic baseline technique as a means of improving the sensitivity of presence self-report scales.*, *6th International Workshop on Presence*, Aalborg, Denmark, 2003.

[22] B. Insko, *Passive haptics significantly enhances virtual environments*, *Department of Computer Science*, University of North Carolina at Chapel Hill, 2001.

[23] S. Nichols, C. Haldane and J. R. Wilson, *Measurement of presence and its consequences in virtual environments.*, International Journal of Human Computer Studies, 52 (2000), pp. 471-491.

[24] K. E. Bystrom and W. Barfield, *Effects of participant movement affordance on presence and performance in virtual environments*, Virtual Reality: The Journal of the Virtual Reality Society, 2 (1996), pp. 206-216.

[25] W. Barfield and S. Weghorst, *The sense of presence within virtual environments: A conceptual framework.*, in G. Salvendy and M. Smith, eds., *Human-Computer Interaction: Software and Hardware Interfaces*, Elsevier, 1993, pp. 699-704.

[26] J. Neter, W. Wasserman and G. A. Whitmore, *Applied statistics*, Allyn and Bacon, Boston, MA., 1988.

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

[28] R. Rosenthal and R. L. Rosnow, *Essentials of behaviour research: Methods and data analysis.*, McGraw Hill, New York, NY, 1991.

[29] C. Youngblut and B. T. Perrin, *Investigating the relationship between presence and task performance in virtual environments.*, *IMAGE 2002 Conference*, Scottsdale, AZ., 2002.

[30] H. Stumpf and E. Klieme, *Sex-related differences in spatial ability: more evidence for convergence.*, Perceptual and Motor Skills, 69 (1989), pp. 915-921.

[31] D. Voyer, S. Voyer and M. P. Bryden, *Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables.*, Psychological Bulletin, 117 (1995), pp. 250-270.

[32] E. Coluccia and G. Louse, *Gender differences in spatial orientation: A review*, Journal of Environmental Psychology, 24 (2004), pp. 329-340.

[33] R. S. Kennedy, L. J. Hettinger and M. G. Lilienthal, *Simulator Sickness*, in G. H. Crampton, ed., *Motion and Space Sickness*, CRC Press, Boca Raton, FL, 1988.

[34] P. Graf and L. Ryan, *Transfer-appropriate processing for implicit and explicit memory.*, Journal of Experimental Psychology: Learning, Memory, and Cognition, 16 (1990), pp. 978-992.

[35] P. N. Johnson-Laird, *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.*, Harvard University Press, Cambridge, MA., 1983.

[36] M. L. Ryan, *Cyberspace, Virtuality, and the Text. In Cyberspace Textuality: Computer Technology and Literary Theory*, Indiana University Press, Bloomington, ID., 1999.

[37] C. Botella, R. M. Baños and M. Alcañiz, *A psychological approach to presence*, *6th International Workshop on Presence*, Aalborg, Denmark, 2003.

[38] M. Slater and A. Steed, *A virtual presence counter*, Presence: Teleoperators and Virtual Environments., 9 (2000), pp. 413-434.

[39] M. Meehan, B. Insko, M. Whitton and F. P. Brooks, *Physiological measures of presence in stressful virtual environments.*, *SIGGRAPH 2002*, San Antonio, USA, 2002, pp. 645-652.

[40] M. Slater and S. Wilbur, *A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments.*, Presence: Teleoperators and Virtual Environments., 6 (1997), pp. 603-616.

[41] C. Sas and G. M. P. O'Hare, *The presence equation: An investigation into cognitive factors which underlie presence within non-immersive virtual environments.*, *The 4th Annual International Workshop on Presence*, Philadelphia, PA, 2001.

[42] W. Barfield and C. Hendrix, *The effect of update rate on the sense of presence within virtual environments.*, Virtual Reality: The Journal of the Virtual Reality Society, 1 (1995), pp. 3-16.

Understanding Instant Messaging: Gratifications and Social Presence

Ha Sung Hwang, Matthew Lombard
Hanyang University, Temple University
{hhwang@temple.edu, lombard@temple.edu}

Abstract

This paper explores the Instant Messaging phenomenon within the college sector—the first generation to grow up with the Internet. Drawing on the uses and gratifications approach and social presence theory, a survey of 443 IM users reveals five gratifications sought and obtained: social utility, interpersonal utility, convenience, entertainment/relaxation and information. The results also demonstrate the important role of social presence in IM use. The value of integrating social presence into the uses and gratifications paradigm and other theoretical and practical implications are discussed.

Keywords--- **Instant messaging; social presence; uses and gratifications**.

1. Introduction

Instant Messaging (IM) has become a popular mode of communication for people with access to the Internet. Some people do not use it and never will, but for a rapidly growing number of people IM is a useful communication tool, and for some it is a vital part of their lives. The telephone is no longer necessary for a person to be connected constantly to his or her family and friends. One can simply turn on their computer and log onto IM and hold simultaneous conversations, without long distance fees.

Although the IM phenomenon has received much media attention, the current IM literature is limited. Most studies narrowly focus on IM use in work-related activities within the business community (e.g., [1], [2]; [3]). Even fewer studies have addressed the first generation that grew up with the Internet and the largest group of IM users: college students. This study addresses this limitation by exploring the potential predictors of IM use within the collegiate sector. Guided by two sets of theory, uses and gratifications and social presence theory, this study specifically (1) examines the gratifications sought (GS) and obtained (GO) by college student IM users, and (2) examines the effects and role of social presence in IM use. Consequently, the study aims to extend uses and gratifications and social presence theories.

2. Instant Messaging

Instant messaging is software that allows computer users to send and receive short text messages in real time. Unlike online "chatting," a group activity, instant messages are exchanged between only two users (who can both see what is being typed). All one needs to join IM is a computer and an Internet connection. Most IM programs (the biggest at this writing are AOL Instant Messenger, MSN Messenger and Yahoo! Messenger) are free and easy to download and install.

3. Uses and Gratifications and Instant Messaging

Research about the Internet has utilized the uses and gratifications approach to examine the motivations behind the use of the Internet in general and different online activities such as the World Wide Web, electronic bulletin boards, and email. Studies of motivations for Internet use in general find that, as with television, audiences use it to satisfy needs such as entertainment, escape (or passing time), information seeking, and social interaction (e.g., [4], [5]). In addition, people use particular Internet services like email (e.g., [6]) and electronic bulletin boards (e.g., [7]) to satisfy the need for maintaining relationships and/or social interaction.

Despite the rapid and widespread diffusion of Instant Messaging systems there is little research that explores why people use IM. One study examined the motivations for chatting on ICQ, the first IM application [8] and found motives of relaxation, entertainment, fashion, affection, sociability and escape. Other studies on IM suggest that teenagers primarily use IM to increase socializing opportunities [9].

Of equal importance is understanding what benefits audience members obtain from IM use. Although traditional media are not generally reported to fulfill users' needs fully, one study ([10]) suggests that new media do successfully fulfill audience members' needs. He found that the majority of people sought increased communication as well as information gathering from email and felt that these needs had been gratified by the service. The study also reported that respondents sought both entertainment and information from the Web and felt that they had obtained both gratifications.

The limited literature on IM use in the workplace suggests IM serves as an effective communication tool for geographically distributed work teams. IM also plays a key role for teenagers in communicating with friends [11]. Although college students constitute a vast majority of IM users, few studies have examined why college students use IM and what they get out of it.

Based on the literature discussed above, the following research question is proposed.

RQ 1. What gratifications are college students seeking and obtaining from using IM? S*pecifically, what is the relationship between gratifications sought (GS) and gratifications obtained (GO) from IM use?*

4. Social presence and Instant Messaging

Social presence is a "sense of being with another in a mediated environment" ([12], p. 14). Researchers have examined social presence in a variety of mediated communication contexts, from low bandwidth interactive text technologies (e.g., email, bulletin board) to high bandwidth audio-visual technologies (e.g., video and computer teleconferencing systems).

Although it was initially studied in relation to such traditional media as audio and closed-circuit television (as well as face-to face communication), social presence is increasingly being acknowledged as an important factor for understanding the effects of new media. Studies suggest that interactive virtual environments created by some Internet technologies evoke social presence. For instance, [13] examined the users' experience of "being with others" in text-based virtual environment (e.g., MUDs). They surveyed 207 users and 69% reported that they felt a sense of presence with others when they use this technology.

[14] examined the relationship between social presence and various types of CMC. Email was perceived to possess the highest level of social presence, followed by real-time discussion and bulletin board. Such different degrees of impact on social presence, he insisted, "not only come from the attributes of CMC systems, but also the uses and various perceptions of CMC systems" (p. 21).

Other studies suggest that social presence affects the outcomes of CMC use. Some find that social presence is a significant predictor of the user's satisfaction with interactive television classes [15] and text-based computer conferences [16].

Such findings have implications for the current study. Despite the very limited ability to transmit social cues in short text messages, it is possible that Instant Messaging creates social presence, as email and bulletin boards do, and that this feeling of social presence can lead to heavy usage of and satisfaction with IM.

Based on the literature discussed, the present study attempts to answer the following research question:

RQ 2. What is the role of social presence in IM use? Specifically, how is students' experience of social presence in IM use related to gratifications sought (GS) and obtained (GO) from IM use?

Finally, the present study aims to investigate how three important concepts—gratifications sought (GS), gratifications obtained (GO), and social presence are linked to IM use.

RQ 3. How do gratifications sought (GS), gratifications obtained (GO), and social presence predict IM use?

5. Method

An online survey of college students was conducted to address the research questions.

5.1. Sampling

The majority of respondents were drawn from undergraduate communication classes at three universities in the United States (N=508). A total of 602 surveys were completed. Four hundred and forty-three respondents were IM users and 159 were non-users. Among IM users, 45.5% (N=200) were males and 51.3% (N=225) were females. The age of these respondents ranged from 18 to 39 years (*M*=20.26), with three-fourths of them between 18 and 24 years. The sample included freshmen (43.5%, N=188), sophomores (24.1%, N=104), juniors (21.1%, N=91), seniors (8.3%, N=36), and graduate students (3%, N=13).

5.2. Measurement

5.2.1. IM Use behavior. To assess IM use behavior respondents were asked to report (1) how long they have been using IM, (2) how many minutes they use IM in a typical day, (3) how many days they use IM in a week, (4) the average number of minutes they spend on an IM session, and (5) how many IM sessions they have in a typical day.

5.2.2. Gratifications sought (GS) from IM use. To measure motivations or reasons why college students use Instant Messaging, selected motivation items used in previous research on new media such as the Internet [17], e-mail [6] and ICQ [8] were adapted and modified. Additional items from a pilot study were also included in the survey. The final questionnaire consisted of 28 motivation statements (see Table 1). Respondents were asked to state their levels of agreement with the statements on a 7-point Likert scale (1= strongly disagree, 7=strongly agree).

5.2.3. Gratifications obtained (GO) from IM use. To measure gratifications obtained from IM use, the present study followed [18], which distinguish between the measurement of gratifications sought and obtained. In their study, gratifications obtained (GO) were measured after gratifications sought (GS). The same items used to measure GS were used to measure GO, but were slightly reworded. The present study followed this approach. For instance, the first statement for GS, "I use IM to keep up with news" was reworded to "IM actually helps me keep up with news" to measure GO. The same 28 items that were used to measure GS were included in the survey questionnaire (see Table 2). Respondents were asked to reply using the same 7-point scale employed to measure gratifications sought (1=strongly disagree, 7=strongly agree).

5.2.4. Social presence. To assess the feeling of social presence through IM use the study used 7 items selected from previous studies ([19], [13]) and modified for this context. Respondents were asked to report their levels of agreement with 7 statements on a 7-point Likert scale (1=strongly disagree, 7=strongly agree) (see Table 3).

Table 1. Means, Standard Deviations, and Confirmatory Factor Analysis Results for Gratifications Sought from IM use (N=443)

GS items	M	SD	Load-ings
Social Utility	5.45	.98	
▪ I use IM to see what others are up to	5.84	1.28	.76
▪ I use IM to exchange information with people I know	5.84	1.34	.74
▪ I use IM to pass information on to other people	5.27	1.45	.72
▪ I use IM to keep in touch with friends or family members	6.18	1.36	.70
▪ I use IM to keep in touch with friends or relatives who live far away	5.88	1.51	.64
▪ I use IM to let others know I am concerned about them	4.35	1.59	.60
▪ I use IM to feel involved with what's going on with other people	4.78	1.65	.56
Cronbach's alpha			.80
Eigenvalue			3.23
Interpersonal Utility	3.53	1.24	
▪ I use IM because I need someone to talk to or be with	3.84	1.72	.85
▪ I use IM to feel less lonely	4.00	1.79	.82
▪ I use IM to be reassured that someone is there	3.42	1.88	.77
▪ I use IM to get interesting things to talk about	3.55	1.78	.71
▪ I use IM to avoid going out	2.46	1.74	.57
▪ I use IM because it's like fact-to-face conversation	3.54	1.85	.52
Cronbach's alpha			.80
Eigenvalue			3.08
Convenience	5.46	1.08	
▪ I use IM because it's convenient	5.87	1.17	.85
▪ I use IM because it's fast	5.91	1.27	.79
▪ I use IM because it's simple and easy	5.66	1.26	.77
▪ I use IM because it's easier than email	5.51	1.51	.76
▪ I use IM to talk to many people at the same time	5.42	1.77	.68
▪ I use IM because it's easier than making a phone call	4.98	1.73	.65
▪ I use IM to save money without long distance fees	4.86	2.05	.50
Cronbach's alpha			.82
Eigenvalue			3.65
Entertainment/ Relaxation	4.40	1.16	
▪ I use IM because it's entertaining	4.94	1.44	.83
▪ I use IM because it's fun	4.85	1.46	.81
▪ I use IM to pass time when I am bored	5.44	1.55	.71
▪ I use IM to forget about other things	3.18	1.83	.64
▪ I use IM because it relaxes me	3.55	1.58	.63
Cronbach's alpha			.78
Eigenvalue			2.98
Information	3.09	1.25	
▪ I use IM to get information I am looking for	3.08	1.72	.79
▪ I use IM to keep up with the news	2.52	1.62	.73
▪ I use IM to express my personal feelings and opinions freely	3.42	1.88	.68
Cronbach's alpha			.68
Eigenvalue			1.63

Note: Response options ranged from strongly disagree (1) to strongly agree (7).

6. Results

6.1. IM Use Behaviors

The college students who use IM had been doing so for an average of five and a third years (range: 1 - 12, M=5.29, SD=2.48). The majority of respondents (59.7 %, N=258) reported that they use IM daily, 17.4% (N=75) reported using it 5-6 days a week, 10% (N=43) reported 3-4 days, 9.3% (N=40) reported 2-3 days and only 3.7% (N=16) reported using IM once a week.

Students in this sample reported that on a typical day they spend an average of one hour and 40 minutes (SD=118.35 min.) using IM, but responses ranged from five minutes to 16 hours. When asked how many times their IM window pops up on a typical day, students reported that they had from one to 80 IM sessions with an average of 7.78 sessions (SD=9.73). Respondents also reported that they spent an average of 30.06 minutes (SD=47.34) on a typical IM session, but responses ranged from 2 minutes to 7 hours and 17 minutes.

6.2. Construction of Factors for GS and GO

As described above, the same 28 items were used to measure gratifications sought and obtained (although the 28 items used to measure GS were slightly reworded to measure GO). Therefore, it was necessary to match GS factors with their corresponding GO factors. To accomplish this goal, first an exploratory principle component factor analysis (with varimax rotation) for GS was conducted. Confirmatory factor analysis for GO was then conducted to fit the GS factor analysis results. However, this approach failed to produce reliable factors for both GS and GO. Therefore, the reverse approach was used—that is, after an exploratory principle component factor analysis (with varimax rotation) for GO items, confirmatory factor analyses for GS were conducted.

The following criteria were applied for the selection of GO factors: (1) An eigenvalue of 1.0 or greater was a necessary condition for each factor and (2) each factor had to contain at least two items meeting a 60/40 criterion –i.e., two primary factor loadings of at least .60 with no cross loadings over .40 on any other factor. The analysis yielded five factors. The index based on each GO (and corresponding GS) factor was then tested with Cronbach's alpha to ensure inter-item or scale reliability. (The resulting GO and GS indices are presented in the next section.)

6.3. Gratifications and IM use

RQ 1. What gratifications are college students seeking and obtaining from using IM? Specifically, *what is the relationship between gratifications sought (GS) and gratifications obtained (GO) from IM use?*

The first research question concerns identifying GS (motives) and GO (benefits) of Instant Messaging and relationships between GS and GO. As described above, an exploratory principle component factor analysis with varimax rotation was conducted to determine GO factors. The analysis accounted for 60% of total variance and the results are summarized in Table 2.

Table 2. Means, Standard Deviations and Factor Analysis Results for Gratifications Obtained from IM use (N=443)

GO items	M	SD	Load-ings
Social Utility	5.34	1.00	
▪ IM actually helps me to keep in touch with friends or family members	5.88	1.37	.74
▪ IM actually helps me to keep in touch with friends or relatives who live far away	5.50	1.27	.63
▪ IM actually helps me to exchange information with people I know	5.42	1.47	.62
▪ IM actually helps me see what others are up to	5.14	1.26	.61
▪ IM actually helps me to feel involved with what's going on with other people	4.92	1.52	.60
▪ IM actually helps me pass information on to other people	5.30	1.39	.58
▪ IM actually helps me let others know I am concerned about them	4.65	1.51	.56
Cronbach's alpha			.88
Eigenvalue			4.00
Interpersonal Utility	3.87	1.27	
▪ IM actually helps me to feel less lonely	3.38	1.83	.78
▪ IM actually helps me to be reassured that someone is there	4.20	1.65	.76
▪ IM actually helps me to have someone to talk to or be with	4.39	1.63	.74
▪ IM actually helps me to avoid going out	3.31	1.77	.65
▪ IM is actually like fact-to-face conversation	3.56	1.87	.63
▪ IM actually helps me to get interesting thing to talk about	4.05	1.73	.53
Cronbach's alpha			.83
Eigenvalue			3.90
Convenience	5.64	1.04	
▪ IM is actually easier than email	5.47	1.49	.78
▪ IM is actually easier than making a phone call	4.97	1.75	.75
▪ IM is actually fast	6.03	1.19	.73
▪ IM is actually convenient	6.00	1.23	.60
▪ IM is actually simple and easy	6.01	1.22	.57
▪ IM actually helps me to save money without long distance fees	5.26	1.76	.55
▪ IM actually helps me to talk to many people at the same time	5.72	1.48	.53
Cronbach's alpha			.82
Eigenvalue			3.49
Entertainment/ Relaxation	5.37	1.26	
▪ IM actually helps me to pass time when I am bored	5.43	1.55	.70
▪ IM actually helps me to be entertained	4.77	1.54	.68
▪ IM actually helps me to have fun	4.32	1.64	.61
▪ IM actually helps me to forget about other things	3.74	1.82	.63
▪ IM actually helps me to relax	3.54	1.66	.50
Cronbach's alpha			.83
Eigenvalue			3.23
Information	3.51	1.29	
▪ IM actually helps me to keep up with then news	3.42	1.88	.76
▪ IM actually helps me to get information I am looking for	3.38	1.83	.72
▪ IM actually helps me to express personal feelings and opinions freely	3.78	1.79	.62
Cronbach's alpha			.67
Eigenvalue			2.40

Note: Response options ranged from strongly disagree (1) to strongly agree (7).

As indicated in Table 2, the five GO factors suggest that IM functions to fulfill needs for 1) "social utility"—a need for a sense of community, 2) "interpersonal utility"—a need for individual connection, 3) convenience, 4) entertainment and relaxation, and 5) information.

As indicated in Table 3, all GO factors are positively and significantly correlated with all GS factors. The strongest correlations are between corresponding gratifications sought and obtained, particularly for convenience and interpersonal gratifications (r=.83, p<.01 in both cases).

Table 3. Bivariate Correlations between GS factors and GO factors

GS \ GO	Social utility	Interpersonal utility	Convenience	Entertainment/ Relaxation	Information
Social utility	.78** (.73**)	.24** (.02)	.59** (.54**)	.39** (.26**)	.38** (.28**)
Interpersonal utility	.35** (.13**)	.83** (.77**)	.18** (.04)	.53** (.41**)	.51** (.27**)
Convenience	.64** (.56**)	.26** (.08**)	.83** (.81**)	.46** (.36**)	.37** (.27**)
Entertainment/ Relaxation	.45** (.32**)	.46** (.32**)	.46** (.40**)	.81** (.78**)	.45** (.36**)
Information	.41** (.31**)	.42** (.31**)	.28** (.20**)	.45** (.36**)	.77** (.73**)

Note: **p<.01 (2-tailed). Partial Correlations between GS and GO, controlling for social presence, appear in parentheses.

6.4. Social Presence and IM Use

RQ 2. What is the role of social presence in IM use? Specifically, how is students' experience of social presence in IM use related to gratifications sought (GS) and obtained (GO) from IM use?

Means for the social presence measures indicated that the students did experience social presence while using instant messaging (see Table 4).

Table 4. Means and Standard deviations for Social Presence items and index

Items	M	SD
Social presence index	4.42	1.20
▪ I often smile in response to the IM messages that the other person sends in an IM session	5.60	1.33
▪ I often make a sound out loud (e.g., laugh or speak) in response to IM message that the other person sends in an IM session	5.36	1.61
▪ IM messages express feeling and emotion	4.21	1.67
▪ I feel a senses of actually being together with the person I am communicating with when IM am using IM	4.17	1.66
▪ I feel emotionally connected with the person I am communicating with during IM use	4.13	1.59
▪ I feel that I am present with others and the others are present with me during IM use	3.96	1.61
▪ During IM use I feel as if I and the other person are located in the same room	3.52	1.69

Note: Response options ranged from strongly disagree (1) to strongly agree (7).

Pearson correlations indicate that all GS factors were positively and significantly related to social presence (see Table 5). When the college students used IM for social utility, interpersonal utility, convenience, entertainment/relaxation, and information reasons, they experienced social presence. The strongest correlations were between social presence and using IM for interpersonal utility needs (r= .48, p<.01), between social presence and social utility needs (r=41, p<.01), and between social presence and entertainment/relaxation needs (r=.38, p<.01). These findings suggest that social and interpersonal motives for using IM are highly related to sense of social presence. In addition, respondents in the sample indicated that they felt a high degree of social presence when they used IM for entertainment/relaxation needs.

Table 5. Pearson Correlations between GS, GO and Social presence

Gratifications	Correlations	
	GS VS. Social Presence	GO VS. Social Presence
Social Utility	.41**	.49**
Interpersonal Utility	.48**	.54**
Convenience	.37**	.32**
Entertainment/Relaxation	.38**	.44**
Information	.32**	.37**

**p<.01 (2-tailed)

Social presence was also positively and significantly related to gratifications obtained from IM use (see Table 5). Correlations revealed significant relationships between social presence and all GO factors. The strongest correlations were between social presence and GO for interpersonal utility (r=.54, p<.01), and between social presence and GO for social utility (r=.49, p<.01). Although significant relationships were reported, GO for information (r=.37, p<.01) and GO for convenience (r=.32, p<.01) were less related to social presence than the other GO factors.

It is interesting to note that IM users who either seek or obtain convenience and information gratifications from IM use perceived less social presence, while using IM in order to seek social, interpersonal, and diversion entertainment gratifications and obtaining these gratifications were linked to higher levels of social presence.

To investigate the mediating role of social presence in the relation between gratifications sought and obtained from IM use, partial correlations were calculated (see Table 4). When social presence was statistically controlled, all correlations between GS factors and GO factors were reduced to some degree. The correlations that were most reduced were the correlation between GS for social utility and GO for interpersonal utility (from .24 ($p<.01$) to .02 (n.s.)) and the correlation between GS for interpersonal utility and GO for convenience (from .18 ($p<.01$) to .04 (n.s.)). Overall, these findings suggest that social presence plays a significant role in the relationships between gratifications sought and obtained from IM.

RQ 3. How do gratifications sought (GS), gratifications obtained (GO), and social presence predict IM use?

To investigate how the overall concepts of GS, GO, and social presence predict IM use, a path analysis was conducted (using Amos software). The model that best fit the data is presented in Figure 1; the relatively good fit is indicated by the chi-square (3.26), p-value (.19), GFI (.998) and RMSEA (.038). The model shows a strong relationship between GS and IM use (ß=.71), which suggests that GS has a direct and significant effect on IM use. On the other hand, GO has an indirect effect on IM use, an effect mediated by gratifications sought; in this model GO directly influences GS rather than IM use.

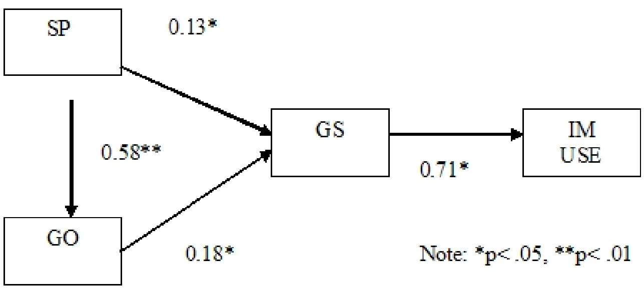


Figure 1. Path model of relationships among GS, GO, and social presence in predicting IM use

The role of social presence in the model is especially interesting. Social presence has indirect effects on IM use via GS and GO. But social presence had a more direct impact on GO (ß=.58) than GS (ß=.13). Again, being influenced by social presence, GO directly affects GS and GS directly affects IM use.

Overall, it can be said that GS was the most important variable that directly influenced IM use, while GO and social presence indirectly affected IM use.

7. Discussion

The first objective of this study was to examine the gratifications sought and obtained by IM users among college students. The study found five salient reasons for using IM among this largest segment of users: social utility, interpersonal utility, convenience, entertainment-relaxation, and information. It was also found that IM gratified those needs successfully.

The second objective of this study was to examine the role and effects of social presence in IM use. The results indicate that even with its low bandwidth, text-only format, IM evokes a sense of social presence—i.e., of "being together" or "emotional connectedness"—and has an important impact on what college students seek and obtain from IM use.

Although all sought and obtained gratifications were positively linked to social presence, the study found that interpersonal utility was the gratification most related to the experience of social presence. College students who use IM for seeking someone to talk to or be with, for being reassured that someone is there, for avoiding loneliness, and/or for getting interesting things to talk about with others perceive to a greater extent that IM messages convey feeling and emotion and feel more of a sense of "being together" and "emotional connectedness" (as if "they are located in the same room") than those who use IM for seeking other gratifications. Consistent with this, the more college students obtain the interpersonal utility benefits from IM use, the more they feel social presence in comparison to other gratifications.

While IM users seeking interpersonal utility gratifications may be more receptive to feeling a sense of "being together" during IM use, it seems clear that IM's ability to allow users to hold real time conversations through exchanging short messages with people they know, along with the ability to use emoticons to indicate emotional states, encourages a sense of "being together" among many IM users, despite the limitations of a text-based medium.

The important role of social presence is also seen in its effect on the relationships between GS and GO. Partial correlations between GS factors and GO factors with social presence controlled show that all correlations between GS and GO were reduced to some degree. And in the path analysis, while social presence was an insignificant direct predictor of IM use, with its great influence on GO, social presence turns out to be an important variable in the relation between GS and IM use. In other words, the more college students feel a sense of "being together" or "emotional connectedness" when chatting with another through IM, the more they perceive that IM gratifies their needs, and the more college students feel IM gratifies their needs, the more motivations for their IM use, and so IM use itself, increase.

The findings of the present study have several important implications. First, the results provide empirical evidence that IM users feel a sense of "being together" or "emotional connectedness" when chatting with another person through IM, a "lean" or low bandwidth channel of communication. This counter-intuitive result suggests the need to better understand social presence and its role in CMC environments. Even when they acknowledge the centrality of the user's perceptions previous CMC studies have followed pioneers Short et al. ([20]) and examined social presence as a characteristic of a medium, while the measures used here emphasize the user's subjective media experience. If supported by future studies, this new conceptualization can in turn help us develop more valid and standardized social presence measures.

Second, the social presence results provide important extensions of the uses and gratifications approach. The demonstrated key role of this new variable in the motivations for and benefits from IM use suggests that uses and gratifications researchers should consider the concept of social presence as an important variable in explaining what audience members do with a variety of media. Future empirical inquiry is required to establish a theoretical model that integrates the concept of social presence into the uses and gratifications paradigm.

Finally, the results suggest that the designers and marketers of IM systems should take into consideration the role of social presence in IM uses and gratifications. By using the concept of social presence as a marketing strategy and in developing new features (e.g., personalized "buddy" icons), the IM industry may be able to encourage more college students, and perhaps others, to use IM to fulfill social and interpersonal needs.

Two limitations in this study should be noted. First, respondents were all college students, most majoring in communication, from three large urban universities in the United States. Clearly, they do not represent the typical home IM user. Therefore, the research results cannot be generalized to the wider potential home-based IM user population. Second, the use of non-probabilistic sampling for universities chosen and self-selection of students taking the survey are acknowledged as restricting the generalizability of this study. However, the results constitute a first step toward a better understanding of the reasons for, the nature of, and the benefits produced by the use of the increasingly important new medium of IM.

References

[1] Hansen, K., & Damm, C. (2002). *Instant collaboration: Using context-aware instant messaging for session management in distributed collaboration tools.* Paper presented at NordiCHI, October 19-23, 2002, Denmark.

[2] Herbsleb, J. D., Atkins, D. J., Boyer, D.G., Handel, M., & Finholt, T. A. (2002). Introducing Instant Messaging and Chat in the workplace. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp.171-178). New York: ACM Press.

[3] Nardi, A. B., Whittaker, S. & Bradner, E (2000). Interaction and outeraction: Instant messaging in action, *Proceedings of CSCW' 2000* (pp. 79-88). New York: ACM Press.

[4] Ferguson, D. A. & Perse, M. E. (2002). The World Wide Web as a functional alternative to television. *Journal of Broadcasting & Electronic Media, 44*(2), 155-174.

[5] Kaye, B. K. (1998). Uses and Gratifications of the World Wide Web: From Couch Potato to Web Potato. *New Jersey Journal of Communication, 6*(1), 21-40.

[6] Stafford, L., Kline, L. S., & Dimmick, J. (1999). Home E-mail: Relational Maintenance and Gratification Opportunities, *Journal of Broadcasting & Electronic Media, 4*3(4), 659-669.

[7] James, M. L. (1992). *An exploratory study of the perceived benefits of electronic bulletin board use and their impact on other communication activities.* Unpublished Doctoral Dissertation. Florida State University.

[8] Leung, L. (2001). College student motives for chatting on ICQ, *New Media & Society, 3*(4), 483-500.

[9] Grinter, E. R. & Palen, L. (2002). Instant messaging in teen life. In *Proceedings of the 2002 ACM conference on Computer Supported Cooperative Work* (pp. 21-30). New York: ACM Press.

[10] Hunter, C. D. (1996). *The uses and gratifications of Project Agora.* Unpublished undergraduate thesis. Boston College. [Electronic Version] Retrieved May 2, 2004, from http://www.asc.upenn.edu/usr/chunter/agora_uses/

[11] Eldridge, M. & Grinter, R. (2001). *Studying text messaging in teenagers.* Retrieved May 2, 2004 from http://socio.ch/mobile/index_mobile.htm

[12] Biocca, F., Harms, C., Burgoon, J. (2003). Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators and Virtual Environments, 12*(5), 456-480.

[13] Towell, J. & Towell, E. (1997). Presence in text-based networked virtual environments or "MUDS." *Presence: Teleoperators and Virtual Environments, 6*(5), 590-595.

[14] Tu, C. H. (2002). The impact of text-based CMC on online social presence. *The Journal of Interactive Online Learning, 1*(2). Retrieved July 13, 2004 from http://www.ncolr.org

[15] Hackman, M. Z., & Walker, K. B. (1990). Instructional communication in the televised classroom: The effects of system design and teacher immediacy on student learning and satisfaction. *Communication Education, 39*(3), 196-206.

[16] Gunawardena, C. N, & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *The American Journal of Distance Education, 11*(3), 8-26.

[17] Papacharissi, Z. & Rubin, A. M. (2000). Predictors of Internet Use. *Journal of Broadcasting and Electronic Media, 44*, 175-196.

[18] Palmgreen, P., Wenner, L. A., & Rayburn, J. D., II (1980). Relations between gratifications sought and obtained: A study of television news. *Communication Research, 7*, 161-192.

[19] Lombard, M., & Ditton, T. B. (1997). At the heart of it all: The concept of Presence. *Journal of Computer-Mediated Communication, 3*(2). Retrieved March 10, 2004 from http://www.ascusc.org.jcmc/vo13/issue2/

[20] Short, J.A., Williams, E, & Christie, B. (1976). *The social psychology of telecommunications.* London: John Wiley & Sons, Ltd.

Increasing the Motion of Users in Photo-realistic Virtual Environments by Utilising Auditory Rendering of the Environment and Ego-motion

Rolf Nordahl
Medialogy, Aalborg University Copenhagen
2750 Ballerup, Denmark
{rn@media.aau.dk}

Abstract

An occurring problem of the image-based-rendering technology for Virtual Environments has been that subjects in general showed very little movement of head and body. Our hypothesis is that the movement rate could be enhanced by introducing the auditory modality. In the study described in this paper, 126 subjects participated in a between-subjects experiment involving six different experimental conditions, including both uni- and bi-modal stimuli (auditory and visual). The aim of the study was to investigate the influence of auditory rendering in stimulating and enhancing subjects' motion in virtual reality.

The auditory stimuli consisted of several combinations of auditory feedback, including static sound sources as well as self-induced sounds. Results show that motion in virtual reality is significantly enhanced when moving sound sources and sound of ego-motion are rendered in the environment.

1. Introduction

Although sound is one of the fundamental modalities in the human perceptual system, it is still a largely undiscovered area to researchers and practitioners of Virtual Reality. While research has provided insights into aspects of how the multimodal nature of the human may consist, many questions still remain in how one can utilize e.g. audio-visual phenomena when building new media products.

In previous research, we investigated if self-sound enhanced sense of presence in virtual reality [1]. These environments were based on the Image Based Rendering technique [13]. While the users were exposed to photorealistic virtual environments, where they could explore the environment within the constraints of the Region of Exploration (with a radius of 60 cm), no temporal information was delivered. Such virtual environments could be suggested to be a good example of the current state of Presence and VR research, where the feeling of Presence is considered by some to be mostly linked to the spatial domain [14]. However, we believe that such feeling of Presence may be heightened by adding temporal or spatio-temporal information, e.g., created by auditory feedback.

In order to examine this assumption we previously designed parametric sound effect [2], controlled by the motion of a user in the virtual environment. This was done using sound synthesis by physical models. Such physical models described the sonic impact between the user wearing a shoe and the different surfaces that were designed to accommodate the visual stimulus. These models were driven by a footstep controller which the user was asked to wear before entering the VR experience. The addition of sound effects created some interesting dynamic variations in the otherwise static environment, and the experiments showed how interactive sounds enhanced the sense of presence in VR [1].

In the previous simulated virtual environment no other sounds but the one generated by the footsteps controller and synthesizer were present. This was done because the visual feedback proposed to the subjects was an empty technical museum in Prague, where no sounds other than the noise of a fan could be heard. Drawing a parallel to current film sound theory we therefore assumed that introducing other sounds, would act as non-diegetic [12], i.e., sounds which perform as something outside the "reality" of the virtual world, and therefore conflict with the overall goal of creating a close-to-reality experience.

We observed that subjects appreciated the quality of the images reproduced by image-based rendering techniques. However, they were not very stimulated to visit the environment, since no dynamic events were happening. We are therefore interested in investigating if by adding auditory feedback, and therefore a temporal dimension, it is possible to enhance the interest of subjects to such a degree that they actively investigate and explore the environment.

In particular, we extended our previous research by adding environmental sounds as well as simulation of moving sound sources to the virtual reality experience. The sources of such sounds, which might be located out of the subject's field of view, are diegetic, i.e., belonging to the constructed reality and therefore sustain the experience. Our goal is to understand if enhancing the environment with dynamic sounds increases the motion of subjects in virtual environments. Our hope is that sound creates dynamic variations which make the environment more interesting to visit and investigate.

In order to achieve this goal, we decided to monitor the motion of subjects in the virtual environment. We believe that a higher degree of motion represents a higher degree of engagement and interest in exploration of the simulation.

2. Experiment design

2.1 PC and peripherals for Visual Delivery

In the following section the hardware and software configuration used in the experiments is described. The visual stimulus was provided by a standard PC running Suse Linux 10. This computer was running the BENOGO software using the REX disc "Prague Botanical Garden".

The Head-Mounted-Display used was a VRLogic V8[2]. It features Dual 1.3" diagonal Active Matrix Liquid Crystal Displays with resolution per eye: ((640x3)x480), (921,600 color elements) equivalent to 307,200 triads. Furthermore the HMD provides a field of view of 60° diagonal.

The tracker used was a Polhemus IsoTrak II[3]. It provides a latency of 20 milliseconds with a refresh rate of 60 Hz.

2.2 PC and peripherals for audio delivery

The audio system was created using a standard PC running MS Windows XP SP 2. All sound was run through Max/MSP 4.5.1[4] and as output module a Fireface 800 from RME[5] was used. Sound was delivered by eight Dynaudio BM5A speakers[6].

2.3 Audio-visual setup

In the laboratory eight speakers were positioned in a parallelepipedal configuration. Current commercially available sound delivery methods are based on sound reproduction in the horizontal plane. However we decided to deliver sounds in eight speakers and thereby implementing full 3D capabilities. By using this methodology we were allowed to position both static sound elements as well as dynamic sound sources linked to the position of the subject. Moreover we were able to maintain a similar configuration to other virtual reality facilities such as CAVEs, where eight channels surround is presently implemented. This is the reason why 8-channels sound rendering was chosen compared to e.g. binaural rendering.

As described two computers were installed in the laboratory, one running the visual feedback described in the following section, and one running the auditory feedback. A Polhemus tracker, attached to the head mounted display, was connected to the computer running the visual display, and allowed to track the position and orientation of the user in 3D. The computer running the visual display was connected to the computer running the auditory display by TCP/IP. Connected to the sound computer there was the interface RME Fireface 800 which allowed delivering sound to the eight channels, and the wireless shoe controller. The mentioned controller, developed specifically for these experiments [10], allowed detecting the footsteps of the subjects and mapping these to the real-time sound synthesis engine. The different hardware components are connected together as shown in Figure 1.

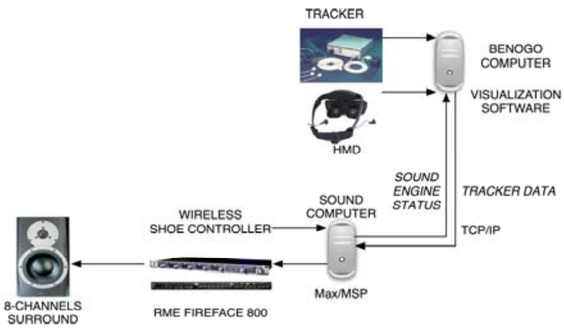


Figure 1. Technical configuration to connect the different hardware components necessary for the experiment

3. Audio-Visual feedback

3.1 Visual feedback

The visual feedback used in these experiments was created under the BENOGO[7] project.

The idea behind this project is the creation of photorealistic visual environments obtained by taking pictures of a specific location at different angles, and building a reconstruction of the same place at the computer using image based rendering techniques.

[2] http://www.vrlogic.com/html/virtualresearch/v8.html (accessed January 17th, 2006)
[3] http://www.inition.co.uk/inition/pdf/isotrak.pdf (accessed January 17th, 2006)
[4] http://www.cycling74.com/ (accessed January 17th, 2006)
[5] http://www.rmeaudio.com/english/firewire/ff800.htm#TECHNICAL (accessed January 17th, 2006)
[6] http://www.dynaudioacoustics.com/Default.asp?Id=3680&AjrNws=663&AjrNwsPg=1(accessed January 17th, 2006)

[7] www.benogo.dk

Figure 2: A view of the botanical garden in Prague reconstructed using image based rendering techniques.

One of the peculiarities of this approach is the fact that no moving objects have to be present in the environment when the pictures are taken, since this would affect the visual reconstruction.

This also implies that the reconstructed scenarios do not vary over time, which means that one could be concerned with that the exposure to the environment becomes tedious and uninteresting for the users to explore. As such we regard the exploration of auditory feedback as a good way to cope with these limitations, as explained in the following section.

3.2 Auditory feedback

Different levels of auditory rendering were implemented, in order to test their effect on stimulating motion of subjects in virtual reality.

A soundscape accompanying the visualization of the Prague botanical garden shown in Figure 2 was designed. Such soundscape contained sounds of birds, water dropping sounds, and other kinds of environmental sounds typical of a botanical garden. The soundscape was delivered using 8-channel surround sound speakers in a damped laboratory via Dynaudio BM5A loudspeakers.

Additionally, moving sound sources such as mosquitoes flying around the room were added, to enhance the sense of immersion in the soundscape.

Furthermore, ego-sounds of the subjects walking around the environment were synthesized in real-time. Such sounds were created by having the users wear a pair of pressure sensitive sandals, connecting to a PC via a wireless interface. Walking around the room, subjects generated in real-time walking sounds synthesized using modal synthesis.

To summarize, four kinds of auditory stimuli were provided to the subjects:

1. “Static” soundscape, reproduced at max. peak of 58dB, measured c-weighted with slow response. This soundscape was delivered through the 8-channels system.
2. Dynamic soundscape with moving sound sources, developed using the VBAP algorithm [9], reproduced at max. peak of 58dB, measured c-weighted with slow response.
3. Auditory simulation of ego-motion, reproduced at 54 dB. (This has been recognised as the proper output level as described in [1])
4. A piece of classic music [11], reproduced at max. peak of 58dB, measured c-weighted with slow response.

Figure 3: A subject during the experiment.

4. Test description

126 subjects took part to the experiment. All subjects reported normal hearing and visual conditions. Figure 3 shows one of the subjects participating to the experiment.

Before entering the room, subjects were asked to wear a head mounted display and the sandals enhanced with sensors. Subjects were not informed about the purpose of the sensors-equipped footwear. Before starting the experimental session the subjects were told that they would enter a photo-realistic environment, where they could move around if they so wished. Furthermore, they were told that afterwards they would have to fill out a questionnaire, where several questions would be focused on what they remember having experienced. No further guidance was given. Each subject was exposed to one of the six conditions shown in Table 1 for 3 minutes. The experimental conditions are described in the following section.

4.1 Experimental Conditions

The experiment was performed as a between-subjects study including the following six conditions:

1. Visual only. This condition had only uni-modal (visual) input.

2. Visual with footstep sounds. In this condition, the subjects had bi-modal perceptual input (audio-visual) comparable to our earlier research [1].

3. Visual with full sound. This condition implies that subjects were treated with full perceptual visual and audio input. This condition included static sound design, 3D sound (using VBAP algorithm [9]) as well as rendering sounds from ego-motion (the subjects triggered sounds via their footsteps).

4. Visual with full sequenced sound. This condition was strongly related to condition 3. However, it was run in three stages: the condition started with bi-modal perceptual input (audio-visual) with static sound design. After 20 seconds, the rendering of the sounds from ego-motion was introduced. After 40 seconds the 3D sound started (in this case the sound of a mosquito).

5. Visual with sound + 3D sound. This condition introduced bi-modal (audio-visual) stimuli to the subjects in the form of static sound design and the inclusion of 3D sound (the VBAP algorithm using the sound of a mosquito as sound source). In this condition no rendering of ego-motion was conducted.

6. Visual with music. In this condition the subjects were introduced to bi-modal stimuli (audio and visual) with the sound being a piece of classical music [11]. This condition was used as a control condition, to ascertain that it was not sound in general that may influence the in- or decreases in motion. Furthermore it enabled us to deduce if the results recorded from other conditions were valid. From this it should be possible to deduct how the specific variable sound design from the other experimental conditions affects the subjects.

Table 1 summarizes the different experimental conditions together with the mean and standard deviation of the age of the participants. The first column of Table 1 represents the name of the condition, in the same order as described above, while the second column outlines the auditory feedback of each condition as described in Section 3.2. As an example, the condition "full" represents the situation in which subjects were exposed to auditory feedback 1,2 and 3, which means a static soundscape, a dynamic soundscape where moving sound sources were introduced, and the sound produced by their footsteps while they were walking around the virtual environment.

CONDITION	AUDITORY STIMULI	NUMBER SUBJECT	MEAN (AGE)	ST.D. (AGE)
Visual	None	21	25.6	4.13
Visuals w. foot	3	21	25.7	3.75
Full	1 + 2 +3	21	25	4.34
Full seq	1 + 2 + 3	21	22.8	2.58
Sound + 3D	1+2	21	22.9	2.5
Music	4	21	28	8.1

Table 1: Description of the seven different conditions to which subjects were exposed during the experiments. The number in the second column refers to the auditory feedback described in Section 3.2.

It is worth reminding that in the fourth condition called full sequenced (abbreviated with "full. Seq."), subjects were exposed to the same auditory stimuli as in the first condition. The difference consists in the order of appearance of such stimuli. For example, the ego-sound generated by the footsteps was introduced only 20 seconds after the subjects started the experiment, whereas in the full condition the ego-sound was present from the beginning of the experiments.

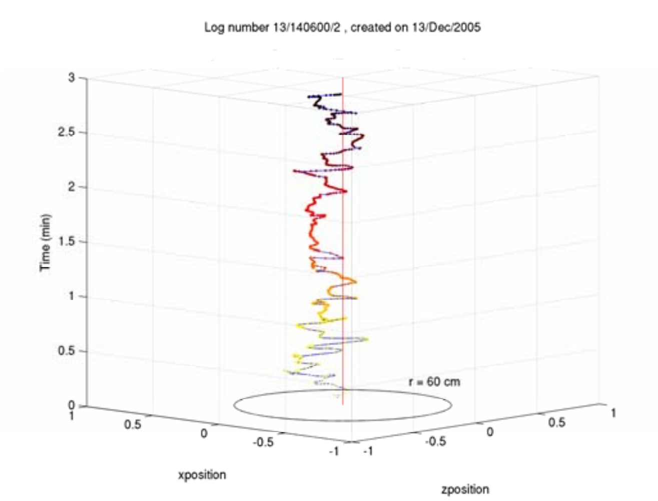


Figure 4: Visualization of the tracker data for one subject exposed to the "visual only" condition.

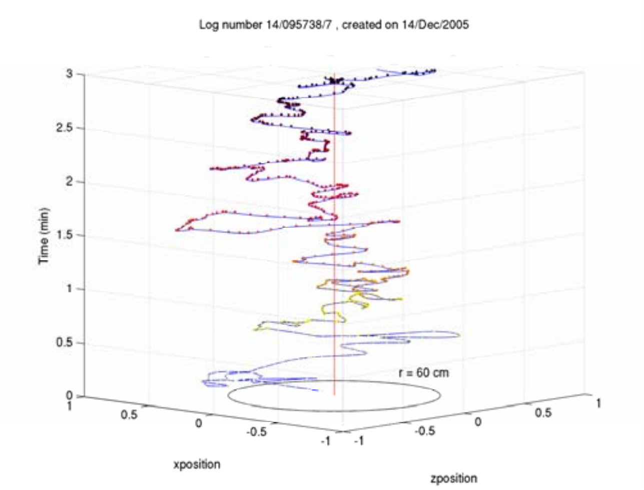


Figure 5: Visualization of the tracker data for one subject exposed to the "full" condition.

Figure 4 shows the visualization of the tracker data over three minutes for one subject exposed to the "Visual only" condition. This visualization was designed to have a way to display in a meaningful way the behavior of the subjects during the experiments.

The circle on the bottom of Figure 4 represents the region of exploration for the visual feedback. Due to the technology used to produce the visual feedback, outside that circle no visual information was present. The vertical line represents the center of the region of exploration. The vertical axis of the figure represents time, while the x and z axis represent the surface where the subjects could walk around. To clarify the visualization, the color information represents the evolution of the motion over time.

We found this visualization rather useful, since it represents clearly quantity of motion of different subjects. Figure 5 shows the same visualization as Figure 4 but in this case for a subject exposed to the condition "Full". Notice how the motion pattern in Figure 4 is significantly less pronounced than in Figure 5.
This is consistent with the test results reported in the following section.

5. Results

Table 2 shows the results obtained by analysing the quantity of motion over time for all subjects for the different conditions. Such analysis was performed by calculating motion over time using the tracker data, where motion was defined as Euclidian distance over time for the motion in 3D.

Our results show that there is clear evidence of the impact of sound in relation to visual only stimuli. As can be seen in Table 2, motion is higher in conditions were the auditory stimuli is "Full" (mean=26.47, st.d=5.6) or "Full sequenced" (mean=25.19, st.d=5.91).

The other conditions show a significant reduction in motion.

Tracked movement	**Full**	**Music**	**Full Seq**	**Visuals only**	**Visuals w. footsteps**	**Sound + 3D**
Mean	26,47	20,95	25,19	21,41	22,82	21,77
Median	26,54	20,79	24,31	21,61	25,66	21,87
st.d.	5,6	6,38	5,91	6,39	6,89	6,74

Table 2: Motion analysis for the different conditions.

The significance of the results is outlined in Table 3. In this table, each condition was tested among the others.

	Full	**Music**	**Full seq.**	**Visual only**	**Visual w. foot**	**Sound + 3D**
Full						
Music	0.003					
Full seq.	0.243	0.018				
Visual Only	0.006	0.41	0.03			
Visual w. foot	0.04	0.197	0.132	0.26		
Sound + 3D	0.011	0.347	0.048	0.431	0.32	

Table 3: Comparison of the motion analysis for the different conditions.

As can be seen from Table 3, there exists a clear connection between the stimuli. First of all it is interesting to notice that the condition of "Music" elicits the lowest amount of movement, even less than the condition "Visual Only". However, the difference between the condition "Visual Only" and "Music" is not significant (p=0.410), which translates into that we cannot state that using sounds not corresponding to the environment (such as music), should diminish the amount of movement. The fact that music shows less movement indicates that it is important which sound is used. The condition "Music" was in fact used as control condition for this very purpose.

Results also show that footsteps sounds alone do not appear to cause a significant enhancement in the motion of the subjects. When comparing the results of the conditions "Visual only" versus "Visuals w. footsteps" (no significant difference) and the conditions "Full" versus "Sound+3D" (significant difference) there is an indication that the sound of footsteps benefits from the addition of environmental sounds.
This result shows that environmental sounds are implicitly necessary in a virtual reality environment and we assume that their inclusion is important to facilitate

the subjects in accepting the faithfulness of the simulation.

6. Conclusions

In this paper we investigated the role of dynamic sounds in enhancing motion in virtual reality.

Results show that 3D sound with moving sound sources and auditory rendering of ego-motion enhance the quantity of motion of subjects visiting the VR environment.

It very interesting to notice that it is not the individual auditory stimulus that affects the increase of motion of the subjects, but rather that it is the combination of soundscapes, 3-dimensional sound and auditory rendering of one's own motion that induces a higher degree of motion.

We are currently extending these results to environments were the visual feedback is more dynamic and interactive, such as computer games and virtual environments reproduced using 3D graphics.

7. References

[1] Nordahl, Rolf, "Design and evaluation of a multi-modal footstep controller with VR-applications", Proceeding of 2nd International Conference on Enactive Interfaces, Genoa, Italy, 2005.

[2] Cook, P.R. Real sound synthesis for interactive applications. AK Peters, 2002.

[3] Shilling, Russel and Krebs, Eric, "Videogame and Entertainment Industry Standard Sound Design Techniques and Architectures for Use in Videogames, Virtual Environments and Training Systems", MOVES Institute, Naval Postgraduate School.

[4] Freeman, J. and Lessiter, J., "Here, There and Everywhere: The Effects of Multichannel Audio on Presence", Proceedings of the 2001 International Conference on Auditory Display, 2001.

[5] Storms, Russel L. and Zyda, Michael J., "Interactions in Perceived Quality of Auditory-Visual Displays", Presence, Vol. 9, No. 6., December 2000.

[6] Shinn-Cunningham, B., "Applications of Virtual Auditory Displays", Proceedings of the 20th international Conference of the IEEE Engineering in Biology and Medicine Society, Vol. 20, No. 3, 1998.

[7] Larson, Pontus et al., "The Actor-Observer Effect in Virtual Reality Presentations", Cyberpsychology and behaviour, Volume 4, Number 2, 2001.

[8] Kahn, James K. et al., " Integrating Sounds and Motions in Virtual Environments", Presence, v. 7, No.. 1, pp. 67-77, February 1998.

[9] Pulkki,V., "Spatial sound generation and perception by amplitude panning techniques". PhD thesis, Helsinki University of Technology, Espoo, Finland, 2001.

[10] Dimitrov, Smilen, "A simple practical approach to wireless distributed data acquisition", 9th semester project on the Medialogy education, Aalborg University Copenhagen, 2005 .

[11] Mozart, Wolfgang Amadeus, "Piano Quintet in E flat, K. 452, 1. Largo Allegro Moderato", Philips "Digitals Classics, 446 236-2, 1987.

[12] Bordwell, David and Thompson, Kristen., "Film Art", McGraw-Hill Companies; 7th edition (June 1, 2003)

[13] Bakstein, Hynek and Pajdla, Tomas, "Visual Fidelity of Image Based Rendering", Proceedings of the Computer Vision Winter Workshop 2004, 2004

[14] Luciani, A., "Dynamics as a common criterion to enhance the sense of presence in virtual environments", Proceedings of 7th International Conference on Presence, Presence 2004, 96-103. Valencia, Spain, 2004

Mapping the way to fun: The effect of video game interfaces on presence and enjoyment

Paul Skalski
University of Minnesota Duluth
{pskalski@d.umn.edu}

Ryan Lange
Michigan State University
{langerya@msu.edu}

Ron Tamborini
Michigan State University
{tamborini@msu.edu}

Abstract

This study examines the potential for natural mapping to affect presence-related outcomes of video game exposure. Interactivity in the form of natural mapping has been suggested to be a possible contributor to presence, but few studies to date have investigated this potential, particularly as it applies to video games, which are expected to make extensive use of naturally mapped interfaces in the future. The present study addressed how the control system used to interact with a video game impacts the experience of the game, with the expectation that natual mapping would relate positively to spatial presence and game enjoyment. A total of 48 subjects took part in an experimental study manipulating the type of interface used to play a PC driving video game (steering wheel, joystick, or keyboard). They then completed measures of perceived interface naturalness (included as a manipulation check), spatial presence, and game enjoyment. Results of the study were consistent with expectations, though questions remain about the independence of the endogenous variables and causal direction of the observed relationships.

1. Introduction

The popularity of video game entertainment has soared in recent years. In 2005, the game industry reaped record profits of $10.5 billion in the U.S. [1], due in part to advances in game technology. Game industry growth has traditionally been fueled in part by technical innovation [2], and many such developments are on the horizon, including High Definition (HD) graphics and new game playing interfaces, the focus of this investigation.

Game playing interfaces have progressed considerably in capabilities over time, from single-button joysticks to the multiple-button and stick controllers of today [3]. Newer control devices allow players to interact with a game using a wider, more realistic range of actions conducive to the sensation of *presence*, “the perceptual illusion of non-mediation” [4]. Few studies to date have examined the impact of different game interfaces on presence, however, despite the considerable attention interfaces have received from the video game industry in recent years. Sony, for example, is expected to make the motion-tracking *EyeToy* interface a prominent feature of its upcoming *Playstation 3* console, and Nintendo’s next generation *Wii* system is banking on a revolutionary new control device that works through natural hand and arm movements such as swinging [5]. These developments highlight the increasing role in gaming of an especially important type of interactivity, natural mapping.

Natural mapping is defined as “the ability of a system to map its controls to changes in the mediated environment in a natural and predictable manner” [6]. It has been suggested to be a key contributor to spatial presence experienced during video game play [7], in part because more naturally mapped interfaces allow players to access more complete mental models of real-world behavior [8]. More naturally mapped interfaces (such as a club for a golfing game or racket for a tennis game) should lead players to perceive more control naturalness, which should in turn relate positively to spatial presence, a mediating variable in this study. When players feel more spatially present in a video game environment, they should indicate more enjoyment of the game post exposure. Enjoyment is a particularly important consideration of the interactive entertainment industry, simply because enjoyed games are typically the most successful and profitable ones.

Based on the above rationale, the following hypotheses and research question are posited:

Hypothesis 1: Individuals who play a video game with more naturally mapped controls will perceive greater interface naturalness than those who play the same game with less natural mapping.
Hypothesis 2: Perceived interface naturalness predicts spatial presence experienced while playing a video game.
Hypothesis 3: Spatial presence predicts video game enjoyment.
Research Question 1: What else predicts video game enjoyment?

2. Methodology

A single-factor, between-subjects experiment was conducted manipulating the type of control device used to play a PC driving video game, *Need for Speed Underground II*. A total of 48 participants played the game for ten minutes using either a steering wheel controller (more natural mapping), a joystick (less natural mapping), or a keyboard (less natural mapping). They then completed measures of perceived interface naturalness (which also served as a manipulation check), spatial presence, enjoyment, skill, prior game use, and demographic characteristics.

3. Analysis and Results

ANOVA was used to test the first hypothesis, and multiple regression analyses were used to test the last two hypotheses and research question. Regression variables were entered in three blocks: skill and demographics, prior game use, and media condition/study exposure outcomes. By block three in both the spatial presence and enjoyment regression analyses, a significant portion of variance in the dependent measure was explained.

Hypothesis 1, which predicted that players of a game with a more naturally mapped control device would perceive greater interface naturalness than those who played the game with less natural mapping, was supported. Participants in the driving wheel condition perceived a higher level of interface naturalness (7-point-scale $M = 4.76$, $SD = 1.00$) than those who played the game using a joystick ($M = 3.09$, $SD = 1.59$) or keyboard ($M = 3.13$, $SD = 1.53$), $F\ (2, 45) = 7.83$, $p<.01$, $eta=.26$. A Tukey HSD post-hoc test confirmed that this effect was driven by the wheel condition.

Hypothesis 2, predicting that perceived interface naturalness predicts spatial presence, was also supported. When the final block of variables was added, the only significant predictor of spatial presence was the perceived naturalness variable, $\beta = .72$, $p < .01$.

Hypothesis 3, predicting that spatial presence relates to enjoyment, was also supported. When the final block was added, the strongest predictor of enjoyment was spatial presence, $\beta = .59$, $p < .01$.

In answer to Research Question 1, additional predictors of enjoyment included prior game play in general ($\beta = .43$, $p < .01$), prior driving game play in particular ($\beta = .43$, $p < .01$), and use of the Nintendo Entertainment System ($\beta = -.25$, $p < .05$).

4. Discussion

This study provides preliminary empirical support for the predicted positive relationship between natural mapping and presence [5, 7, 9]. It also offers evidence that spatial presence mediates the relationship between natural mapping and video game enjoyment. However, high observed associations between the perceived naturalness, spatial presence, and enjoyment scales raise questions about the dimensionality of these measures—it may be that all three scales are tapping aspects of presence. In the case of perceived naturalness, this seems especially likely, though it was treated as separate in this investigation to provide a manipulation check and more varied form of natural mapping. In the case of enjoyment, this would be consistent with conceptualizations of spatial presence that include enjoyment-related concepts [9].
This also raises questions about causality—if presence and enjoyment are second-order unidimensional, it may be that enjoyment is a necessary condition for a state of presence to be maintained, and not the reverse. Although the present study treated presence as a mediating variable influencing media enjoyment, future research should attempt to untangle these closely-related concepts to gain a better understanding of how both operate in media users.

Despite these questions/limitations, the core finding of this study—that perceived interface naturalness, spatial presence, and enjoyment are closely and positively associated—is important by itself. It points to the exciting potential of naturally mapped interfaces, both in gaming and in other contexts. Natural mapping and resultant spatial presence may not only positively impact video game enjoyment, but also other positive outcomes of presence such as training. In the case of training, natural mapping in simulators can provide users with a more complete mental model for how to perform the real-life actions they are learning, resulting in more skills transference. At the same, the downside of natural mapping must also be addressed. If natural mapping can enhance the learning of prosocial skills, it can also create stronger mental models for antisocial behavior, such as firing a weapon or punching and kicking, as players of highly-mapped violent video games might do. Future work should address these and other outcomes of natural mapping as it relates to presence.

References

[1] NPD Group. (2006). *The NPD group reports annual 2005 video game industry retail sales.* Retrieved April 20, 2006 from http://www.npd.com/dynamic/releases/press_060117.html
[2] Williams, D. (2002). A structural analysis of market competition in the U.S. home video game industry. *International Journal on Media Management*, 4(1), p. 41-54.
[3] Skalski, P. (2004) *The quest for presence in video game entertainment.* Presented on presence panel at the Central States Communication Association Annual Conference, Cleveland, OH.
[4] Lombard, M., & Ditton, T. B. (1997) At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2). Retrieved April 30, 2002 from http://www.ascusc.org/jcmc/vol3/issue2/lombard.html
[5] Terdiman, D. (2006). Nintendo shows off Wii controller. *ZDNet News.* Retrieved June 15, 2006 from http://news.zdnet.com/2100-1040_22-6070233.html
[6] Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, 42(4), 73-93.
[7] Tamborini, R., Eastin, M., Skalski, P., Lachlan, K., Fediuk, T. & Brady, R. (2004). Violent virtual video games. *Journal of Broadcasting and Electronic Media,* 48(3), 335-357.
[8] Tamborini, R. & Skalski, P. (2006). The role of presence. In Vorderer, P. & Bryant, J. (Eds.), *Playing computer games: Motives, responses, and consequences.* Mahwah, NJ: Lawrence Erlbaum.
[9] Wirth, W., Bocking, S., Hartmann, T., Klimmt, C., Schramm,H., & Vorderer, P. (2003). *Presence as a process: Towards a unified theoretical model of formation of spatial presence experiences.* Unpublished manuscript.

Does "Being There" Improve Memory: The Impact of Presence on Recall

James Denny, David Atkin
Cleveland State University, University of Connecticut
(j.denny@csuohio.edu, david.atkin@uconn.edu)

Abstract

This study intends to investigate the roles of color vs. black-and-white in film and their impact on the level of presence experienced by viewers. The study also seeks to understand the relationship between color vs. black-and-white in terms of their effects on visual and factual recall, as well as the role of presence on these types of recall.

Keywords---**presence, visual recall, factual recall, color vs. black-and-white film**

Past research suggests that the existence of color in a mediated presentation leads to changes in the way a person may view that mediated offering. Newspaper readers were found to prefer color over black-and-white in four-color halftones. [1] Also, the inclusion of color photographs in newspapers was discovered to alter the manner in which people read that paper. It was found that color photos drew people's attention more than black-and-white photos. [2] In researching mediated presentations of moving images, studies utilizing what was then the advent of color television found that viewers felt that they were more participatory in televised events than those viewers who watched the same programs in black-and-white. [3; 4] The color group was described as being more absorbed and emotionally involved than the black-and-white group. [3; 4] The black-and-white group was found to have paid greater attention to the commentators of the programs. [3; 4].

Based on these findings, it appears that the color group was more immersed and engaged in the visual presentations than was the black-and-white group. Immersion and Engagement are two of the three elements of Presence, a communication concept which has also been described as a sense of "being there" for audience members. Through prior research, it has been suggested that color should evoke more presence than black-and-white. [5] Also, color should add a greater element of reality for viewers. [6] Realness is the third element of Presence. With recall also having been discussed in terms of color vs. black-and-white [3; 4], three research questions were developed.

Our three research questions consist of, *(1) Will color lead to greater recall? (2) Will color lead to a greater sense of presence?* and *(3) Will presence play a role in recall?* To explore these questions, two groups of college students were shown footage from a World War II documentary [7], which utilized film from the war that was originally shot in color. One group was shown a ten-minute segment in its original color, while the second group was shown the same segment with the color having been stripped away, leaving it in black-and-white. Following the viewing sessions, both groups were administered a questionnaire that included measures for immersion, engagement, realness, visual recall, and factual recall. A statistical analysis of the relationships between these variables was then conducted using SPSS.

(1) It was found that color was significant in its relationship to factual recall, but not for visual recall. (2) Participants who watched the segment in color reported did not report a greater level of presence, though this relationship was also found to be approaching significance. This relationship was also found to be significant. (3) Also, participants who experienced a higher sensation of presence scored higher in their factual recall score. Visual recall was found to not significantly be affected by color, black-and-white, or a greater level of Presence.

References

[1] Click, J. W., & Stempel, G.H., III (1976). Reader response to front pages with four-color halftones. *Journalism Quarterly*, *53*, 736-738.

[2] Josephson, S. (1996). Questioning the power of color. *News Photographer*, *51*(1), 4-8.

[3] Scanlon, T.J. (1967). Color television: New language? *Journalism Quarterly, 44*, 225-230.

[4] Scanlon, T. J. (1970). Viewer perceptions on color, black and white TV: An experiment. *Journalism Quarterly, 47*, 366-368.

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

[6] Sherman, B.L., & Dominick, J.R. (1988). Perceptions of colorization. *Journalism Quarterly, 65*, 976-980.

[7] Palmer, G. And S. Pearson. (Producers). (2003). *The perilous fight: America's world war II in color* [Motion Picture]. United States: PBS Home Video.

Presence, Efficacy, and the Net: Exploring Patterns in Political Participation from a Comparative Perspective

Mehpare Selcan Kaynak, Cheryl Campanella Bracken
Bogazici University, Cleveland State University
{selcan.kaynak@boun.edu.tr, c.bracken@csuohio.edu }

1.0 Introduction

This paper aims to examine the concepts of presence, efficacy and political participation as interlinked phenomena that might explain the impacts of internet usage on political behavior. More specifically, it first introduces presence as a new variable to consider political participation, and then examines its connections to political efficacy.

In previous studies, participants have reported experiencing a sense of presence (the "illusion of non-mediation" [Lombard & Ditton, 1997]) while using the Internet (Tu & McIssac, 2002). In addition, the concept of presence has been established as a mediating variable in audiences' perceptions of credibility (Bracken, 2006). Building from this perspective, we examine whether or not presence and efficacy can be linked in explaining motivations to political participation. In this context, the particular experiences that evoke feelings of presence are also expected to increase feelings of efficacy.

A final contribution of this study is its comparison of American and Turkish college age students (representing the only generation that has grown up with the Internet so far), in determining how much the technology itself, as opposed to particularities of the respective political cultures, contribute to our understanding of internet's impact on political participation.

RQ1: Can presence sensations predict political participation?

RQ2: Can presence sensations predict internal efficacy?

RQ3: Does nationality/culture influences presence sensations?

2.0 Methodology

A survey was conducted in both the United States and in Turkey. The 221 respondents were students attending university in both countries. The survey included items relating to presence, political efficacy, media use, and political participation.

2.1 Presence

Two measures of presence were included in the survey: Immersion and social richness [1].

2.2 Political Participation

Participation was measured using an additive scale in which respondents were asked if they ever participated in political behaviours ranging from going to a political website to running for a political office.

3.0 Results

RQ1 was tested using forced entry multiple regression. The overall model was significant (p =.045). However, the presence dimensions were not predictors of political participation. RQ2 was also tested multiple regression. But the presence dimensions did not significantly predict feelings of internal efficacy.

The last RQ inquired about the influence of nationality/culture on general presence sensations. To test this RQ a t-test was employed with the respondents' home institution serving as an indicator of "culture" as the independent variable. The two presence dimensions were the dependent variables. The t-Test for both presence dimensions was significant (p = .001) for both DVS. The American students reported higher levels of immersion (M = 4.51) than the Turkish students (M= 3.80). The American students also gave higher ratings for social richness (M = 4.88) than did the Turkish students (M = 4.23).

4.0 Discussion

This exploratory study found unique differences between American and Turkish respondents for general sensations of both immersion and social richness. While the relationship between general sensations of presence, efficacy, and political participation was not significant, it might be because the respondents were not exposed to a particular political content in which to respond. Further investigation is necessary.

References

[1] Lombard, M., & Ditton, T. B. (1997) At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2), Retrieved April 30, 2002 from http://www.ascusc.org/jcmc/vol3/issue2/lombard.html

[2] Tu, C. H., & McIssac M. S., (2002). An examination of social presence to increase ineraction in online classes. *American Journal of Distance Education, 16*(2), 131- 150.

[3] Bracken, C.C. (2006, in press). Source credibility and presence: The role of television form. Journal of Broadcasting & Electronic Media.

I felt like it happened to me: Television audience perceptions of televised conflict

Renée A. Botta, Jill Rudd
University of Denver, Cleveland State University
{rbotta@du.edu, j.rudd@csuohio.edu}

Abstract

An increasing number of studies have linked presence to other media effects, typically as a mediating variable. Scholars have also examined individual differences in the experience of presence. For this study, the authors tested whether real world experiences would impact presence sensations when similar scenarios were acted on in television programming. A random sample telephone survey was utilized to ask 439 respondents questions about their personal experiences of conflict in their home life, the severity of the conflict, the genre they watch most often on television, and the extent to which they experience a sense of immersion and perceptual realism when viewing conflict on television.

The results indicate that men and women who reported experiencing family conflict that escalated to a threat of physical violence or to actual physical violence also reported feeling more immersed when viewing conflict on television and that women with these real world experiences also reported a greater sense of realism when viewing conflict on television. A finding that approached significance may indicate that conflict on nonfiction genres such as reality programming, news, and sports feels particularly real to those who have experienced family violence.

The implications are that such highly emotional experiences can be "re-lived" through television and lead such viewers to experience heightened states of presence. Given the studies in which presence is a mediator of other media effects, these results may also suggest that viewers who have experienced family violence get a double dose effect of television aggression through such re-lived experiences.

A Markerless Augmented Reality System for the treatment of phobia to small animals

Juan, M.C.[1], Joele, D.[2], Baños, R.[3], Botella, C.[4], Alcañiz, M.[1], van der Mast, Ch.[2]
[1] MedICLab, Technical University of Valencia (Spain)
[2] Mediamatics, Delft University of Technology (The Netherlands)
[3] Departamento de Personalidad, Evaluación y Tratamientos Psicológicos. University of Valencia (Spain)
[4] Departamento de Psicología Básica y Psicobiología, Universitat Jaume I (Spain)

Abstract

Augmented Reality has proved to be effective in the treatment of phobia to small animals. The first system developed for this purpose uses visible markers for tracking. In this paper, we present a second version of this system, where the markers are not visible to the user. We use ARToolKit for tracking. The system incorporates two cameras: one colour camera for capturing the real world without visible markers and one infrared camera for capturing the real world whereby it is possible to distinguish markers drawn with a special ink. We are carrying out a study to compare the sense of presence and reality judgment in non clinical populations using both systems.

Keywords--- **Augmented Reality, invisible markers, phobia to small animals, virtual therapy**

1. Introduction

People who suffer from arachnophobia or other types of phobia to small animals become anxious when they are in a situation where these animals can appear. They suffer an unrealistic and excessive fear that makes life miserable. They are always frightened of seeing the animal they fear. Until now these phobias have been treated using exposure therapy and Virtual Reality (VR). Carlin et al. [1] used immersive Virtual Reality for exposure therapy. The first experiment was carried out at the U.W. Human Interface Technology laboratory (HITLab) (www.hitl.washington.edu/projects/exposure). The first patient treated with this system needed 12 VR therapy sessions of one hour. First, she started at completely the other end of the virtual world from the virtual spider. Slowly, she got a little closer. In later sessions, after she had lost some of her fear of spiders, she was sometimes encouraged to pick up the virtual spider and/or web with her cyberhand and place it in positions that were most anxiety provoking. Other times, the experimenter controlled the spider's movements. Some virtual spiders were placed in a cupboard with a spiderweb. Other virtual spiders climbed or dropped from their thread from the ceiling to the virtual kitchen floor. Eventually, after getting used to them, she could tolerate holding and picking up the virtual spiders without panicking. She could also pull the spider's legs off.

Renaud et al. [2] compared tracking behaviour with a virtual spider and a neutral target in fearful and non-fearful subjects.

Garcia et al. [3] explored whether VR exposure therapy was effective in the treatment of spider phobia. They compared a VR treatment condition vs. a waiting list condition (participants waiting for treatment, but without treatment) in a between group design with 23 participants. Participants in the VR treatment group received an average of four one-hour exposure therapy sessions. VR exposure was effective in treating spider phobia compared to a control condition as measured with a "fear of spiders questionnaire", a behavioural avoidance test, and severity ratings made by the clinician and an independent assessor. 83% of patients in the VR treatment group showed clinically significant improvement compared with 0% in the waiting list group, and no patients dropped out.

Botella et al. [4] presented a telepsychology system that uses VR to treat the phobia to small animals (cockroaches, spiders and rats). Patients follow the treatment in their own house. In this system, a typical kitchen was modelled. The system had different levels at which one or more small animals could appear. The animals randomly appeared when the user opened a door of a cupboard. It was possible to kill the animals and dispose of them in a dustbin.

Juan et al. [5] and Botella et al. [6] used a visible marker Augmented Reality system for treating 10 real patients with phobia to cockroaches and spiders. These treatments reduced the fear and avoidance of the feared animal in only one hour-long session [7]. This system uses a visible marker (a white square with a black border containing symbols or letters) to determine the position and orientation of where the animal has to appear. Although that system has proven its effectiveness, in this paper we present a second version of the system that uses invisible markers. Why is it important that the marker appears or not? One of the steps of our protocol to treat patients with phobias is that patients have to search for the feared animal in the same way they would do in their house. In our case, to simulate this search we use three boxes, under one of which a marker is randomly placed. This is where the animals will appear. When the patient first sees part of the marker, the system still does not show the animals, but the patient knows they will appear. If the marker is not visible, the patient does not know that the animals are going to appear and when they appear they will produce the desired surprise. This is the main reason why we have developed this second version of the system. Our hypothesis is that markerless systems could increase the sense of presence and reality judgment. We are carrying out a study to confirm this.

The paper is organized into four sections. Section two gives an overview of the system. Section three presents some results and in section four we include some conclusions.

2. Markerless Augmented Reality System for the treatment of phobia to small animals

2.1. Hardware

The hardware components and their characteristics are:

- An infrared (IR) camera, IR bullet camera with 715 nm IR filter. The IR bullet camera comes in a lipstick sized tube 2.5 inches long, with a diameter of 0.8125 inches. The diagonal FOV of the camera is 92 degrees. The image sensor is 1/3" CCD with 290,000 CCIR pixels, capable of delivering a video stream at a frame rate of 30 fps in several image formats, among them 640x480. The output of the camera is a composite video signal. A regulated 12 VDC power supply is needed for proper operation. USB2.0 Video Grabber converts a video composite signal into a USB 2.0 signal. It delivers a frame rate of up to 30 fps at a resolution of 640x480. The device is compliant with DirectShow.
- A colour, Dragonfly camera. The dimensions of the Dragonfly camera are 63.5x50.8 mm. This type of camera has a Sony 1/3" progressive scan CCD sensor, which delivers uncompressed 24-bit true colour images at a resolution of 640x480, with a maximum frame rate of 30 fps. The computer connection consists of a 6-pin IEEE-1394 interface. Camera parameters, such as white balance and exposure time, can be changed through image acquisition software, giving it a horizontal field of view (FOV) of 42.2 degrees, and a vertical FOV of 32 degrees. This corresponds to a diagonal FOV of 52 degrees.
- A Cy-visor HMD. The Cy-visor HMD is a closed headset, which integrates two SVGA 800x600 micro displays that have a diagonal FOV of 31 degrees. It supports several video input modes, but can also be configured to display the output of a computer monitor. The distance between both microdisplays can be changed manually.

All the above-mentioned components are shown together in Figure 1.

In order to draw the invisible marker we have used a special ink. We also considered powder, but it was discarded because it needs an external light source. Both have a finite durability. This implies that markers have to be drawn every week. The ink we have used is invisible to ultraviolet light and the human eye. Ultraviolet light is below 400 nm. The human eye can see light between 400 nm and 750 nm. The ink emits 840 nm frequency light and has a 793 nm absorption frequency, which lies in the infrared range. There exist three classifications for infrared light: near-IR, mid-IR and far-IR. This classification is based on ascending light frequencies. The ink falls in the near-IR spectrum, 750 nm to 3000 nm. The ink itself is delivered in the form of an ink pen (IR1PenSM). The tip of the pen is fluorescent green to the human eye. The ink is influenced by ultraviolet light. Anything written on white paper can not be seen by the human eye but, when viewed with a modified camera with infrared functionality, the ink becomes visible. An external light source is not required.

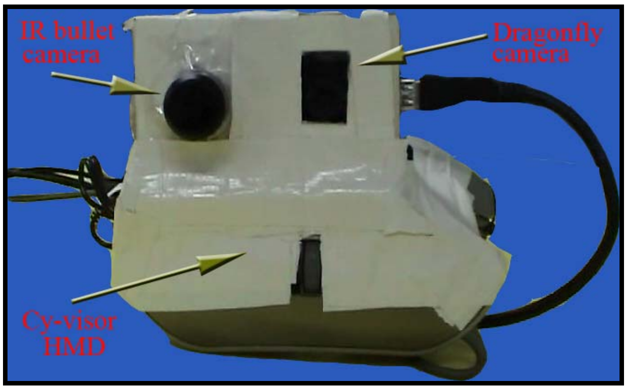


Figure 1. Capture and visualization system (Dragonfly camera, Infrared camera and the HMD)

a) b)

Figure 2. VRML models of: a) the static cockroach b) model of the medium spider

2.2. Software

We programmed the system using Microsoft Visual Studio C++ version 6.0 as the development environment. We used ARToolKit [8] version 2.65 [9] with Virtual Reality Modeling Language (VRML) support to incorporate AR options. The three-dimensional models of the virtual elements were constructed using Autodesk 3D Studio Max, version 5.0. These models were exported to VRML format and edited with VRMLPad, version 1.2. Textures were created in Adobe Photoshop, version 7.0. The graphical user interface was created with the OpenGL Utility ToolKit (GLUT)-based user interface library (GLUI). Sound support is provided by the OpenAL sound library.

The second version of the system uses the same four VRML models as the first system used. We have three different spiders and one cockroach. For each type of animal, three models have been created; a non-moving, a moving and a dead animal. To obtain as real a result as possible the moving cockroach is animated with moving legs and moving tentacles, and the spiders move their legs. In Figure 2.a) the resulting model of a static cockroach is shown, after the texture has been applied. The model of the medium spider is shown in Figure 2.b).

When the animals are killed the system produces a sound similar to that of a real animal being killed. The system includes two sounds: a squirting sound similar to the sound of a real can of insecticide; and a squishing sound similar to that of a real cockroach or spider being crushed.

2.3. Description of the system

The system uses a video see-through HMD with one colour camera and one infrared camera to view a scene concurrently. It is not possible to detect markers in images captured by the colour camera but it is in images of the infrared camera. The video of the infrared camera is treated by the system and using ARToolKit it is possible to establish the position and orientation of the infrared camera with respect to the marker. As the relationship between the infrared camera and the colour camera is known, it is possible to determine the position and orientation of the invisible marker in the video of the colour camera. In this way, the virtual element is superimposed over the place where the invisible marker is situated. As a result, the user sees a scene where no marker exists, but the virtual element appears in the right position and orientation. The resulting image is finally shown on the microdisplays of the HMD. Figure 3 shows the set up of the system and the place where tests are being carried out.

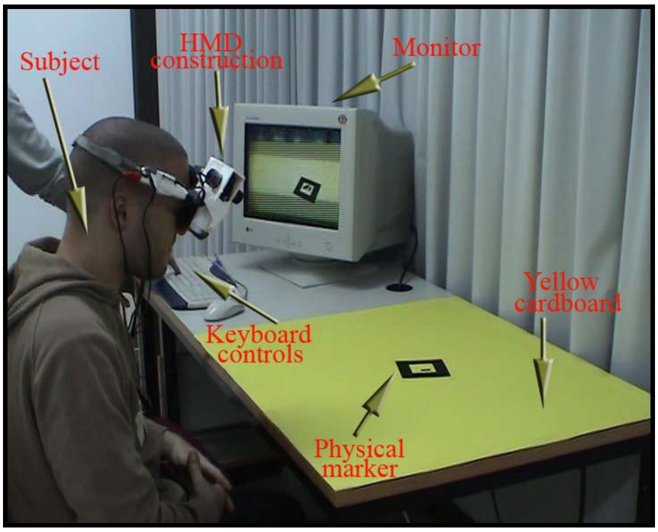


Figure 3. Set up of the system and the place where tests are being carried out

The functionality of the system is the same as the marker system [5]. Therefore, in this paper we only comment on its functionalities briefly. The user can select the number of animals to appear: one animal, increase/reduce in three animals, increase/reduce in 20 animals. The animal/s can increase/reduce its/their size. The animals can move or stop. It is possible to kill animals using two different elements. When this occurs the system plays a sound related to the tool used and one or more dead animals appear. If the spider system is used, it is possible to choose among three types of spiders.

3. Results

In this section, several images of the marker and markerless systems running are shown. Figure 4 shows two similar situations. In Figure 4.a) the user is using the markerless system. Figure 4.b) shows the view of the user that appears in Figure 4.a). Figure 4.c) shows a similar image, but in this case, the user is using the marker system. If Figures 4.b) and 4.c) are compared, we consider that Figure 4.b) is more natural than Figure 4.c).

We believe that the markerless system could be even more effective than the marker system, but before testing the system with real patients, we are carrying out a study to determine the sense of presence and reality judgement when participants use both systems.

The study is still in progress and at the moment it includes participants recruited by advertisements in the University campus, all of whom are students, scholars or employees at the Technical University of Valencia (age range from 21 to 40). All participants fill out the fear and avoidance to cockroaches and spiders questionnaires (adapted from Szymanski and O'Donohue questionnaire [10]). Participants are divided into two groups. The first group uses the marker system first and then the markerless system. The second group uses the markerless system first and then the marker system.

In order to check the sense of presence that participants experience using both systems, they fill out a questionnaire adapted from the Slater et al. questionnaire [11]. For checking the reality judgment that participants experience using both systems, they fill out a questionnaire adapted from the Reality Judgment and Presence questionnaire of Baños et al. [12].

The study is still in progress and for the moment we do not have enough participants in order to extrapolate any conclusions, but in a preliminary analysis of the data, we can say that participants seem to have a greater sense of presence and reality judgment using the markerless system than using the marker system. Moreover, these scores seem to be greater if participants use the marker system first and then the markerless system. We will be able to extrapolate the final conclusions once the study is finished.

4. Conclusions

We have presented the first Markerless Augmented Reality system for the treatment of phobia to small animals. We believe this technique will be suitable for any Augmented Reality System where it is important that users do not see the marker and thus do not know where the virtual elements will appear.

Our next project is to use it to treat real patients. The marker version has been used to treat one patient with phobia to cockroaches [6], five patients with phobia to cockroaches and four with phobia to spiders [5]. In all cases, patients reduced their fear and avoidance of the feared animal in only one session [7] of treatment using the visible marker Augmented Reality system. Moreover, all of them were able to kill the real animal after the treatment.

Before the treatment none of them were able to approach or interact with the live animal without fear.

We are now carrying out a study comparing our Augmented Reality system with markers with the system presented here. While we finish our study, we can say that participants that have tested the system seem to have greater sense of presence and reality judgment with the markerless system than they experience using the marker system. The study is still in progress and we will be able to present our final conclusions once it is finished.

References

[1] Carlin, A., Hoffman, H. y Weghorst, S., "Virtual reality and tactile augmentation in the treatment of spider phobia: a case study", Behaviour Research and Therapy, Vol. 35. N. 2, pp. 153-158, 1997

[2] Renaud, P., Bouchard, S., Proulx, R., "Behavioral avoidance dynamics in the presence of a virtual spider", IEEE Trans. On Information Technology in Biomedicine, Vol. 6. N. 3, pp. 235-243, 2002

[3] García-Palacios, A., Hoffman, H. G., Carlin, A., Furness, T., & Botella, C., "Virtual Reality in the treatment of spider phobia: A controlled study. Behaviour", Research and Therapy, Vol. 9, pp. 983-993, 2004

[4] Botella , C., Baños, R., Quero, S., Perpiñá, C. & Fabregat, S. "TelePsychology and Self-Help: The treatment of phobias using the Internet", Cybertherapy, San Diego (EEUU), 2004

[5] Juan, M.C., Alcañiz, M., Monserrat, C., Botella, C., Baños, R., Guerrero, B., Using Augmented Reality to treat phobias, IEEE Computer Graphics and Applications, Vol. Nov-Dic, pp.31-37, 2005

[6] Botella, C., Juan, M.C., Baños, R.M., Alcañiz, M., Guillen, V. & Rey, B., Mixing realities? An Application of Augmented Reality for the Treatment of Cockroach phobia. Cyberpsychology & Behavior, Vol. 8, pp. 162-171, 2005

[7] Öst, L., Salkovskis, P. & Hellstroöm, K. "One-session therapist directed exposure vs. self-exposure in the treatment of spider phobia", Behavior Therapy, Vol. 22, pp. 407-422, 1991

[8] Kato, H., Billinghurst, M., "Marker tracking and HMD calibration for a video-based augmented reality", Conferencing system. 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99), San Francisco (California), pp. 85-94, 1999

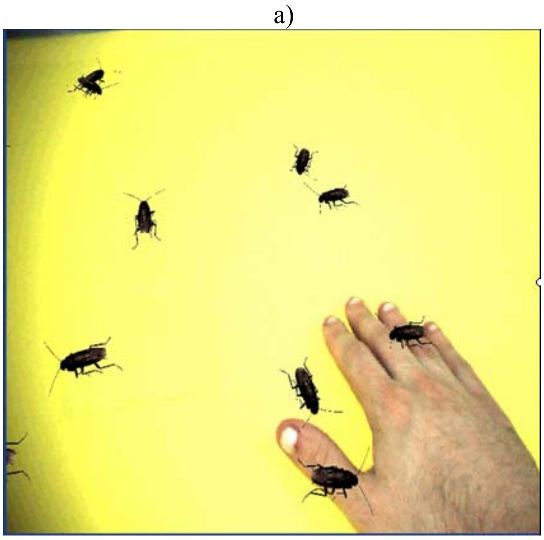
[8] ARToolKit version 2.65vrml, http://www.hitl.washington.ed/artoolkit, 2004

[10] Szymanski, J., O'Donohue, W., Fear of spiders questionnaire. Journal of Behavior Therapy and Experimental, Psychiatry, Vol. 26, N. 1, pp. 31-34, 1995

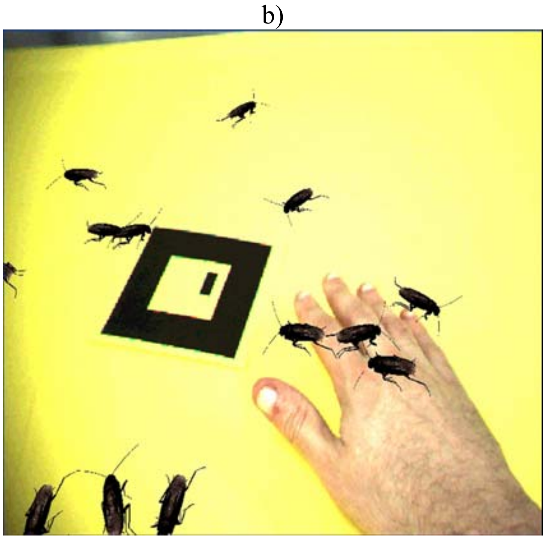
[11] Slater, M., Usoh, M., & Steed, A., Depth of presence in virtual environments. Presence: Teleoperators and Virtual Environments, Vol. 3, pp. 130-144, 1994

[12] Baños, R.M., Botella, C., García-Palacios, A., Villa, H., Perpiña, C., Alcañiz, M., Presence and Reality Judgment in Virtual Environments: A Unitary Construct?, Cyberpsychology & Behavior, Vol. 3, N. 3, 2000

a)



b)



c)

Figure 4. a) User using the system. b) User using the markerless system. c) View of the user using the marker system.

The Effects of Witness Viewpoint Distance, Angle, and Choice on Eyewitness Accuracy in Police Lineups Conducted in Immersive Virtual Environments

Jeremy N. Bailenson, Alexandra Davies, Jim Blascovich, Andrew C. Beall, Cade McCall, and Rosanna E. Guadagno
Department of Communication, Stanford University, Department of Psychology, University of California, Santa Barbara
{Bailenson@stanford.edu, AyDavies@gmail.com, Blascovich@psych.ucsb.edu, Beall@psych.ucsb.edu, McCall@psych.ucsb.edu, Guadagno@psych.ucsb.edu}

Abstract

The current study investigated the value of using immersive virtual environment technology as a tool for assessing eyewitness identification. Participants witnessed a staged crime and then examined sequential lineups within immersive virtual environments that contained three-dimensional, virtual busts of the suspect and six distractors. Participants either had unlimited viewpoints of the busts in terms of angle and distance, or a unitary view at only a single angle and distance. Furthermore, participants either were allowed to choose the angle and distance of the viewpoints they received or were given viewpoints without choice. Results demonstrated that unlimited viewpoints improved accuracy in suspect-present lineups but not in suspect-absent lineups. Furthermore, across conditions, post-hoc measurements demonstrated that when the chosen view of the suspect during the lineup was similar to the view during the staged crime in terms of distance, accuracy improved. Finally, participants were more accurate in target-absent lineups than in suspect-present lineups. Implications of the findings in terms of theories of eyewitness testimony are discussed, as well as the value of using virtual lineups that elicit high levels of presence in the field. We conclude that digital avatars of higher fidelity may be necessary before actually implementing virtual lineups.

1. Introduction

Our goal was to examine the use of immersive virtual environments (IVEs) as a tool for eyewitness identification and to explore theoretical issues underlying the behavior of eyewitnesses. Indeed, a recent issue of the journal *PRESENCE: Teleoperators and Virtual Environments [5]* was dedicated to papers exploring the interplay of IVEs and the legal system. Given the unique capability of IVEs to provide experimental control in terms of the amount and type of facial visual information presented to witnesses, the ability to correctly identify a perpetrator should increase as a result of better matching of visual cues sensed by the witness during the lineup to those sensed during the crime. Given that virtual humans have reached a state technologically to achieve high levels of copresence, it is possible to begin studying their use in applications that take advantage of the presence that IVEs provide.

1.1. Eyewitness Identification and Police Lineups

Police and prosecutors have long relied on eyewitnesses to crimes to identify criminals, and their testimony has a unique and powerful influence on juries and judges during criminal trials. In fact, such identifications are considered direct evidence, rather than circumstantial evidence—to put this in perspective, Wells [48] points out that even fingerprints are not considered direct evidence. Unfortunately, eyewitness testimony is not only one of the most compelling types of direct evidence to which a jury is exposed, but as many studies have shown, it can also be quite unreliable [49] . If the results of experiments using staged crimes are any indication, false identification (i.e., instances in which the witness incorrectly identifies someone other than the actual perpetrator in a lineup) occurs at rates potentially as high as 70% [31] , though this percentage varies according to testing situations. This problem is compounded by the fact that false identifications are often asserted with as much, if not more, confidence than accurate identifications [6] . While there has been much work examining ways to improve on these techniques, here we focus on reviewing research that relates theoretically to the manners in which IVES can potentially ameliorate the deficiencies of eyewitness testimony in lineups.

1.2. Prior Research on Improving Lineups Relevant to IVEs

1.2.1. Context Reinstatement. It has often been proposed that providing contextual cues to eyewitnesses of a crime (e.g., taking the witness back to the scene of the crime) will help them remember details about the crime and, in particular, facilitate accuracy in identifying the perpetrator of a crime. Context is expected to enhance recall because it provides memory retrieval cues [28] .

The idea that context reinstatement will enhance the accuracy of eyewitness recognition grew out of research on context effects on recognition memory. Research in this area has indicated that pairing faces or words with contextual cues will enhance performance on recognition tasks [35] [36] [37] [39] Thus, it seems reasonable to expect that these results may generalize to eyewitness identification.

There are several dimensions relevant to context reinstatement. The dimension most relevant to the present investigation is the physical or visual environment in which the encoding occurred. Methods commonly utilized to examine context reinstatement include the following: photographs of the environment [16] , objects from the environment [28] , guided recollection [33] , and mental contextual reinstatement cues [24] [17] .

Relevant to the current work, a small number of studies have examined physical context reinstatement and found that returning participants to the scene of the crime does enhance performance in facial recognition. For instance, two studies by Smith and Vela [45] examined this issue by staging a memorable event and asking participants to identify the confederate while in the context in which the event took place or in a different context. The results revealed that recognition performance was better when the recognition task took place in the same place as the memorable event.

However, results have been mixed across different context reinstatement methods. For instance, some have found that context reinstatement facilitates recognition performance (e.g., [34]) and others have found a null or weak relationship between context and recognition accuracy [17] [18] [40] . A meta-analysis on facial identification and eyewitness identification studies indicated that there was a significant effect for context reinstatement in two directions [42] . Context reinstatement increases overall correct identifications but it also increases false alarms. Additionally, the authors reported that the magnitude of improvement via context reinstatement was greater in lifelike situations than in laboratory studies.

Often times, it is impractical or impossible to return an eyewitness to the scene of a crime owing to changes in season, weather, lighting, etc. However it is possible that a virtual recreation of the crime scene could produce the same context reinstatement effects as the less practical option of returning an eye-witness to the physical scene of a crime even to the extent of recreating seasonal, weather, and lighting effects. The idea of using IVEs to recreate the scene of a crime was first investigated by Guadagno, Bailenson, Beall, Dimov, and Blascovich [21] . These authors staged a live crime and the examined eye-witness accuracy in the same or different context in which the crime took place that was either a physical room or a virtual recreation of the room. The results revealed that simultaneous line-ups (i.e., lineups with all the suspects present at once) that took place in either the virtual or physical scenes of the crime produced more accurate identifications of the perpetrator of the staged crime than did comparable line-ups in a different physical or virtual room.

1.2.2. Facial Information Processing. Research suggests that while the memory of a familiar face is sufficiently abstracted to allow for recognition despite a variety of transformations, the memory of an unfamiliar face (a face seen for the first time) depends heavily on the image upon which that memory is based [10] . An early meta analysis [42] of the literature on the recognition of unfamiliar faces concluded that across studies successful recognition was significantly influenced by the degree to which the images used in training materials matched those used at testing. Specifically, consistency of pose, context, and viewing angle proved important. Even in matching tasks where memory was not an issue, participants frequently failed to match images of unfamiliar faces at varying angles and lighting [25] . Given this sensitivity, we might expect an improvement in recognition rates if subjects are given the opportunity to select an angle that matches the angle of their original exposure.

Another potential advantage of active exploration of a three dimensional head is that it allows transformations of view. Schiff, Banka, and de Bordes Galdi [41] found that participants were better at recognizing targets in dynamic “mug shots” (in which the targets were videotaped rotating 180 degrees) than from static mug shots. Similarly, Christie and Bruce [11] found a slight advantage for recognition of unfamiliar faces when those faces were moving (either by nodding or shaking from side to side) than from nonmoving faces. If, as these data suggest, movement across views provides additional information that aids recognition [11] , we might expect the viewpoint transformations created by the witness’s own movement around the digital version of the suspect to produce similar effects.

1.2.3. Comparing Media of Suspect Representation. Some of the theoretical questions of the current study concerning amount of facial information presented in the lineup have been addressed via other media. As technology has advanced, particularly in the last twenty years, photographic and, more recently, video lineups have largely begun replacing live lineups for the purpose of eyewitness identification. A recent survey has found that a majority of lineups conducted by police jurisdictions in the US are not live (27% live, 73% photographic; [51]). In terms of the effectiveness of the various types of lineups, research comparing the methodologies is quite mixed, with most studies finding no large differences while others finding advantages and disadvantages.

For example, some studies have demonstrated that live lineups typically lead to minimally higher levels of accuracy than mediated lineups [19] [13] On the other hand, photographic lineups have been shown to lead to more correct identifications compared to live lineups, the latter leading to more false identifications [42] . According to Cutler, Berman, Penrod and Fisher [12] , the conservative conclusion based on available research findings at the time is that witnesses viewing live lineups, videotaped lineups or photo arrays perform similarly (see also [43] [14] .

On the other hand, Valentine and Heaton [47] found that video lineups are more fair than their live counterparts for a variety of reasons, including facilitation of large databases for distracters, avoidance of the physical collocation of witness and suspect in order to reduce witness anxiety [1] , and provision of video editing to prevent subtle, unintentional behaviors as in live lineups that could potentially bias the witness [51] . Furthermore, there is the risk that some of the important details of a suspect's appearance are lost in a photograph compared to a live or video lineup, as evidenced by lower accuracy rates found in studies such as Egan, Pittner and Goldstein [19] and Cutler and Fisher [13] .

In sum, there do not seem to be extremely large differences in patterns of results between live, photographic and video lineups. However, in the following section we discuss the use of IVEs, which provide a media that is a hybrid of photographic, video, and live lineups, perhaps combining the various advantages of the aforementioned three types of lineups.

1.2.4. The Potential of IVEs in Police Lineups. Bailenson, Blascovich, Beall, & Noveck [4] have discussed the potential for IVEs as a tool to improve eyewitness identification as well as other legal processes. An IVE lineup might feature the witness entering some type of digital environment that includes *virtual busts*, digital reconstructions of the suspects and distracters. Previous work has outlined the specific methodology of constructing IVEs and virtual busts [2] , and has provided empirical evidence that current technology is sophisticated enough to produce busts that highly resemble suspects. Immersive virtual environments offer a number of possibilities for improving lineups that are not as easily or effectively achieved with live and photographic lineups.

For example, in terms of context, reconstructing a virtual representation of the crime scene is easily and effectively achieved. Rather than being shown photographs of the crime scene prior to viewing the lineup or given instructions to imagine the crime scene (tactics often utilized to generate context reinstatement) - the witness could be placed in a virtual re-creation of the crime scene, and view the lineup from within that virtual world. Thus, if the crime occurred in a liquor store, the witness could view the lineup in that liquor store virtually without ever having to physically return to the scene of the crime. This is particularly useful when it would be traumatic for the witness to return to the actual crime scene, or when the crime scene no longer exists; for example if the offense is arson.

If one considers context on a perceptual level, IVEs offer an extremely valuable advantage in terms of representing suspects and distractors. With databases of photographs used in lineups, there is often variance in terms of the distance from the camera to the person as well as the camera angle. However, using IVEs one can lock the viewpoint and make it identical across all lineup members, thus maximizing the ability to control lineup fairness. Furthermore, using IVEs, lineup members can be made to appear identical in all ways except for the criteria on which the eyewitnesses should be trying to differentiate them; for example, all of the lineup members can be shown wearing identical clothing, hairstyles, and accessories. To illustrate, if the perpetrator had a beard at the time the crime was committed, but shaved it prior to the lineup, that might pose a problem for the eyewitness trying to identify him [38] . In a virtual lineup, that beard can be easily reinstated on the suspect, and all the lineup members, so that the witness can see them as they would have looked when the crime was committed.

Most importantly, IVEs are unique in allowing eyewitnesses to view suspects from any angle or distance without compromising the witness's anonymity or forcing the witness to get near the criminal. In live lineups, witnesses are unable to approach lineup members and view them close up because being in the same room as the suspect could be dangerous or emotionally traumatic. While video lineups potentially allow this action to take place, it is unrealistic to expect stock video footage to cover every single angle and distance between potential witnesses and the foils. On the other hand, from a single IVE digital model, an infinite pattern of examinations are easily and safely achieved, as IVEs allow witnesses to view suspects from any angle or distance chosen—from six centimeters away if they prefer—without ever placing the witness and suspect in the same physical room. Such features allow witnesses an active, unlimited exploration of lineup members that would never be possible in the physical world or from stock video footage.

Currently a majority of studies, as well as actual police procedures, rely on photo lineups or live lineups where the eyewitness views the targets from a distance [6] . These types of lineups afford the eyewitness small degrees of visual information (i.e., limited viewing angles and level of detail) compared to a virtual lineup where they can see the target from any angle and any distance.

One might expect that unlimited examination (i.e., examining the lineup members from whatever viewpoints the witness chooses) would allow for better recognition cues than for limited exploration (i.e., looking at a single snapshot of a suspect). On the other hand, unlimited visual information about the witness may be counterproductive if the information available at time of retrieval was not actually present at the time of encoding. For example, if an eyewitness witnessed a crime and only saw the face of the perpetrator from one specific angle, seeing the lineup members from other viewpoints might prove distracting; using IVEs we can ensure that the original viewpoint is the only angle from which both the suspect and all the distractors are displayed. Consequently, witnesses would not receive extraneous information from mug shots which likely contribute to false identifications, and on the other hand they would receive information that may be essential but not included in a mug shot (e.g., the back of someone's head). Therefore, IVEs may provide a mechanism to achieve a new level of context reinstatement.

One unique aspect of IVEs is that users may actively and freely move about an environment to examine the suspects and distractors. Consequently, a lineup may be improved by allowing an eyewitness to recreate the dynamic motion of the suspect's criminal behavior (or of the eyewitness himself) during the IVE lineup procedure. Research demonstrates that the processes governing perception of human faces have a substantial spatial gestalt component [20] [46] . These researchers provide evidence that not only are the local features of a face important in recognition but also the global configuration of those local features, which shifts at different angles and distances. Consequently, exploration of faces may be similar to exploration of other objects, to which studies show that *active* navigation results in superior performance compared to passive measures [26] . By allowing an eyewitness to explore a digital environment containing suspects, as opposed to merely looking at photographs, IVEs could take advantage of the manners in which humans encode and retrieve information about faces.

2. Overview of Experiment

We examined the possibility that allowing witnesses to set the specific angle and distance between themselves and the suspects during lineups would improve accuracy. Participants witnessed a staged crime in vivo and subsequently participated as an eyewitness in a sequential lineup, containing 7 suspects, within a digital IVE.

Three independent variables were manipulated via the sequential lineup: *view, exploration mode,* and *perpetrator presence*. There were two levels of *view (unlimited, limited)* such that participants in the unlimited condition were free to view the suspect from any distance and any angle (similar to a face-to-face examination). Participants in the limited condition were only able to view the suspect from a single distance and angle (similar to a 3D photograph). There were two levels of *exploration mode* (*active, passive*) such that participants in the active condition could walk and turn their heads naturalistically to manipulate their viewpoint. Participants in the passive condition were yoked (i.e., linked) to the archived movements of a previous participant and could not choose their distance or viewpoint; their viewpoints were updated by the system as if they were watching a stereoscopic movie based on the movements of another participant. There were two levels of *perpetrator presence (present, absent)*.

We hypothesized that lineup accuracy would be greater for participants with unlimited views than those with a limited view because this would increase the chance that the visual cues they encoded during the staged crime would actually be available for recognition cues during the lineup.

We also predicted better accuracy for participants in the active exploration mode than the passive exploration mode. Previous work has demonstrated better memory with active examination of stimuli than with and passive navigation through both physical space [44] and in VEs [23] [27] [9] . Furthermore, if there was variance in the angles in distances from which participants viewed the staged crime due to sitting in different seats, leaning in different directions, etc., then an active viewpoint choice should allow witnesses to best match the viewpoint (i.e., angle and distance between them and the suspects) of the lineup to that of the crime.

We predicted that participants would be more accurate in target absent lineups than in target present lineups, as previous work utilizing sequential lineups had indicated high accuracy of correct rejections [15] [30] .

3. Method

3.1. Participants

Ninety-eight adult participants (52 males, 46 females) were recruited from the undergraduate population at Stanford University. They received either course credit or a payment of $10. Their mean age was 19.80 (*SD* = 1.83). Participants were randomly assigned to one of eight conditions based on the crossing of the independent variables with approximately the same gender ratio in each condition.

3.2. Design

The experimental design was a 2 (view: limited vs. unlimited) x 2 (exploration mode: active vs. passive) x 2 (suspect-presence: present vs. absent) between-subjects factorial with twelve participants in each of the eight conditions. Each participant in the passive condition was yoked to a previous participant in the active condition, such that across the active and passive conditions the amount of visual information received was identical. Two confederates served as suspects for this study, with half of participants seeing a staged crime committed by Confederate A, and the other half seeing the crime committed by Confederate B. The confederates were balanced across conditions such that each one appeared in each of the eight between-subjects conditions for approximately half of the participants. Participants witnessed the staged crime either alone or in groups of two.

3.3. IVE System

The specific technology utilized to render the IVE is described in detail in previous research [3] and is depicted in Figure 1. Participants' physical movements along X,Y,Z

Figure 1: A depiction of the immersive virtual environment system. The components are 1) WorldViz PPT position tracking cameras, 2) Virtual Research V8 head-mounted display and Intersense orientation tracking sensor, and 3) image generator.

spatial dimensions were tracked by optical sensors (Worldviz PPT X4, update rate 60 Hz) and head rotations were tracked by accelerometers in the physical room (Intersense IS250, update rate of 150 Hz). Perspectively-correct stereoscopic images were rendered by a 1700 MHz Pentium IV computer with an NVIDIA 5950 graphics card, and were updated at an average frame rate of 60 Hz. The system latency in the head-mounted display (HMD) was 45 ms maximum. The software used to assimilate the rendering and tracking was Vizard 2.17. Participants wore an nVisor SX HMD that featured dual 1280 horizontal by 1024 vertical pixel resolution panels that refreshed at 60 Hz. The display optics presented a visual field subtending approximately 50 degrees horizontally by 38 degrees vertically.

3.4. Perpetrators (i.e., Confederates)

The two confederates were physically similar (see Figure 2) allowing for a single set of suspect foils (i.e., distractors), weighing approximately 88 kilograms and having a height of approximately 1.8 meters.

Foils were selected using a match-description strategy [50] . Thirteen judges unfamiliar with the confederates read a brief written physical description of the perpetrator: "Caucasian male, between the ages of eighteen and twenty-two, sandy blond or light brown hair, blue or green eyes, athletic build, 5'10 to 6'2". The judges (7 males and 6 females) then examined a set of forty screenshots of male virtual busts (i.e., a virtual three- dimensional model of head and shoulders) created using modeling software (3dMeNow) and frontal and profile photographs (see [2] for more information about the virtual busts).

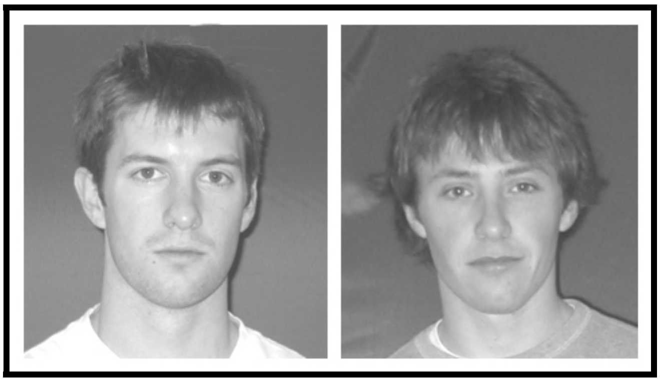


Figure 2: Photographs of the two confederates used in the study.

The software created a three-dimensional mesh to fit the shape of a person's head and then wrapped a texture which is the result of two photographs stitched together around the mesh. The forty busts were selected by the experimenters a priori from a large database because they roughly fit the physical description of the suspect. Each judge rated how

Figure 3: The eight heads used in the lineups. The rightmost two heads on the bottom are models of the two confederates depicted in Figure 2.

well they thought each of the forty faces matched the description on a seven-point scale with higher numbers indicating a better fit. The six faces that received the highest mean scores (M = 5.16, SD = .42) were used as foils (see Figure 3). In the perpetrator-absent condition, the confederate who had not committed the staged crime was included as a foil as well.

3.5. Procedure

Upon arrival at the lab, each participant was taken into a lab room, seated, and asked to read and complete a consent form as well as a demographic questionnaire containing items concerning race, gender, and other personal characteristics. Participants were run either individually or in groups of two.

Next, the experimenter told the participant(s) that one more participant would be arriving. After waiting two minutes, the experimenter said that she would go downstairs to find the last participant. The experimenter left a purse with a wallet sticking out of it on a chair approximately 2 meters in front of the participants. Immediately after leaving the room, the experimenter cued the confederate who waited one minute before he entered the room. The confederates were blind to experimental condition.

The confederate walked into the room where the participant(s) were waiting, established eye contact, asked "Is this the virtual reality study?" and then asked "Are you the person running the study?" The participant typically replied that he or she was only participating, and that the experimenter had gone to look for the last participant and would be back shortly. The confederate proceeded to get visibly agitated, raising his voice and exclaiming that he had "no time to just wait around for the experimenter to return." While speaking, the confederate began to stare at the purse and wallet. After proclaiming that he could "not wait any longer!" the confederate grabbed the wallet and ran from the room. Overall, the confederates spent between 45 and 60 seconds in the room with the participants.

Immediately after the confederate exited, the experimenter re-entered and explained that the theft had been staged. The experimenter indicated that for the time being the participants should behave as if they had actually witnessed a crime, and explained that the participant(s) would be asked to identify the perpetrator in a lineup. When there were two participants, one sat in a chair in the hallway during the ten to fifteen minutes it took the other participant to view the lineup. Eighty of the 98 participants were run in pairs; the ratio of participants who went first to participants who went second was held constant across the eight between-subjects conditions.

The virtual room that contained the lineup was a digital representation of the room in which the crime occurred (same dimensions, blank walls, similar carpeting, etc.). Once they donned the HMD, participants were given a brief tutorial that explained how to use the gamepad to enter responses to questions while inside the virtual environment. They were then given a practice session to ensure that they understood how to use the gamepad. Next, all participants received the same basic instructions:

We are interested in whether you can accurately identify the perpetrator who you saw steal the wallet. To do this, you will view a lineup of suspects in a virtual room. The perpetrator may or may not be present in that lineup. You will view each member of the lineup individually, and after seeing each one it will be your job to determine whether or not he is the perpetrator. Once you make a decision, you cannot change it. After making each selection, you will be asked how confident you are in your response on a scale of 1 to 7, 1 being "not at all confident" and 7 being "extremely confident."

Participants were not told the number of suspects that would appear in the lineup in advance. After hearing the instructions, participants were asked if they had any questions. Most did not, and those that did almost uniformly wanted clarification that the presentation of the lineup would be sequential rather than simultaneous. Participants were then instructed to notify the experimenter if at any point they began to feel uncomfortable in the virtual environment. Then, depending on each participant's experimental condition, they were read one of the following additional sets of instructions:

Passive-Limited:

We would like you to stand here, without moving, and view the suspects. You will not be able to look around the room, or view the suspects from any angle other than the one you see initially, so there is no need to move your head or body.

Passive-Unlimited:

We would like you to stand here and view the suspects. You will be able to see the suspects from different angles and distances, but you will not be able to choose what viewpoints you see them from, so there is no need to move your head or body.

Active-Limited:

We would now like you to view the suspects. You may walk around to view the suspects from different angles and distances. At first you will only be able to see a plaster head. Once you decide from what angle and distance you would like to view the suspect, press the "Right" button on your gamepad and you will then be able to see the suspect clearly from your chosen viewpoint, for as long as you like. You will be able to choose a different viewpoint for each suspect you see.

Active-Unlimited:

We would now like you to view the suspects. You may walk around the suspects and view them from whatever angle or distance you like, for as long as you like.

After seeing the lineup, participants were given a suspicion probe to determine if they knew about the staged crime in advance. None of the 98 participants had any prior knowledge of the crime. The whole process took approximately one hour.

3.6. Measures

3.6.1. Accuracy. Participants received a score of one if they answered no in the target absent condition (correct rejections) or if they correctly identified the suspect in the target present condition (hits). Otherwise they were scored with a zero. Overall, the accuracy rate was 34 percent. Although it is not possible to do any formal signal detection analysis because each subject only contributed a single score, the Appendix breaks down the responses across subjects to allow interpreting the data by hits, misses, false alarms, and correct rejections.

3.6.2. Confidence. Participants rated how confident they were in their decision after viewing each person in the lineup on a scale from one to seven, with higher numbers indicating higher confidence. We took the mean confidence score of the seven responses (Cronbach's alpha =.61). The mean score was 3.57 (*SD* = 5.59). This factor was included to assess the relationship between confidence and accuracy across the different conditions.

3.6.3. Interpersonal Distance. We recorded the minimum distance in meters between each participant and the suspect over the seven trials he or she traversed (or was led in the passive condition) the virtual lineup. The mean minimum distance was 1.32 (*Max* = 3.00, *Min* = .28, *SD* = .73).

3.6.4. Suspect Presence. We also recorded the number of times participants indicated the suspect was present (i.e. "yes" responses) in the one of the seven sequential trials versus absent in the lineup. Participants had a slight bias to give "no" responses 55 percent of the time.

4. Results

Post experimental interviews indicated that not a single participant was aware prior to the experiment that a staged crime would be occurring and that none of the participants were aware that the confederate was acting. There were neither statistically significant differences nor notable trends between participants run alone or those run in groups of two. We also tested for a) order of receiving lineup (first versus second) for instances in which there were two participants, and b) confederate (A or B) and found no significant differences on any of our measures.

We ran a 2 (view: limited or unlimited) x 2 (exploration mode: active or passive) by 2 (perpetrator presence: present or absent) ANOVA with accuracy score as the dependent variable. The dependent variable was either a 0 or 1, and consequently was not normally distributed. While ANOVA is resilient against this assumption violation [29] , to be sure we repeated every analysis reported using a nonparametric, binary logistic regression. The patterns of significance (and non-significance) were identical. For the sake of simplicity, we report the ANOVA data.

There was a main effect for perpetrator presence, $F(1, 90) = 19.95$, $p<.001$, partial Eta Squared = .19. As Figure 4 depicts, participants were more accurate in suspect-absent lineups than in suspect-present lineups. There was also a significant interaction between view and perpetrator presence, $F(1, 90) = 6.40$, $p<.01$, partial Eta Squared = .07.

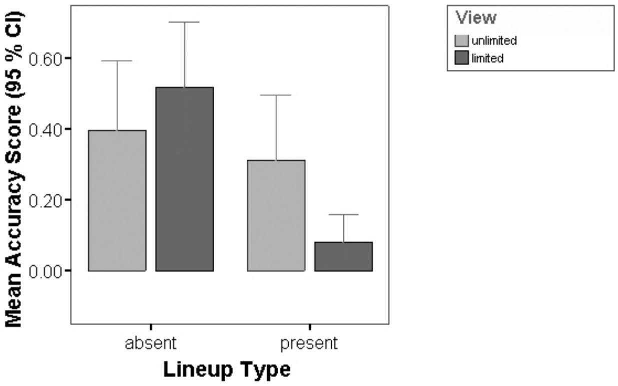


Figure 4: Accuracy scores by condition.

As Figure 4 demonstrates, having unlimited information improved participants score in suspect-present lineups but not in suspect-absent lineups. There was no significant effects involving exploration mode.

We had hypotheses concerning the distance participants maintained between themselves and the virtual suspects; consequently we repeated the above ANOVA with including interpersonal distance as a linear covariate. None of the effects from the original ANOVA changed, however there was a significant effect of interpersonal distance, $F(1,85) = 3.98$, $p<.05$ partial Eta Squared = .05, with large minimum distances signaling high accuracy. Note that the degrees of freedom are slightly lower in this analysis because distance data from three subjects was lost due to equipment failure.

To further explore the notion of distance, we computed the viewpoint from the average distance and angle from which all participants (in the active condition) chose to view the suspect. Figure 5 depicts this mean viewpoint. Interestingly, this viewpoint features a smaller ratio of head-to-image than most photographs utilized in most traditional photo-lineups. Indeed, participants were more accurate when their distance was farther away from the face of the suspect. In order to examine this effect further, we examined a random sample of the videos of the confederates committing the staged crimes, and established that on average, the confederates had left between approximately two and three meters between themselves and the participant(s) due to the placement of the chair with the money on it. The confederates had been instructed to maintain such a large distance in order to prevent a situation in which a participant might feel tempted to physically attempt to stop him from

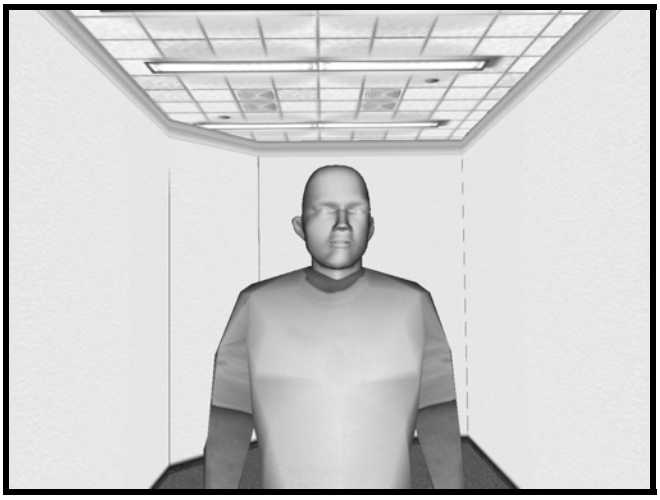


Figure 5: A suspect with a plaster head displayed to participants before they chose an angle and distance in the limited condition shown at the overall average viewpoint chosen by participants.

taking the money. Consequently, one explanation for the distance effect is that participants who maintained similar interpersonal distance levels while witnessing the crime and the lineup (i.e., between two and three meters) were more accurate than those who received disparate visual information.

It is important to note that there are other potential explanations for this result. For example, it could be the case that the virtual busts have visual artifacts that are inaccurate only when viewed up close. Alternatively, people who are truly accurate may not need to move close to the suspect because they have already made their decisions. Our distance/viewpoint matching explanation, though certainly intuitive given the data, is ad-hoc.

We next repeated the above ANOVA with confidence scores as the dependent variable. No effects approached significance (All *F*'s < .5, all *p*'s > .5). However the correlation between confidence and accuracy was positive, $r(98) = .24$, $p < .02$, such that people who were correct were more confident in their decision.

In terms of suspect presence (i.e., yes or no responses), we ran an ANOVA with view (limited or unlimited), exploration mode (active or passive) and perpetrator presence (present or absent) and accuracy score as the dependent variable. There was a marginal effect of view, $F(1,90) = 3.20$, $p < .08$, partial Eta Squared = .03, with a trend towards more yes responses with unlimited views (See the Appendix). No other effects approached significance (All *F*'s < 1.5, all *p*'s > .2).

5. Discussion

In the current experiment, we examined the use of IVEs for police lineups, and predicted that active compared to passive exploration would increase eyewitnesses' correct identifications of criminal suspects, and that an unlimited view would provide better recognition than a limited view. Both of those predictions share a common underlying process assumption—that unlimited, active view choices would best match visual information during encoding and during retrieval on the part of witnesses. Our results showed some support for this underlying prediction. An unlimited view led to better accuracy than a limited view in the suspect-present condition, regardless of whether or not the witnesses chose the viewpoint. Similarly, matching the lineup distance (between the suspect and the witness) and the staged crime distance improved accuracy. Finally, similar to previous research utilizing sequential lineups, participants were more accurate in suspect-absent lineups than in suspect-present lineups.

Interestingly, participants often chose to examine viewpoints that were different from the viewpoints at which they actually witnessed the crime. Hence, giving a witness active control over their viewpoint may be counterproductive without instructing them to only attempt to view the potential suspects from their original viewpoint. Future research should examine this phenomenon by providing eyewitnesses with specific instructions about viewpoint choice.

One unexpected finding was the extremely large size of the main effect of suspect presence. Participants were much more accurate when the suspect was absent (i.e., correct rejections) than when the suspect was present (i.e., hits). Our explanation for this disparity is necessarily ad hoc. Previous research that has utilized a sequential lineup procedure has also demonstrated high rates of correct rejections [15] [30] . What is surprising in the current work is the extremely low rate of correct identifications (i.e., hits). Participants had an extremely difficult time identifying the suspect in the limited condition in which they could only see the virtual busts from a single view.

This finding potentially stems from the fact that our three-dimensional models were not perfect analogs of the suspects, as a comparison between Figures 2 and 3 reveals relatively large disparities between the photographs of the confederates and the avatars of the three-dimensional models of the confederates. Undoubtedly, as technology improves, IVEs will become an increasingly viable alternative to both photographic and live lineups. In the meantime, IVEs have the capability to aid researchers in illuminating the positive and negative aspects of live and photographic lineups because it allows the isolation of variables such as viewpoint and context. The results of this study indicate that there are attributes of virtual lineups (such as recreating a range of viewing distances) that are beneficial to eyewitness identifications that cannot be equivalently reproduced using other, more traditional techniques.

The ease with which lineup members can be depicted in a variety of outfits, hairstyles and even locations is unparalleled, and has potentially profound advantages for eyewitness identification. In the future, it will be possible to program lineup member avatars to literally go through the motions of crime commission, exactly as the witness remembers the crime to have occurred, in a high-presence, virtual re-creation of the original context of the crime scene. The possibilities that arise with IVE technology can potentially revolutionize lineup creation and eyewitness identification. It is imperative, then, as technology advances, to periodically evaluate virtual lineups compared to real-world lineups to determine when virtual reality can be considered a feasible, perhaps even superior, alternative to live and photographic lineups.

Conclusions

In the current study, we demonstrated the potential for improving lineups with IVEs. Although there were limitations with the current data due to the fidelity of the virtual busts given current technology, the possibilities that arise with IVE technology can potentially revolutionize lineup creation and eyewitness identification. It is imperative, then, as technology advances, to periodically evaluate virtual lineups compared to real-world lineups to determine when virtual reality can be considered a feasible, perhaps even superior, alternative to live and photographic lineups.

Acknowledgements

This research was partially supported by grants from Stanford University's Office of Technology and Stanford University's Media-X center to the first author, and from NSF ITR grant 0219399 to the fourth author. We would like to thank Ariana Young for her helpful comments on this paper, Aaron Sullivan, Andrew Orin, and Nick Yee for programming assistance, and Eric Ford, Bryan Kelly and Drew Taylor for their assistance collecting data.

References

[1] Ainsworth, P. B., & King, E. (1988). Witnesses' perceptions of identification parades. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues, 1*. Chichester: Wiley.

[2] Bailenson, J. N., Beall, A. C., Blascovich, J., & Rex, C. (2004). Examining virtual busts: Are photogrammetrically-generated head models effective for person identification? *PRESENCE: Teleoperators and Virtual Environments, 13(4),* 416-427.

[3] Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2001). Equilibrium revisited: Mutual gaze and personal space in virtual environments. *PRESENCE: Teleoperators and Virtual Environments, 10*, 583-598.

[4] Bailenson, J. N., Blascovich, J., Beall, A. C., & Noveck, B. (2006). Courtroom applications of virtual environments, immersive virtual environments, and shared virtual environments. *Law and Policy,* 28(2), 249-270.

[5] Barfield, W., Lauria, R., Mann, S., & Kerr, I. (2005). Special Section: Legal, Ethical, and Policy Issues Associated with Virtual Environments and Computer Mediated Reality Guest Editors' Introduction. *PRESENCE: Teleoperators and Virtual Environments, 14(6)*, iii-v.

[6] Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law & Human Behavior, 25*(5)*,* 475-491.

[7] Beizer, D. (2005). Tech success: Round up the usual suspects. *Post News Week Tech*.

[8] Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13,* 103-124.

[9] Brooks, B. M., Attree, E. A., Rose, D., Clifford, B. R., & Leadbetter, A. G. (1999). The specificity of memory enhancement during interaction with a virtual environment. *Memory, 7*(1), 65-78.

[10] Bruce, V., & Burton, M. (2002). Learning new faces. In Fahle, M., & Poggio, T., (Eds.), *Perceptual Learning* (pp. 317-334). Cambridge: MIT Press.

[11] Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition*, *26,* 780-790.

[12] Cutler, B. L., Berman, G. L., Penrod, S. D., & Fisher, R. P. (1994). Conceptual, practical, and empirical issues associated with eyewitness identification test media. In Ross, D., Read, J.D., & Toglia, (Eds.). *Adult Eyewitness Testimony: Current Trends and Developments.* New York: Cambridge University Press.

[13] Cutler, B. L., & Fisher, R. P., (1990). Live lineups, videotaped lineups, and photoarrays. *Forensic Reports, 3*, 439-449.

[14] Cutler, B. L., Fisher, R. P., & Chicvara, C. L., (1989). Eyewitness identification from live versus videotaped lineups. *Forensic Reports, 2,* 93-106.

[15] Cutler, B. L., & Penrod, S. D. (1988). Improving the reliability of eyewitness identifications: Lineup construction and presentation. *Journal of Applied Psychology, 73*, 281-290.

[16] Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology, 72*, 629-637.

[17] Cutler, B. L., Penrod, S. D., O'Rourke, T. E., & Martens, T. K. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behavior, 1,* 113-134.

[18] Davies, G., & Milne, A. (1985). Eyewitness composite production: A function of mental or physical reinstatement of context. *Criminal Justice and Behavior*, *12*, 209-220.

[19] Egan, D., Pittner, M., & Goldstein, A. (1977). Eyewitness identification: Photographs vs. live models. *Law and Human Behavior, 1*(2), 199–206.

[20] Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482-498.

[21] Guadagno, R. E., Bailenson, J. N., Beall, A. C., Dimov, A., & Blascovich, J. (2005, May). Virtual line-ups: Can immersive virtual context reinstatement facilitate eye-witness recall? Hot topic talk presented at the *American Psychological Society* Convention, Los Angeles, CA.

[22] Guadagno, R.E., Blascovich, J., Bailenson, J.N., McCall, C. (2006, in press). Virtual Humans and Persuasion: The

Effects of Agency and Behavioral Realism. *Media Psychology*.

[23] Harman, K. L., Humphrey, G. K., & Goodale, M. A. (1999). Active manual control of object views facilitates visual recognition. *Current Biology, 9*, 1315-1318.

[24] Hershkowitz, I., Orbach, Y., Lamb, M. E., Sternberg, K. J., & Horowitz, D. (2001). The effects of mental context reinstatement on children's accounts of sexual abuse. *Applied Cognitive Psychology, 15*, 235-248.

[25] Hill, H., & Bruce, V. (1996). Effects of lighting on matching facial surfaces. *Journal of Experimental Psychology: Human Perception & Performance, 22*, 986-1004.

[26] James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., & Goodale, M.A. (2001). Active and passive object recognition in a virtual environment [Abstract]. *Journal of Vision*, *1*(3), 292a, http://journalofvision.org/1/3/292/, doi:10.1167/1.3.292.

[27] Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R.G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science, 9*(4), 293-298.

[28] Krafka, C., & Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality and Social Psychology*, *49*, 58-69.

[29] Lindman, H. R. (1974). *Analysis of variance in complex experimental designs.* San Francisco: Freeman & Co.

[30] Lindsay, R. C. L., Lea, J., Nosworthy, G., Fulford, J., Hector, J., LeVan, V., & Seabrook, C. (1991). Biased lineups: Sequential presentation reduces the problem. *Journal of Applied Psychology, 76*(6), 796-802.

[31] Lindsay, R. C. L., & Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior*, *4*, 303-313.

[32] Malpass, R. S. & Devine, P. G. (1981a). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology*, *66*(4), 482-489.

[33] Malpas, R. S. & Devine, P. G. (1981b). Guided memory in eyewitness identification. *Journal of Applied Psychology*, *66*(3), 343-350.

[34] Malpass, R. S. & Devine, P. G. (1984). Research on suggestion in lineups and photo-spreads. In G. L. Wells & E. F. Loftus (Eds.), *Eyewitness Testimony: Psychological Perspectives* (pp. 64-91). Cambridge: Cambridge University Press.

[35] McKenzie, W. A., & Tiberghien, G. (2003). Context effects in recognition memory: The roll of familiarity and recollection. *Cognition and Consciousness*, *13*, 20-38.

[36] Memon, A. & Bruce, V. (1983). The effects of encoding strategy and context change on face recognition, *Human Learning*, *2*, 313-326.

[37] Parker, A., Gellatly, A., & Waterman, M. (1999). The effect of environmental context manipulation on memory: Dissociation between perceptual and conceptual implicit tests. *European Journal of Cognitive Psychology*, *11*, 555-570.

[38] Pozzulo, J. D., & Lindsay, R. C. L. (1995, June). Eyewitness identification procedures when appearance has been changed. *Society for Applied Research in Memory and Cognition*, Vancouver.

[39] Rainis, N. (2001). Semantic contexts and face recognition. *Applied Cognitive Psychology*, *15*, 173-186.

[40] Sanders, G. L. (1984). The effects if context cues on eyewitness identification responses. *Journal of Applied Social Psychology*, *14*, 386-397.

[41] Schiff, W., Banka, L., & de Bordes Galai, G. (1986). Recognizing people seen in events via dynamic mug-shots. *American Journal of Psychology*, *99*, 219-231.

[42] Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100,* 139-156.

[43] Shepherd, J. W., Ellis, H. D., & Davies, G. M. (1982). *Identification Evidence*. Great Britain: Aberdeen University Press.

[44] Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9*(4), 315-320.

[45] Smith, S. M., & Vela, E. (1992). Environmental context-dependent eyewitness recognition. *Applied Cognitive Psychology*, *6*, 125-139.

[46] Tanaka, J. W., & Sengco, J. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25*, 583-592.

[47] Valentine, T., & Heaton, P. (1999). An evaluation of the fairness of police line-ups and video identifications. *Applied Cognitive Psychology, 13*, S59–S72.

[48] Wells, G. L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14*(2), 89-103.

[49] Wells, G. L., & Loftus, E. F. (2003). Eyewitness memory for people and events. In A. Goldstein, (Ed.), Comprehensive handbook of psychology, Volume 11, Forensic psychology. New York: John Wiley and Sons.

[50] Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology, 78*(5), 835-844.

[51] Wogalter, M. S., Malpass, R. S., & McQuiston, D. E., (2004). A national survey of US police on preparation and conduct of identification lineups. *Psychology, Crime & Law, 10*(1), 69–82.

Appendix:

Proportions of Accuracy, Proportions of "Yes" Responses, and Mean Confidence Scores by Condition. The denominator of the ratios indicates the total number of participants in that condition.

	Confederate A				Confederate B			
Exploration:	active		passive		active		passive	
View:	limited	unlimited	limited	unlimited	limited	unlimited	limited	unlimited
	Proportion of accurate responses							
Perpetrator present:	0/6	2/6	1/7	2/6	0/6	1/6	0/6	1/6
Perpetrator absent:	4/7	1/6	6/6	4/6	4/6	2/6	2/6	3/6
	Proportion of "yes" responses to perpetrator presence							
Perpetrator present:	1/6	4/6	3/7	4/6	4/6	2/6	1/6	2/6
Perpetrator absent:	3/7	5/6	0/6	2/6	2/6	4/6	4/6	3/6
	Mean Confidence Scores							
Perpetrator present:								
mean	5.86	5.88	5.65	4.88	5.52	5.05	5.55	6.12
SD	0.48	0.49	0.75	0.69	0.72	0.99	0.78	0.71
Perpetrator absent:								
mean	5.80	5.74	5.76	5.76	5.02	5.60	5.67	5.60
SD	0.74	0.67	0.89	0.82	1.06	0.73	0.88	1.06

Looking At or Looking Out: Exploring Monocular Cues to Cre ate a See-Through Experience with a Virtual Window

Wijnand IJsselsteijn, Willem Oosting, Ingrid Vogels, Yvonne de Kort, Evert van Loenen

[w.a.ijsselsteijn@t ue.nl; w.oosting@tue.nl; ingrid.m.vogels@philips.com; y.a.w.d.kort@tue.nl; evert.van.loenen@philips.com]

Abstract

In indoor environments, having a view from a window plays an important role in human physical and psychological well-being – particularly if the view contains natural elements. In places where physical windows are absent or the view is highly artifact-dominated, virtual windows can potentially play a beneficial role. The current paper presents a research experiment on the efficacy of three monocular depth cues, i.e., motion parallax, blur, and occlusion, in engendering a window-like 'see-through experience' using projected photorealistic scenes. Results indicate that all three cues have a significant main effect on the viewer's 'see-through experience', with motion parallax yielding the greatest effect size. These results provide a first step in identifying and testing the perceptual elements that are essential in creating a convincing virtual window.

Keywords **Virtual windows, restorative environments, depth perception, motion parallax, view replacements**.

1. Introduction

The ubiquity of windows in places designed for human occupancy suggests that people attach special importance to the collection of functions that windows may serve in their environment. Indeed, several studies have shown the importance of windows on people's psychological and physical well-being. For example, Finnegan and Solomon [1] report that employees in work environments without windows report higher levels of work-related stress and lower levels of job satisfaction than those with access to windows.

Of all window functions (e.g., view to the outside, light source, knowledge of weather and time of day, regulating air quality, providing situational and orientation cues, etc.), the view out appears among the most significant [2]. In cases where windows are available, the content of the view outside is of particular importance. Several studies have reported on beneficial and restorative[8] effects of views onto a natural scene [3,4], whereas views onto human-built environments yield effects which are similar to having no window at all [5]. For example, Ulrich [7], in one of the classic studies in this area, reports on an experiment where patients recovering from gallbladder surgery where either assigned to rooms overlooking a natural setting or facing a brick wall. The ones with the view to nature had shorter post-operative stays, received less negative evaluation from nurses, and took fewer narcotic analgesics after surgery, than matched patients in rooms facing a brick wall.

Heerwagen and Orians [8] investigated how people compensate for the lack of windows in their office. They found that people without access to a real window view tended to use more visual materials (e.g., posters) than those with access to a window (even when compensating for the potential effect of extra wall space available through the absence of a window), and that the content of such materials was more likely to contain natural scenes and landscapes than urban scenes. For a detailed review of studies on the effects of windows on work and well-being, we refer to [9].

In addition to office environments, a number of specialised environments exist in which access to a window view is limited or entirely absent. For instance, in prison, many cells do not have a window view for control or safety reasons. In a number of hospital settings, such as intensive care units, patients only have a very limited view outside, if at all. And in spite of Jules Verne's visionary descriptions of the *Nautilus* providing an all-around view to its crew through several large bi-convex glass windows, submarines are in fact quite claustrophobic places where the only direct view outside is provided through a small periscope, and people have to live and work in such a confined space for long periods of time. What is noteworthy here is that these particular environments are almost inherently stressful to their inhabitants, for a variety of reasons, implying that a stress-reducing and restorative effect of a window view onto nature could potentially be quite significant, even if such a view is provided artificially.

As the majority of the world's population lives and

[8] The preference of people to have views onto a natural scene has been explained from an evolutionary perspective, in terms of potential access to resources such as food and safety [3]. Others have argued for the role that natural scenes may play in the restoration of depleted attentional resources [6].

works in increasingly urbanised environments looking out of their windows (if they have that luxury) onto concrete buildings, parking lots and roads, there is a solid basis for investigating whether artificial views from virtual windows could provide beneficial effects similar to those of real views. However, before such comparisons can be realistically made, we need to identify and test the perceptual elements that constitute a 'window experience'.

1.1. Creating virtual windows

Perhaps the earliest examples of pretend views from nonexistent windows can be found in antiquity. Trompe l'oeil paintings, dating as far back as 400 BC, were meticulously detailed paintings created to look entirely realistic in texture and dimensionality when observed from a particular vantage point.

Figure 6 A modern trompe l'oeil [10]

More recently, a number of commercial efforts are being developed aimed at (re)placing a window view. One example, the TESS Round Skylight, is shown in Figure 2. Semi-transparent photographs are placed in front of a light source to simulate a window in windowless medical

facilities.

Figure 2 TESS Round Skylight [11]

Such a virtual view to the sky and trees is expected to have a relaxing effect on patients as they undergo treatment. In line with findings of Ulrich [7], TESS has specifically targeted its virtual window application towards windowless

healthcare environments, such as critical care units and MRI environments (see Figure 3).

Figure 3 Virtual window at an intensive care unit [11]

Other innovative ways to bring the benefits of nature views into health care settings include the electronic window of nature that simulates the passage of daylight from dawn to dusk, created by Joey Fischer/Art Research Institute Limited and used first in the United States at Stanford [12]. SensoryScapes Panels even provide multisensory stimulation, combining nature views, soothing sounds and botanical aromas. They are specifically targeted towards windowless healthcare and interior office environments (see Figure 4).

Figure 4 SensoryScape panel at an office environment [13]

Figure 5 Armas Magic Window [14]

Figure 6 – Magic Window MW1000 [15]

Figure 7 Philips DreamScreen prototype at the Philips HomeLab [16]

The examples discussed thus far typically rely on transparent photo sheets with a light source placed behind them. Replacing content in these cases requires additional effort, and printing new photographic sheets can also be time-consuming and costly. In contrast, other manufacturers utilize electronic displays for the purpose of simulating window-views in a more flexible manner. For example, the Armas Magic Window systems, as presented in Figures 5 and 6, are linking an array of eight TFT-screens per window, enabling the user to change the content of each window view dynamically using a computer. Note how the physical separation between the different TFT screens is resolved by implementing a window frame.

The current project was carried out in the context of the larger Dreamscreen project of Philips Research, which has the aim to study how wall- and window-sized video displays in combination with directional audio cues and other sensory stimulation may create a convincing and beneficial immersive experience to users in their homes. The DreamScreen prototype presented in Figure 7 is based on five front-projected displays.

When comparing real space to virtual space, limiting ourselves to visual media for the time being, we find that real world perception has several critical features [17], which we will briefly discuss here in the context of creating virtual windows.

a) Static depth information is provided via several independent mechanisms (e.g., linear perspective, occlusion, texture density gradients, binocular disparity) that are consistent with each other and the observer's viewpoint.

Although none of the examples provided earlier employ binocular disparity as a cue, most other static depth cues are consistent in relation to each other, which is of course more challenging in the case of trompe l'oeil paintings than in cases where photographic projections are being employed. Most current window substitutes, however, do not take into account that the frame surrounding the window is perceived at a different depth layer than the view from the window. Assuming one is attending to the view outside, the frame will be disappearing in increasing optical blur as a consequence of the accommodation of the eye. Although adaptive blur generation based on gaze tracking would be the ideal solution, there are other, more practical options. One has been explored some years ago by CRL, a company that produced the Vistral screen surround which was placed over the screen edge like a picture frame. It generated a Moiré effect from two layered patterns of dots on either side of a glass plate and was extremely difficult to focus the eyes on. This tricked the accommodative system to signal that the Vistral screen surround and the screen itself belonged to different depth layers. As a consequence, it made the image on the screen appear to float in depth.

b) The resolution and intensity of the image is only limited by the sensitivities of our visual system.

In the case of backlit photographic sheets, the resolution is high enough for our visual system not to detect the grain. For electronic displays, the resolution will be lower, but may be compensated by the fact that one view is composed of a combination of multiple displays, as well as the likelihood that these images will be viewed from distances of more than half a meter or so, which substantially increases the resolution per visual angle. However, low brightness levels can still be a serious showstopper for electronic windows at this point in time, although arrays of ultra-bright LEDs show promise in this respect.

c) The effective image size fills our entire field of view, limited only by our facial structures, but without an externally imposed frame.

This is true in general when perceiving the world around us, with the interesting exception of windows, where we have learned to perceive a framed view as part of reality, and not as a mediated representation. The effect of different types of framing on the virtual window experience is a matter of empirical research. Some guidelines are available from prior work in office settings where shape and size of windows were manipulated using apertures [18]. Windows occupying less than 10% of the window wall were regarded as extremely unsatisfactory, whereas window sizes of 20% and larger were deemed most satisfactory. In addition, participants preferred a wide lateral scan, selecting wider windows over taller ones.

d) Dynamic depth information (i.e., motion parallax) is coupled to observer movement.

None of the examples of virtual windows discussed earlier presently support motion parallax, that is, a change in the relative position of objects as a result of observer movement in front of the display. Markus [19] argued that two-dimensional 'artificial windows' (screens that presented nature scenes) are ultimately unsatisfactory as a view replacement because of a lack of dynamic depth cues. He states: "Another criterion for successful window design might be a dynamic one – i.e., the amount of change in the view that takes place for a given change in the viewing position of the observer. As a result of this movement parallax, not only do objects at a different distance within the view change their relative position, but also the window-view relationship changes. This is why two-dimensional artificial windows, even when very carefully contrived, are unrealistic and soon cease to satisfy; they lack the 'depth' within the view and the parallax of window aperture-view is also absent."

Of course, in the area of interactive computer graphics known as virtual reality, head-tracked or head-coupled displays have been in use since they were introduced in the 1960s by Ivan Sutherland, providing the user with the movement parallax cue. Later, head-tracked desktop systems [20], sometimes also referred to as fish-tank virtual reality [21], provided the user with a window-like view onto a computer-generated, virtual world. However, only when combining the elements of photorealism with appropriate viewpoint-dependent transformations of the displayed scene can a window-like 'see through' experience become convincing. However, with the limits in current tracking and rendering speeds, real-time interactivity still trades off against photorealism, making a fully interactive photorealistic views difficult to attain at present, particularly when the content presented on the window is captured 'live'.

1.2. Rationale of the current study

The aim of the current study was threefold. First, we wanted to investigate if we could create a convincing see-through experience using a simplified approach to generating motion parallax in relation to a photorealistic scene, that is, only transforming the relationship between the window frame and the outside view, based on head movements, without transforming the relation between objects contained within the view. Secondly, we wanted to investigate the potential effects of window framing, as the addition of a frame is expected to provide additional depth information regarding the position (depth layer) of the frame vis-a-vis the outside view, via the occlusion or interposition cue, particularly in the case where motion parallax is present. Thirdly, we wanted to investigate whether the addition of blur to the boundaries of the frame would add to the 'see-through experience' as it would signal to the visual system that the frame was located at a different depth layer than the view being displayed.

Recent work most similar in spirit to our own is that of Radikovic et al. [22], who created a window substitute using a tracked wall-mounted display. Using a repeated measures design, they had 14 students assess this interactive virtual window showing a nature scene against a static picture of that same scene. They found that the virtual window supporting motion parallax was considered superior as a window substitute, also having a stronger effect on well-being (positive mood) than the static picture.

Although the Radikovic study usefully demonstrated the added value of motion parallax in simulating a view from a virtual window, in line with Markus' [19] prediction, their experimental manipulation of a head-tracked versus a static image was quite a basic one. We wanted to investigate whether in addition to motion parallax, other monocular depth cues could be usefully deployed in creating a window substitute, giving rise to a more convincing illusion that one is looking through a window at an outdoor scene, rather than at a flat image projected on the wall.

Figure 8 Schematic representation of the experimental 2x2x2 design: 8 conditions varying in blurring of the frame, the presence of an occluding cross-shaped frame, and the presence of motion parallax

2. Method

2.1 Design

The effects of the three types of monocular depth cues on the reported see-through experience were tested in an experiment with a 2 (Motion parallax: off vs. on) x 2 (Occlusion: off vs. on) x 2 (Blur: off vs. on) x 5 (Image) within subjects design. Five different images were used as viewing scenes (see Figure 9), the remaining manipulations are represented in Figure 8.

2.2. Participants

Twenty persons (12 male, 8 female) participated in the experiment, with ages ranging between 19 and 42. All participants had a (corrected) visus of at least 1 and had little or no experience with perception experiments. All were employees or thesis students at Philips.

2.3. Apparatus and setting

The experiment was conducted in a dedicated perception lab at the Philips High Tech Campus. A virtual window prototype was created using a BARCO Reality 6400 beamer, placed under a table draped with black cloth to make the beamer less apparent. The images were projected 1,70 meters wide and 1,28 meters high on a 12m2 plane white wall at a resolution of 1280x960 pixels and 24bits colours. The virtual window thus covered 18% of the wall size, approximately in line with Keighley's [18] recommendations regarding preferred window size. Furthermore a chair was placed behind the table at 5 meters distance from the projected window, resulting in a horizontal viewing angle of 19.3° and a vertical angle of 14.6° and providing participants a desktop to work on.

A Polhemus PATRIOT system was deployed to keep track of the participant's head location. A fixed magnetic field generator in combination with a magnetic field receiver attached to a headphone determined the six degrees of freedom. The system had a refresh rate of approximately 120

Hz and a tracking latency of around 0.7 milliseconds. The end-to-end system latency was approximately 15 milliseconds. Pilot tests showed that when participants did not make highly accelerated head movements, motion parallax could be simulated without too much delay. The range of the transmitter/receiver was about 0.75 meters forcing the participants to remain seated during the experiment.

Custom software was engineered to interface the information supplied by the head tracker over the RST232 port of a PC. Another program was written to interpret the readings of the tracker and superimpose the occluding cross-frame. In a pilot experiment, the optimal gain factor (ratio between image-translation and head-translation) was determined at 0.58. When viewing the scenes, the lighting level in the laboratory was 40 lux (as measured on the desk) to ensure that the projected view would appear brighter than the laboratory and the black occluding cross would form a silhouette similar to that of a window frame.

2.4. Stimuli and monocular cue manipulations

Five different images were chosen, depicting a varied set of scenes and with varying distances at which interesting objects were displayed. The images are depicted in Figure 9.

Motion parallax was created by using digital photographs with a very high resolution, of which the virtual window only showed a small part. The picture was virtually placed at some distance behind the window frame, such that viewpoint-dependent transformations differed between the window frame and the outside view as a whole. Tracked head movements of the participant resulted in matched translations of the image (the picture being moved in the same direction as the head to display the correct view), thus simulating motion parallax.

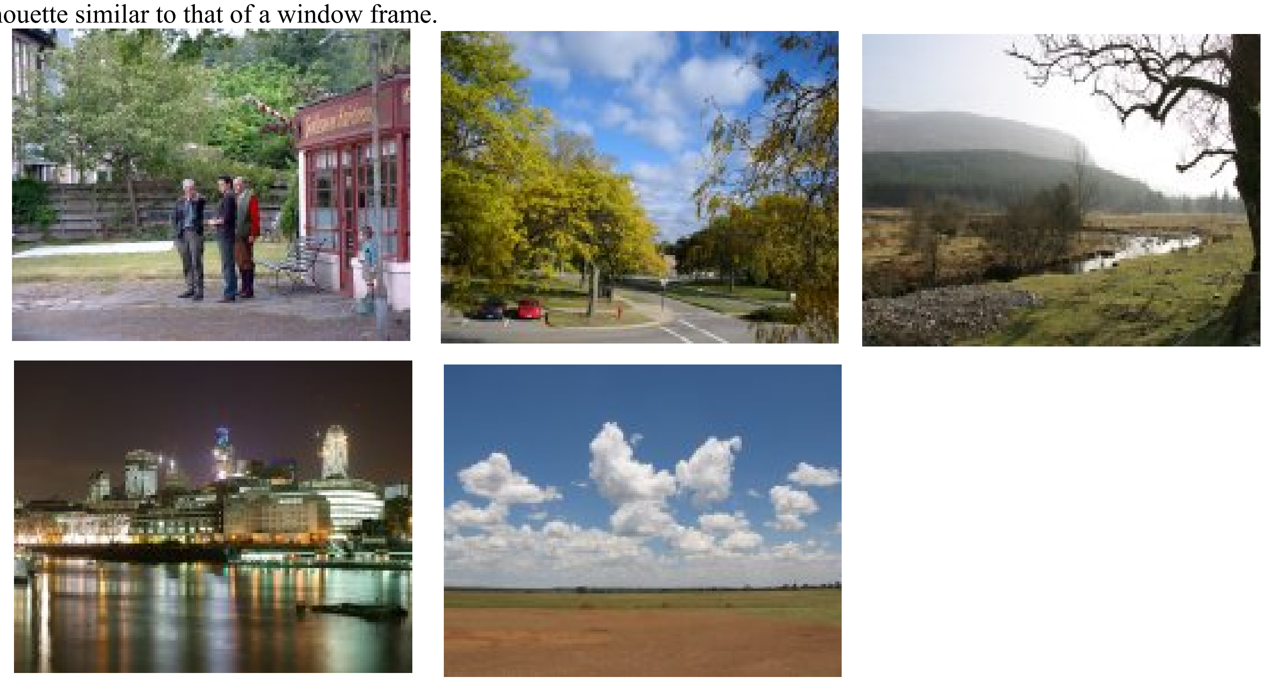


Figure 9 Five different images used as viewing scene: Hairdresser, First floor, Creek, Night Skyline, and Africa

Occlusion was implemented by superimposing a cross-shaped frame, as shown in Figure 10. The black bars of the cross were 5 centimeters wide, as measured on the wall.

Lastly, blur was manipulated by introducing a transparency gradient ranging from 0 to 1 over a distance of 1 centimeter starting at the edges of the frame and the edges of the cross.

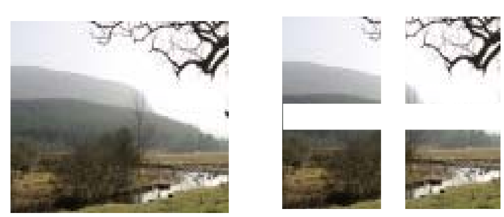


Figure 10 Visible scene, without and with occluding frame

2.5. Measurement

Participants were asked to rate their 'see-through experience', which was defined as 'the feeling that you are watching *through* a window, that is, the feeling that the view is beyond the "window" instead of a slide of a window-view on the wall'. They marked their assessment on the scoring form using the scale depicted in Figure 11. After the experiment the scores were measured with a ruler. The full length of the scale was given 5 points.

2.6. Procedure

Upon entering the room, participants were seated behind the desk and received written instructions to make moderate lateral head movements when viewing a new scene and to watch "out" of the window, not directly at the frame.

Once the participants finished reading the instructions, they placed a headphone with the head-tracker on their head and the experiment leader dimmed the light. The program first presented a training session with examples of views with different monocular depth cues, to allow participants to get used to the setting as well as the task and calibrate their use of the scale based on the range of variance between the different views.

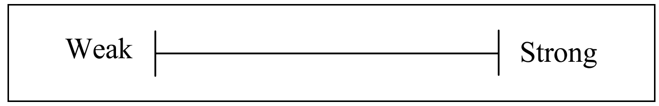


Figure 11 Scale used to assess 'See through experience'

The experiment leader stayed in the room during the training session to answer possible questions and to check whether the participants interpreted the instructions correctly. Participants were encouraged to use the full scale during the actual experiment. Then the experiment leader left the room and the participant commenced with the experiment. The order of the images and views was counterbalanced between participants. During the loading time of each new view an inter-stimulus adaptation field (ISAF) was displayed to eliminate influence from a previous stimulus due to inheritance.

Participants were offered a small token of appreciation (a lollypop) for their time. The experiment lasted approximately 15 minutes.

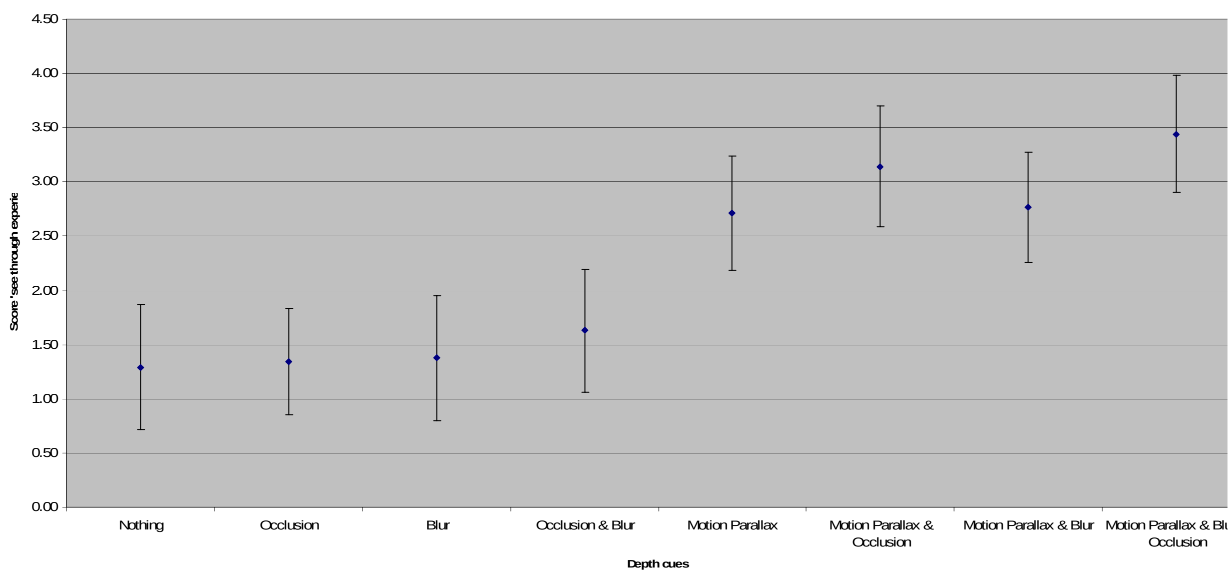


Figure 12 Mean scores and 95% confidence intervals for each of the 8 experimental conditions, averaged across participants and images.

3. Results

The effects of the three types of monocular cues on the reported see-through experience were tested in an experiment with a 2 (Motion parallax) x 2 (Occlusion) x 2 (Blur) x 5 (Image) within subjects design. The average scores for the five images in each of the experimental conditions are reported in Figure 12.

All three monocular cues show positive effects on the see-through experience. These results were tested in a Repeated Measures ANOVA, according to the full model. Motion Parallax showed a significant main effect ($F(1,19)=24.86$, $P<.001$). The experience without motion parallax was rated lower ($M=1.4$) than the experience with motion parallax ($M=3.0$). Although somewhat smaller, the main effect of Occlusion was also significant ($F(1,19)=8.70$, $p=.01$) .Viewing conditions with an occluding frame ($M=2.4$)

were rated higher than those without (M=2.0). The third main effect, of Blurred edges, also reached significance (F(1,19)=6.17, p=.02). Blurring the edges raised the see-through experience from M=2.1 to M=2.3. The last main effect, of Image, did not reach significance.

In addition, the interaction effect between Motion Parallax and Occlusion was significant (F(1,19)=4.71, p=.04), as well as the 3-way interaction between these two variables and Image (F(4,76)=3.68, p=.01. Further analyses showed that the effect of Motion parallax was enhanced for windows with an occluding frame compared to motion parallax without occluding frame for the ‘hairdresser’ image, but not for the remaining images.

Finally, the interaction between Occlusion and Blur almost reached significance (F(1,19)=3.88, p=.06), and the 3-way interaction Occlusion x Blur x Image was significant (F4,76)=2.49, p=.05. Occlusion was very effective for the ‘hairdresser’ and ‘first floor’ scene, less effective for the ‘creek’ and ‘Africa’ scene, and only effective in the ‘night sky’ scene when edges were blurred. No remaining effects proved significant.

4. Discussion

Taking the beneficial effects of windows as a point of departure, this paper presents an investigation on the contribution of three monocular depth cues, i.e., motion parallax, occlusion and blur, to the illusion that a wall-projected scene affords a window-like ‘see-through experience’. These cues were selected on the basis of an analysis of the role they are thought to play in the perception of scenes through a window. Additionally, the three cues chosen were computationally inexpensive, which allowed their implementation in a virtual window showing photorealistic images.

A controlled experiment was performed, manipulating the three monocular cues in a 2x2x2 within-subjects design. The results indicate that all three manipulations had a significant main effect on the ‘see-through experience’. The largest effect was produced by the motion parallax manipulation, in line with our own expectations as well as prior work [22]. Interestingly, the motion parallax effect was highly significant even though the implementation of motion parallax we used was a simplified one, only transforming the relationship between the window frame and the outside view, without transforming the relation between objects contained within the view. This is a promising result, as such a basic motion parallax simulation can be rendered in near to real-time, allowing people to look from their virtual window at any photorealistic scene, be they a static picture or a moving image. It is easy to conceive how such a window could provide a real-time view onto beautiful or exotic natural scenery through connections with various HDTV cameras placed around the world, much like the current widespread use of webcams that capture various kinds of environments, from homes, street corners, and offices, to the Eiffel tower or the African savannah. More realistic motion parallax rendering, such as those based on image-based rendering (IBR), as used by Radikovic et al. [22], are computationally more demanding and do not yet afford the near real-time rendering of such ‘live’ photorealistic imagery. It is an empirical question how our approach compares to IBR in terms of perceptual effect. Our prediction would be that for scenic views without foreground objects to speak of (e.g., views of mountains, deserts, sea, etc.), the simplified motion parallax approach will yield results similar to those that require more computational complexity. However, when foreground objects are salient, the simplified approach we used may provide cues that the surface one is looking at is, in fact, a flat 2D projection.

Although the effects of blur and the occluding frame were both much smaller in terms of effect size than motion parallax, their independent effects were significant nevertheless. This indicates that the ‘see-through experience’ we used as an indicator for how ‘window-like’ the simulation appeared, was indeed influenced by these monocular cues, in line with our expectations. Although none of the virtual window simulations we came across in our research apply selective blurring of the window frame as a depth cue, it appears that this cue yields a moderate effect on the realism of the simulated window, in particular signaling that the frame belongs to a different depth layer than the depicted view. However, we had not anticipated a strong effect of blur, as accommodation is known to be a fairly ineffective source of information for accurate depth discrimination [e.g., 23].

The superimposition of a cross-shaped frame across the entire view yielded a slightly larger effect than blur, but was particularly effective in conditions where motion parallax was present as well. The three-way interaction between Motion Parallax, Occlusion and Image was caused in particular by the ‘hairdresser’ image. This is quite understandable as the occluding frame partially obstructs the view onto the objects of interest in this image, being the people in front of the barber’s shop. Thus, without motion parallax, that is, the ability to look *around* the occluding frame, the view would be irritatingly blocked. More generally, the combination of motion parallax and the superimposed frame gave people a convincing illusion that a stable foreground reference frame was available through which a continuous environment in the background (the scene) could be viewed. This result suggests that providing additional framing in head-tracked virtual windows may enhance the illusion of a continuous environment in the background that is distinct from the wall of the room one is located in. This interpretation is in line with other work on foreground occlusion, in particular that of Mergner and Becker [24] in relation to vection (illusion of self-motion) and Prothero et al. [25] in relation to presence. Thus, in the motion parallax conditions, a stable foreground frame facilitates the perception that the background scene is independent of the window through which it is perceived and continuous beyond the boundaries of the frame. Additional framing then provides more evidence for the stability of the window frame in relation to the outside view.

Based on the results reported in this paper, one area of future research we would like to pursue is the application of virtual windows to settings where access to windows is problematic, as discussed earlier. We have already seen a number of commercial efforts in the healthcare domain. However, such companies typically provide little if any data relating to the clinical effectiveness of their 'healing windows', and usually base their claims on studies of the effects of real window views. It needs to be investigated whether virtual windows will have similar beneficial effects as real windows, and which aspects of the 'window experience' most crucially determine such effects (see also [26, 27]). Based on our results, motion parallax will likely be an important factor, but other factors we did not consider, such as lighting levels, may prove to be equally important.

In addition to applications in health-related environments, virtual windows have great potential for leisure and entertainment, for example as an advanced home theatre system or as an enjoyable view replacement which shows, to quote Basil Fawlty, "herds of wildebeests sweeping majestically across the plains" from a Torquay hotel window. Virtual windows can offer relaxing effects in stress-prone underground environments, such as subways or underground parking lots. For example, the IN-Visible system [28] shows subway travelers a projected view of the exterior urban environment at ground level that one is traveling underneath. Such a virtual subway window can enhance feelings of orientation, but can also make underground traveling much more enjoyable.

The study of artificial windows constitutes a useful case study for presence research, where fundamental and applied issues are intimately linked together. Though clearly delimited, the 'window experience' is quite rich and inherently multimodal. Although the research reported here has only touched upon the investigation of three particular visual cues relevant to window simulations, many other sources of sensory information are of relevance in creating a convincing window substitute, such as stereoscopic imaging, spatial audio characteristics, temperature, light, air quality and flow, and olfactory cues, to name but a few. What is particularly interesting about virtual windows, however, is that they are one of the few simulations where the people confronted with the simulation need not necessarily know in advance that they are looking at a mediated environment. One important characteristic, discussed in this paper, is that windows are generally bounded by a frame, turning one of the intrinsic limitations of most display systems into an advantage. When future virtual windows will use unobtrusive head tracking and will update their high-resolution photorealistic view accordingly in real-time, the next generation of trompe l'oeil artifices will have arrived, fooling both the eye and the mind of unsuspecting viewers looking out.

References

[23] M.C. Finnegan, L.Z. Solomon. Work attitudes in windowed vs. windowless environments. *Journal of Social Psychology,* 115. 291-292. 1981.

[24] E. Nagy, Working in Underground Offices. Doctoral Dissertation, Lund University. 1998.

[25] R.S. Ulrich, R.F. Simons, B.D. Losito, E. Fiorito, M.A. Miles, M. Zelson. Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, 11, 201-230. 1991.

[26] C.M. Tenessen, B. Cimprich. Views to nature: Effects on attention. *Journal of Environmental Psychology,* 15, 77-85. 1995.

[27] R. Kaplan. The role of nature in the context of the workplace. *Landscape and Urban Planning*, 26, 193-201. 1993.

[28] R. Kaplan, S. Kaplan. The Experience of Nature: A Psychological Perspective. New York: Cambridge University Press. 1989.

[29] R.S. Ulrich. View through a window may influence recovery from surgery. Science, 224, 420-421. 1984.

[30] J.H. Heerwagen, G. Orians. Adaptations to windowlessness: A study of the use of visual decor in windowed and windowless offices. *Environment and Behavior,* 18, 623-639. 1986.

[31] K.M.J. Farley, J.A. Veitch. A room with a view: A review of the effects of windows on work and well-being. NRC-CNRC Technical Report IRC-RR-136. 2001.

[32] http://www.trompe-l-oeil-art.com/

[33] http://www.tessusainc.com/

[34] J. Seligmann, L. Buckley. A sickroom with a view. Newsweek, 61. March 26, 1990.

[35] http://www.sensoryscapes.com/

[36] http://www.armas.com.tw/

[37] http://www.superdisplay.com.tw/

[38] http://www.research.philips.com/technologies/misc/homelab/

[39] W.A. IJsselsteijn. Presence in the past: What can we learn from media history? In: G. Riva, F. Davide, W.A. IJsselsteijn, (eds.). *Being There - Concepts, Effects and Measurements of User Presence in Synthetic Environments*, Amsterdam: IOS Press. pp. 17-40. 2003.

[40] E.C. Keighley, Visual requirements and reduced fenestration in offices: A study of multiple apertures and window Building Science, 8, 321-331. 1973.

[41] T.A. Markus. The function of windows: A reappraisal. *Building Science,* 2, 97-121. 1967.

[42] S. Fisher. Viewpoint dependent imaging: An interactive stereoscopic display. *Proceedings of the SPIE*, 367, 41-45. 1982.

[43] C. Ware, K. Arthur, K. Booth. Fish tank virtual reality. *Proceedings of the INTERCHI '93 Conference on Human Factors in Computing Systems*, 37-42. 1993.

[44] A.S. Radikovic, J.J. Leggett, J. Keyser, R.S. Ulrich. Artificial window view of nature. *CHI 2005* Late Breaking Results, 1993-1996. 2005.

[45] J.E. Cutting, P.M. Vishton. Perceiving layout: The integration, relative dominance, and contextual use of different information about depth. In: W. Epstein, S. Rogers (eds.), *Handbook of Perception and Cognition: Vol. 5:*

Perception of Space and Motion. New York: Academic Press. pp. 69-117. 1995.

[46] T. Mergner, W. Becker. Perception of horizontal self-rotation: Multisensory and cognitive aspects. In: R. Warren, A.H. Wertheim (ed.s.). *Perception and Control of Self-Motion*. Hillsdale, N.J.:Lawrence Erlbaum Associates. 1990.

[47] J. Prothero, H.G. Hoffman, D.E. Parker, M. Wells, T.A. Furness. Foreground/background manipulations affect presence. *Proceedings of Human Factors and Ergonomics Society*, pp. 1410-1414. 1995.

[48] A.-M. Sponselee, Y.A.W. de Kort, A.L. Meijnders. Healing Media: The moderating role of presence in restoring from stress in a mediated environment. *Proceedings of PRESENCE 2004*, 197-203. 2004.

[49] Y.A.W. de Kort, W.A. IJsselsteijn. Reality check: The role of realism in stress reduction using media technology. *CyberPsychology and Behavior,* 9, 230-233. 2006.

[50] P. Chatzitsakyris, G. Ducla-Soares, A. Zulas. IN-Visible: Perceiving invisible urban information through ambient media. *CHI 2004* Late Breaking Results, 1348-1350. 2004.

Haptic Thermal Interface: A New Technology for Supporting Presence in Multimodal Virtual Environments?

Chang Nam
University of Arkansas
{cnam@uark.edu}

Abstract

This paper describes how the thermal user interface system can be integrated into other input devices to develop multimodal user interfaces that present multisensory feedback. Multimodal virtual environments supporting both haptic force and haptic thermal feedback will allow researchers to investigate a wide range of research issues involved in presence study. Among presence-related research questions are 1) what role the additional thermal feedback has for presence in multimodal virtual environments, 2) whether it affects task performance in various task environments, 3) what kinds of design principles and guidelines of multimodal user interfaces should be considered to improve the user's perceived presence, and 4) how those principles and guidelines can be applied to the use of multimodal user interfaces in various applications.

1. Introduction

The use of haptic devices as an interaction tool provides users with a new and interesting sensorial experience, *the sense of touch*, through tactile and force feedback. Despite such powerful sensory inputs as tactile and force feedback, however, haptic interface systems are still in an early stage of accomplishing a high degree of realism. One of the haptic elements missing is the ability to present thermal information such as the thermal conductivity and temperature of an object being manipulated. Thermal feedback should be incorporated into haptic interface systems and multimodal virtual environments to deliver a more convincing and intuitive presence in virtual environments or teleoperation systems.

Much more studies are still needed to understand how thermal modality can be used to increase richness of sensory information [1, 2] in virtual environments. It would be interesting to investigate what role the additional thermal feedback has for presence in virtual environments, and whether it affects task performance in various task environments. In addition, more studies need to be done on the effects of multimodal virtual environments supporting multiple modalities including visual, audio, haptic force, and haptic thermal on perceived presence. To help address these two issues, the HCI research team at the University of Arkansas has developed and evaluated thermal interface systems that can present thermal information of an object being manipulated to the human operator interacting with VEs accurately and with no overt time delay [3, 4].

This study reports our ongoing efforts to develop advanced multimodal user interfaces for virtual environments that can provide the user with multiple sensory feedbacks. First, this study will demonstrate how the thermal interface system developed in my previous study [3] can be integrated with other input devices to present multiple sensory feedbacks including thermal feedback. This study concludes with a description on a series of empirical experiments that we are planning to conduct in order to investigate what role the additional thermal feedback has for presence in multimodal virtual environments, and whether it affects task performance in various task environments.

2. Multimodal Interface: Visual-plus-Thermal

***Integration with optical mouse*:** The manipulator end of the thermal interface system can be integrated into a computer mouse to develop a multimodal user interface that can provide both visual and thermal feedbacks to the user interacting with virtual environments. Considering the structure of the mouse system and the way the user operates it, there are two mechanisms to integrate the thermal feedback system onto the mouse: 1) the thumb and the ring finger touch (finger-strategy); 2) the palm touches (palm-strategy). Figure 1 shows an example of the finger-strategy integration, which allows users to position the manipulator where their thumb would lie with normal use of the mouse. This creates a situation where they could use the mouse to initiate thermal change and then could be sensed on the same hand without releasing our grip or moving our hand.

To evaluate the effectiveness of the multimodal user interface system for virtual environments, a three-dimensional desktop virtual environment application was developed in which the user can feel the warmth of objects when touching. For the virtual environment, an office area model was used, consisting of three rooms and various objects such as office furniture and equipments users would expect to find.

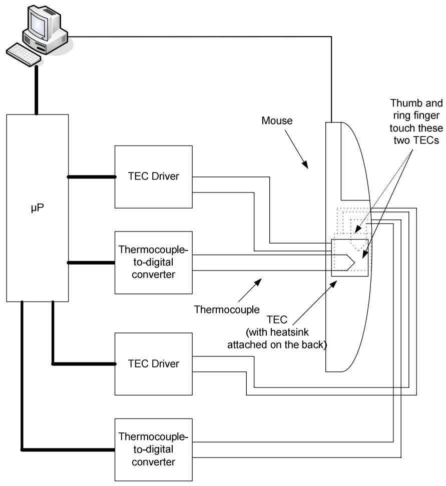


Figure 1. Structure of finger-strategy integration

Temperature changes would be committed when the multimodal interface system entered a 3D object. Our previous work showed promising results that additional thermal feedback helped participants' object identification, as well as enhanced their perceived presence in a 3D desktop virtual environment [4].

4. Multimodal Interface: Thermal-plus-Haptic

Integration with haptic system: The manipulator end of the thermal system can also be integrated into an existing haptic system (e.g., PHANToM Omni, SensAble Technologies). The user then can hold a spherical object and feel both force and thermal feedback while interacting with the VEs. A 3D desktop VE application was also developed to evaluate the system's effectiveness, in which the user can receive thermal as well as haptic feedback (Figure 2).

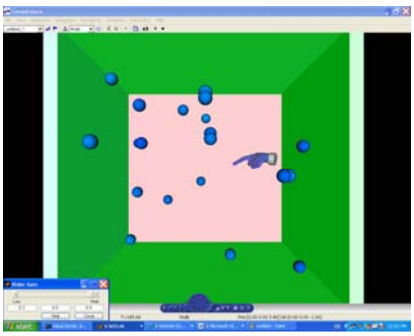


Figure 2: Haptic virtual environment application

Using PHANToM Omni (SensAble Technologies) and proSENSE software (Handshake VR Inc.), we have developed an application in which, through the multimodal user interface, the user can feel both changes in temperature of an object and the movement of its atoms (represented as haptic spheres) as its temperature changes, when experiencing physical properties of temperature and force [5].

Conclusions

This paper described how the thermal interface system can be integrated into other input devices to develop multimodal user interfaces that present multiple sensory feedbacks. Multimodal virtual environments supporting both haptic force and haptic thermal feedback allow researchers to investigate what the addition of thermal feedback and other sensorial feedbacks would contribute to the sense of presence and task performance in multimodal virtual environments.

Our future plan is to conduct a series of empirical experiments in order to investigate a wide range of research issues involved in presence study. Among presence-related research questions are 1) what role the additional thermal feedback has for presence in multimodal virtual environments, 2) whether it affects task performance in various task environments, 3) what kinds of design principles and guidelines of multimodal user interfaces should be considered to improve the user's perceived presence, and 4) how those principles and guidelines can be applied to the use of multimodal user interfaces in various applications.

During the past decades, we have assumed a strong relationship between the number of channels/sensory cues and the degree of presence. As Lombard and Ditton criticized [6], however, the relationship has not been fully investigated. The research questions proposed in this paper, which can be examined by using the multimodal user interface system supporting multiple sensory feedbacks, will help to prove the assumption.

References

[1] R.L. Daft, R.L. Lengel. Information richness: a new approach to managerial behavior and organization design. *Research in Organizational Behavior*, 1984, 6, 191-233.

[2] R.L. Daft, R.L. Lengel. Organizational information requirements, media richness and structural design. *Management Science*, 1986, 32(5), 554-571.

[3] C.S. Nam, J. Di, L.W. Borsodi, W.A. Mackay. A haptic thermal interface: Towards effective multimodal user interface systems. In *Proceedings of IASTED-HCI*, ACTA Press, 2005, 13-18.

[4] C.S. Nam, D. Chung. The influences of haptic thermal feedback on object identification and perceived presence. In *Proceedings of the 16th World Congress of the International Ergonomics Association (IEA)*. 2006. Maastricht, Netherlands: IEA Press (To appear).

[5] C.S. Nam, I. Shafieloo. Haptic virtual environments as a science learning supporting tool: A Preliminary Study. In *Proceedings of the IASTED International Conference on Educational Technology*, ACTA Press, 2006, 40-44.

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

"I Don't Like William Touching My Belly": Gender Differences in Affective Responses to Mediated Social Touch

Christiaan de Nood, Antal Haans & Wijnand IJsselsteijn
Eindhoven University of Technology, Eindhoven, The Netherlands
{A.Haans@tm.tue.nl}, {W.A.IJsselsteijn@tm.tue.nl}

From behind the screen where I hid,
I advance personally, solely to you.
Camerado! This is no book,
Who touches this, touches a man,
(Is it night? Are we here alone?)

- Walt Whitman (1819 - 1892)

Abstract

This poster presents a first exploration to investigate whether the gender differences generally found in unmediated same- and opposite-sex social touch are also present in mediated situations. Such response similarities between real and virtual touch will have implications for the design of haptic communication devices as well as for the study of touch itself. In an experiment, participants were led to believe that a (male or female) stranger was remotely touching them. We presented different types of tactile stimulus patterns to several loci of a participant's upper body through a vest equipped with vibrotactile actuators. Results indicated that affective responses varied with the stimulated body location ($\eta^2 = 9.0\%$; $p < .01$,) and the type (i.e., "stroke" or "poke"; $\eta^2 = 0.4\%$; $p = .06$,) of mediated touch. The effect of dyad composition (i.e., same- vs. opposite-sex), although non-significant, was larger for the male sample than for the female sample, explaining respectively 3.8% ($p = .16$) compared to 0.4% ($p = .67$) of the variance in the total sample. In separate ANOVAs for the male and female sample, the effect of dyad composition did not reach significance, with respectively $\eta^2 = 9.9\%$ ($p = .11$), and $\eta^2 = 0.1\%$ ($p = .73$).

In correspondence with research on unmediated social touch, the effect of dyad composition, although not significant, was about ten times larger for our male participants than for our female participants. Overall, the results contribute tentative evidence for the contention that mediated touch is being perceived in similar ways to unmediated touch, and that touch-like qualities may be attributed to stimulation from an array of vibrating electromechanical actuators.

Keywords--- **physical contact, computer mediated communication, haptic feedback, social presence, human sex differences**

The Effects of Presence and Tactile Illusion on Consumers' Attitudes and Intentions: The Mediating Role of Mental Imagery and the moderators' effects

Yung Kyun Choi
Dongguk University
{choiyung@dgu.edu}

Abstract

This paper investigates the relationship between presence, tactile illusion, vividness of imagery and advertising effectiveness. A total of 207 undergraduate students participated for an experiment. The result of path analysis found that presence and tactile illusion induce more vivid mental imagery. Vividness of mental imagery was a major contributor to consumers' attitudes toward the website. Spatial presence and realness had a partial impact on advertising effectiveness. There were main effects of site type and NFT on some of advertising effectiveness measures with interaction effects for product type and site type, and for product type, site type, and NFT. Theoretical reasoning and practical implications for the results are followed.

Keywords--- ***Presence, cross-modal illusion, mental imagery, 3D advertising, advertising effectiveness***

1. Introduction

Despite the importance of touching in consumers' decision making, this sense is not usually available when we try online shopping. It is possible to virtually experience a product by interacting with 3D graphic images instead. Previous researches [1, 2] show that a sense of presence is an important mediator in this virtual product experience and have a positive impact on web advertising effectiveness. It is also found that, in a virtual environment, presence stimulates synesthesia[3]. Will it be possible that this cross-modal illusion have a compensating effect for deficient sensory cues (e.g., haptic information) in a virtual store? The purpose of this study is to investigate the relationship between presence, tactile illusion, and advertising effectiveness. The paper also examines the difference in advertising effectiveness by product type, site type, and consumer's need for touch (NFT).

2. Presence and cross-modal illusion

Developing mental model in a virtual environment is greatly influenced by the way human process multimodal information. The human sensorimotor system works for an object or an event simultaneously and dynamically. It perceives incoming sensory stimuli from different modalities as a whole not as separate [4]. The multisensory information processing incurs cross-modal interaction and perceptual illusion (e.g. ventriloquism). Both presence and synesthesia may be phenomenal byproducts of the same intermodal integration process, the building of a cognitively consistent mental model of the environment [3]. Therefore, they will be correlated each other.

3. Presence, imagery, and persuasion

Presence experience results from the interpretation of the mental model of the virtual environment. Vivid elements in virtual environments are basically attention getting and they activate relevant information in memory [5] to fill the missing information from the sensory stimuli. When imagery is a mental event visualizing something, presence from concrete mental model may lead to more vivid or richer mental imagery. Recent researches show that presence develops virtual experience to positively influence attitudes and intentions [1, 2]. Presence is also enhanced by 3D advertising and influences product knowledge, brand attitude, and purchase intention by creating a compelling virtual product experience [1]. One interesting finding in their research was that when evaluating material product[9] , Jacket in the study, 3D advertising resulted in more favorable brand attitude and better product knowledge than 2D condition. The beneficial effects of imagery on persuasion has been summarized in 3 ways: (1) by possibly drawing information stored in long term memory into working memory; (2) by anchoring in the personal experience to make it more relevant; (3) by providing multisensory experience [6].

4. The effects of moderating variables

Products differ in the extent to which they possess salient material properties that correspond to texture, hardness, temperature, and weight information [7]. Product categories in which the material properties vary in a diagnostic manner are more likely to encourage touch. In addition to product-based sources of salience for haptic information, the haptic salience is also likely to depend on the person. For example, more haptically oriented consumers with high NFT, consider material properties earlier during product evaluation and have greater chronic accessibility to haptic information [8]. Mode of presentation

[9] Objects with attributes that require touching for understanding are called material products. Comparatively, objects with attributes that can be fully understood through vision are called geometric objects [1, 7].

(e.g., 3D vs. 2D) may do a similar role as it affects presence [1]. Therefore, moderating effects of these variables (product type, consumer's NFT, and presentation format) will be tested in the experiment.

5. Method

For an experiment, 2 (presentation format: 2D vs. 3D) x 2 (product type: watch vs. jacket) × 2(consumer trait: high vs. low NFT) factorial design was employed. A total of 207 undergraduate students (105 males: 50.7%) from a large university in Seoul participated as subjects. The experiment used two different products: watch and jacket. Watch was selected as a geometric product and Jacket as a material one (for this classification, see [1]). Web sites with 2D and 3D format were developed for each product and 3D images were aided by interactive features such as zooming, moving, and rotating. Presence was measured by 3 components: spatial presence, involvement, and realness [9]. Measures for the tactile illusion were developed based on the reasoning from the previous research [7]. All other measures came from previous studies. Most of measures resulted in high reliability ranging from .83 to .96, excepting the realness of presence (.61).

6. Results

Structural equation model was revised until the goodness of fit indices stayed at acceptable level. The major contributor for site attitudes was the vividness of mental imagery that was influenced by spatial presence(β=.17), realness(β=.16), and tactile illusion(β=.14). Tactile illusion was highly influenced by realness(β=.55) but not by spatial presence. Entertaining site attitude was affected by spatial presence(β=.16) and brand attitude by realness (β=.27). As in dual mediation model, brand attitude mediated the effects of site attitudes to purchase intention.The result of multivariate analysis showed significant main effects of site type (Wilks's λ=.78, F=7.68, $p<.01$) with interaction effects between product type and site type (Wilks's λ=.92, F=2.28, $p<.05$). Univariate analysis of variance showed significant main effects of site type on site ent. (F=17.36, $p<.01$), site info (F=14.82, $p<.01$), and VI(F=14.93, $p<.01$). Two-way Interaction for product type and site type is significant on site ent. (F=4.05, $p\leq.05$), brand attitude (F=4.61, $p\leq.05$), PI (F=11.47, $p<.01$), and VI (F=4.71, $p\leq.05$). The main effects of need for touch was significant only on site ent. (F=3.72, $p\leq.05$). Three-way interaction for product type, site type, and need for touch was significant on VI.

7. Conclusion

The vividness of mental imagery is found to be a key resource to persuade consumers in a virtual environment. It directly influenced website attitudes. The result vindicates that imagery is a powerful tool for changing the state of mind and body. Realness contributed a lot for tactile illusion, while spatial presence and involvement did not have any significant effects on the illusion. Spatial presence had a positive effect on vividness of imagery, entertaining attitude toward the site but not on informative site attitude. Mental imagery was composed by presence and cross-modal illusion elements.

Overall, the research found that 3D advertising is more effective than 2D in persuading consumers' attitude and intention in the website. However, due to the interaction for product type and site type, 3D advertising was more effective in geometric products than in material goods. The effects of consumer traits were quite limited. Consumers' NFT had marginally significant effects on entertaining attitude toward the site. By interacting with product type and site type, NFT affected only revisiting intention.

Acknowledgements

The research has been supported by the Korea Research Foundation Grant funded by the Korean Government (KRF-2004-041-B00732).

References

[1] Li, H., Daugherty, T., & Biocca, F.(2002). Impact of 3-D advertising on product knowledge, brand attitude, and purchase intention: the mediating role of presence, *Journal of Advertising, 31*(3), 43-57.

[2] Coyle, J. R. & Thorson, E.(2001). The effects of progressive levels of interactivity and vividness in web marketing sites, *Journal of Advertising, 30*(3), 65-77.

[3] Biocca, F., Kim, J., & Choi, Y.(2001). Visual touch in virtual environments: an exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions, *Presence, 10*(3), 247-265.

[4] Soto-Faraco, S., Spence, C., Lloyd, D., & Kingstone, A.(2004). Moving multisensory research along: motion perception across sensory modalities, *Current Directions in Psychological Science, 13*(1), 29-32.

[5] Nisbett, R. & Ross, L. (1980). *Human inference: strategies and shortcomings of social judgment*, Englewood Cliffs, NJ: Prentice-Hall.

[6] Burns, A., Biswas, A., & Babin, L.(1993). The operation of visual imagery as a mediator of advertising effects, *Journal of Advertising, 22*(2), 71-85.

[7] Klatzky, R., Lederman, S., & Matula, D.(1991). Imagined haptic exploration in judgments of object properties, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(March), 314-322.

[8] Peck, J. & Childers, T. (2002). Individual differences in haptic information processing: on the development, validation, and use of the Need for Touch scale, working paper, Department of Marketing, University of Wisconsin-Madison.

[9] Schubert, T., Friedmann, F., & Regenbrecht, H.(2001). The experience of presence: factor analytic insights, *Presence, 10*(3), 266-281.

An Exploration of Clinicians' Sense of Presence in Critical Care Telemedicine

Leila Alem, Susan Hansen, Jane Li
ICT Centre; CSIRO, Australia
{Leila.Alem@csiro.au, Susan.Hansen@csiro.au, Jane.Li@csiro.au}

Abstract

The Virtual Critical Care Unit, (ViCCU®), is a telemedicine system that allows a specialist at a major referral hospital to direct a team in another, usually smaller and remote hospital. In this study we used a modified version of the Slater-Usoh-Steed (SUS) presence questionnaire to measure clinicians' sense of presence when using ViCCU®. We also explored the relationship between presence experienced when using ViCCU® and personal, usability and media factors. Results indicate that in this context, personal factors influenced clinicians' experience of presence and that a positive relationship between presence and both usability and media factors may exist. Reflections on the appropriateness of the SUS presence measure in this real-world setting are also included.

1. Introduction

Telemedicine is the delivery of healthcare over a distance and for years telemedicine technologies have enabled off-site clinicians to provide healthcare to patients in remote locations. In 2002 a telemedicine application for a critical care setting, the Virtual Critical Care Unit (ViCCU®), was developed in a collaboration between CSIRO through the Centre for Networking Technologies for the Information Economy (CeNTIE) and Sydney West Area Health Service. It was installed in Katoomba and Nepean Hospitals, Australia, in December 2003 for a 2 year clinical trial.

ViCCU® is composed of two main stations: a remote station, Figure 1(a), located at Katoomba Hospital and a specialist station, Figure 1 (b), located at Nepean Hospital.

(a) The remote station (b) The specialist station

Figure 7 The Virtual Critical Care Unit

Using the stations allows a specialist at a major referral hospital and clinicians at a remote hospital to work together as one team to treat and diagnose patients at the remote hospital by transmitting multiple channels of real-time video/audio information of the patient, the clinical team, x-ray/paper documents and patient vital signs from the remote site to the specialist [1].

The literature describes numerous attempts to measure presence in the medical area when using virtual reality or augmented reality, however these have typically been performed in laboratory environments [e.g. 2]. This paper goes further and attempts to measure presence in a real-world clinical setting and explore its relationship with other factors measured in the study.

2. The Study

One of the aims of ViCCU® was to give clinicians located in a remote hospital the feeling that a specialist was physically 'there' at the end of the bed. For this reason we decided to focus our study on physical presence.

As access to the emergency clinicians was limited it was important that the presence measure chosen could be administered easily and quickly. Following a review of existing subjective measures of physical presence, the Slater-Usoh-Steed (SUS) questionnaire [3] was chosen and adapted for this context. The SUS questionnaire was adapted to assess presence using ViCCU® by:

- Replacing direct reference to an experience within a virtual environment with experience with using ViCCU®.
- Generating two versions of SUS questionnaires, one for the doctors and nurses in Katoomba (containing 4 out of the 6 items) and one for the staff specialists in Nepean (containing 5 out of the 6 items).

The adapted versions of the SUS questionnaire were then added to the Katoomba and Nepean Technical Evaluation Questionnaire's. In addition to measuring presence using the SUS questionnaire, we decided to explore the construct of presence specific to three factors: personal – employment category, gender, prior experience with videoconferencing; usability – ease of use, ability to focus on patient, satisfaction with overall design; and media factors – overall satisfaction with video and audio.

Fifty clinicians in total (36 hospital staff from Katoomba and 14 staff specialists from Nepean) took part in this study.

3. Results

A reliability analysis was performed, and the Cronbach's alpha calculated. As sufficiently high alphas were obtained, 0.808 for the Katoomba presence questions and 0.918 for the Nepean presence questions, the items were summed. Spearman correlation tests were performed on non-recoded data to test whether there was a relationship between presence scores, usability and media factors. T tests were used to detect group differences. Results are described below.

Personal Factors Katoomba nurses had significantly higher presence scores (m=18.38, s=5.608, p=0.018) than Katoomba doctors (m=13.64, s=5.329).

Katoomba females reported significantly higher presence scores (m=18.50; p=0.018) than males (m=13.93).

In Nepean, this was reversed with male specialists mean presence scores (m=22.00) higher than females (m=14.67), however this was not significant at the 0.05 alpha level.

In both hospitals, although there were differences in the mean presence scores of those who had prior videoconferencing experience and those that did not, this was not statistically significant.

Usability Factors With the exception of a moderate positive correlation between Nepean presence scores and ease of use (r=0.553), no other usability factors (i.e. ability to focus on patient and satisfaction with overall design) significantly correlated to the overall presence scores.

Media Factors There were no significant correlations between Katoomba presence scores and overall impression of video or audio quality. Interestingly, a subtle negative relationship was observed between the mean presence scores of those who were dissatisfied (m=21.00), neutral (m=17.27) and satisfied (m=16.09) with the overall video quality.

In Nepean, there was a significant positive moderate correlation between presence scores and overall satisfaction with video quality (r=0.549). Although not statistically significant, Nepean specialists who were satisfied with the ViCCU® audio displayed higher mean presence scores (m=21.70) than those who were neutral (m=17.25).

Comparison to Specialist Physically Present Of particular interest to us in this study was how using ViCCU® compared to having a specialist physically present in the room and whether indeed a comparison could be made.

The majority of Katoomba staff reported that using ViCCU® was the same (53%) or better (19%) than the specialist being physically present, and only 28% reported using ViCCU® was worse.

From the specialists station however, the majority of staff specialists reported that using ViCCU® was worse (62%) than being physically present and approximately a third (31%) thought it was the same and only 8% reported that it was better.

4. Discussion

Clinicians reported high levels of presence when using ViCCU®. The study revealed that presence experienced by the clinicians was influenced by some personal factors including employment category and gender. There was some evidence to support a positive relationship between presence and satisfaction with usability and media factors. There was also evidence to suggest that the high levels of satisfaction with ViCCU® expressed by the clinicians could be influenced by the high levels of presence experienced when using ViCCU®, making it comparable to actually 'being there'. The majority of Katoomba staff felt that using ViCCU® was the same or better than the specialist physically being there, but the majority of Nepean staff felt that using ViCCU® was worse.

Whilst the SUS questionnaire may not be as comprehensive as other presence measures, it has shown itself to be a context flexible and an adaptable measure of presence. The major advantage of the SUS questionnaire, particularly in this time-critical telemedicine context, was the length of the questionnaire which lent itself to be easily integrated into a larger questionnaire and could be completed quickly. However, the short length also meant that it may not have necessarily been able to capture the wide range of elements that contribute to presence and provide a more accurate measure. To investigate this, it would be interesting to conduct a study using an adapted SUS questionnaire and a more comprehensive instrument such as the ICT-SOPI [4] in a clinical setting and compare the results.

This study supports the notion that it is possible to measure presence in telemedicine applications; that the SUS presence measure can be used in a real-world clinical context; and that it is sensitive enough to allow the investigation of the influence of different factors such as personal, usability and media factors on presence. While acknowledging the limitations, this study also represents a positive step in measuring presence in telemedicine applications in a clinical context.

References

[1] R.Y. Qiao, K. Bengston, A. Krumm-Heller, M. Hogan. A Critical Care Telemedicine System on Broadband IP network. In *Proceedings of Asia-Pacific Conference on Communications* 852-856. October 2005.

[2] M.C. Juan, D. Perez, D. Tomas, B. Rey, M. Alcaniz, C. Botella, R. Banos. An Augmented Reality system for the treatment of acrophobia. In *Eighth Annual International Workshop on Presence* 315-318. September 2005.

[3] M. Slater. 1999. Measuring presence: a response to the Witmer and Singer presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 8, 560-566.

[4] J. Lessiter, J. Freeman, E. Keogh, J. Davidoff. 2001. A cross-media presence questionnaire: the ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, 10, 282-297.

Presence after Death

Matthew Lombard, Melissa E. Selverian
Temple University
{lombard@temple.edu, MelissaSelverian@comcast.net}

Abstract

This paper examines some of the increasingly sophisticated attempts that humans make to evoke the presence of themselves or others after death. Research possibilities and ethical implications are addressed.

1. Introduction

Presence scholars have connected presence with many diverse aspects of our lives including art, entertainment, business, education and training, physical and psychological health care, and sexuality, but they seldom if ever have considered the concept in relation to our deaths. As a species, we have long used technology, first in the form of drawings, statues and grave markers, to evoke the sense of the presence of a person who is no longer living; but at the beginning of the 21st century, technology permits very sophisticated artificially intelligent and perceptually realistic replications of people who have died. Here we briefly consider technologies that can evoke presence after death, as well as some research opportunities and ethical implications that they suggest.

2. Beliefs and behaviors regarding death and bereavement

Because it is a universal, mysterious, and often disturbing phenomenon, we are naturally interested in death and develop complex beliefs and behaviors regarding it.

Throughout the centuries and the world, a variety of religious, philosophical and psychological theories and beliefs (e.g., life after death, resurrection, reincarnation and immortality [1]) have helped survivors of the deceased feel some sense of closeness with the departed, while shamans, mediums, clairvoyants and "channelers" have claimed to let us communicate with the dead.

"Until the twentieth century, maintaining a bond with the deceased had been considered a normal part of the bereavement process in Western society. In contrast, in the twentieth century the view prevailed that successful mourning required the bereaved to emotionally detach themselves from the deceased" ([2], Detachment Revisited). The current view is that continuing bonds with the deceased is normal and healthy [2].

3. Using technology to evoke presence after death

The living have often relied on technology – defined broadly as "a machine, device, or other application of human industrial arts… includ[ing] traditional and emerging electronic media… and traditional arts such as painting and sculpture" [3] – to help them attain some sense of physical and social presence or connection with the departed.

3.1. Treatment of the body

In many cultures an after-death ritual is the viewing of the body, which is presented to appear as close to the physical likeness of the deceased as possible in order to evoke, among other things, a sense that he/she is present with the survivors [4]. America in the late 1800s was said to be 'Egyptianized' as airtight coffins and arterial-injection embalming promised (but couldn't deliver) permanent preservation of the body [5]. Today thousands of people add their names to the donor roster for the controversial Gunther von Hagens' Body Worlds museum displays that feature "Anatomical Exhibition of Real Human Bodies" [6]. Pet taxidermy "allows pet owners to see, touch and hold their pets, and in a sense, "never have to let go" [7].

3.2. Grave markers, epitaphs and other memorials

The marking of the location that a person has been buried is another way to evoke the presence of that person. Epitaphs ("on the gravestone" in Greek) further identify and help evoke a sense of the deceased. Other types of memorials include roadside memorials, monuments, fountains, buildings named after the deceased, trees planted in their names, even less tangible objects such as endowed academic positions, scholarships, etc.

3.3. Paintings, statues and other art

In many cultures we create or commission drawings, paintings, photographs, sculptures, and statues to provide a sense of closeness with deceased family members as well as national figures.[10] Among the more interesting art forms are

[10] This paper was inspired in part by the effectiveness of a

postmortem paintings and photography [8] [9]. We also recreate the presence of the dead with celebrity look-alikes, impersonators and tribute artists; stage shows (e.g., *Mark Twain Tonight*; theatrical productions (e.g., *Beatlemania*), and "living history presentations" [10].

3.4. Traditional print and electronic media

Americans and others often keep diaries, scrapbooks, photo albums, and audio and video recordings to maintain a connection with family members, living and dead. They write and read eulogies and obituaries after the passing of a loved one and assemble biographies to memorialize the deceased. Families often compile and record oral histories to keep the sights and sounds of the deceased present in the family life.

3.5. New forms of electronic media

Emerging technologies seem likely to evoke the presence of the dead in ways that are even more physically, socially and psychologically vivid and "real" than have traditional technologies.

Memorials sites on the Internet, for instance, combine many features of the earlier media in a convenient, accessible and multisensory venue.

Other new technologies may bring audio recordings and holographic images to gravesites to evoke the presence of the deceased [11] [12].

A disturbing application of new technology for staying connected with the deceased involves installing cameras in caskets before burial and webcasting the images [13].

Evolving film technologies bring deceased actors back to the living, seeming to interact with modern actors and to "[speak] lines [they] never spoke and [make] gestures [they] never made" [14]. Sophisticated techniques of digital manipulation and presentation have even been used to recreate the experience of a live concert by a deceased artist (e.g., "Sinatra at the London Palladium" [15]). Animatronics (mechanized puppets) is used to evoke the presence of American presidents at Disneyworld [16].

The company All Digital has introduced DiNA by Lynn Hersman Lesson, a virtual "person" who comes to "life" on a wall-sized computer screen through artificial intelligence technology [17]. The enormous image of a woman's face nods in anticipation of a conversation and then engages in it. DiNA can conceivably be designed to be whoever a user wants her to be, living or dead.

Hanson Robotics has created an android replication of deceased science fiction writer Philip K. Dick. It has a real-to-life shell and is programmed with an "intellect" via artificial intelligence and "personality" via a mathematically derived extraction of the author's life. It tracks faces with cameras inserted in its eyes and recognizes people in a crowd and perceives expressions [18].

Hiroshi Ishiguro of ATR Intelligent Robotics and Communication Laboratories has created the android Geminoid HI-1, a "silicone-and-steel doppelganger" that is an "exact duplicate" of himself [19].

Meanwhile, Microsoft [20] and the U.S. Government [21] are exploring ways of collecting our individual experiences. It's not difficult to imagine a day when we'll store nearly all of our experiences in a computer and then install them in an artificially intelligent android that looks and acts nearly exactly like us as it interacts with our family and friends after we're gone.

4. Ethical implications

Some of the modern means of evoking the presence of the dead – e.g., the webcam in the casket – raise obvious ethical concerns. The more intriguing issues stem from our growing ability to recreate a person in their living form, as we move beyond simple text, drawings, photographs and even film and video recordings all the way to interactive and intelligent avatars and vivid, artificially intelligent androids. When they are widely available, how will we use them and what effects might they have? Will we choose to make realistic or idealized versions of the deceased? Would it be healthy to live with 'people' who are gone? Could they eventually affect us and the world as do the living? Is using these technologies to recreate ourselves for after we're gone somehow against the 'natural order'? What are the ethical implications of lives 'lived' forever?

5. Research directions and conclusion

We should study which current technologies, and characteristics of technologies such as realism, vividness, movement, interactivity, immersiveness, etc., are more and less effective in evoking the sense of presence of someone who has died, and which are most comforting to the bereaved. We should also study which current technologies are most used and which future technologies are most desired by the public, and why.

Few experiences are as universal as mortality and facing the loss and difficult adjustment when those close to us die. Technology has long provided the means for us to evoke the presence of the deceased, but that evocation is becoming increasingly vivid and realistic. While this may cause substantial harm if the technology is used thoughtlessly or unethically, it also raises the very hopeful possibility of easing the grieving process. Presence, and presence research, is likely a key to accomplishing that important goal.

References

References are available from the first author.

---

charcoal drawing and an oil painting in evoking the sense of presence of the first author's deceased dog (see http://matthewlombard.com/Sidra).

When Mixing Physical Presence with Telepresence: Analysis of a Pilot Study

Cara Stitzlein, Leila Alem
ICT Centre; CSIRO, Australia
{Cara.Stitzlein@csiro.au, Leila.Alem@csiro.au}

Abstract

Current videoconference systems combine face-to-face (F2F) and mediated interactions. We extend the use of a Social Presence measure to a real-world setting combining co-located and remotely located people completing a collaboration activity. Comparisons between physically present and remotely located others did not indicate significant differences in social presence regarding media condition. Post-trial interviews reveal aspects of group member experiences.

1. Introduction

Video conferencing facilitates access to remote places and people through technological mediation. Technology that enables remote and co-located engagement is known as groupware [1]. Numerous approaches attempt to approximate the ideal face-to-face (F2F) quality. For example, Hauber et al. developed an interface with 3-D like interactivity [2]. In another approach to support a realistic telepresence encounter, Yamaashi et al. provide the remote person with two views from the connected site, a foveal (near) and peripheral (far) view [3]. These are just two examples of systems encouraging a natural sampling of visual information of a remote space.

Telepresent encounters have a social component as well as a technological one. Constructs like social presence allow researchers to evaluate the connectedness and interpersonal fluidity of the mediated interaction. A theoretically grounded social presence measure used in telepresence research assesses the extent to which a person feels connected with a remote person (for example, through an interface). The presence field has developed several measures to gauge social presence [4]. Typically compared with F2F, it's commonly theorized that co-located encounters reflect highest ratings of social presence. One standardized measure, The Networked Mind measure of social presence has shown this difference in attributed social presence based on media condition[2;5]. We examine the experiences of a mixed presence group, in which F2F and mediated encounters occur simultaneously, investigating how these findings may be extended.

2. Description of Study

We compared two media conditions, co-located and remotely located. It was assumed that co-located participants would report greater ratings of social presence than remotely located participants. We expected standardized half-structured interviews to reflect qualitative data not captured by the questionnaire but relevant to impressions of "connectedness" with remotely located others.

Two participants sat side by side and worked with a single participant in another room. A videoconferencing connection between the two locations used a high quality audio and video link utilising Digital Video (DV) over IP [6]. The remote participant was provided with two views (Figure 1): a zoomed out, wide angle view and a tight, close up view.

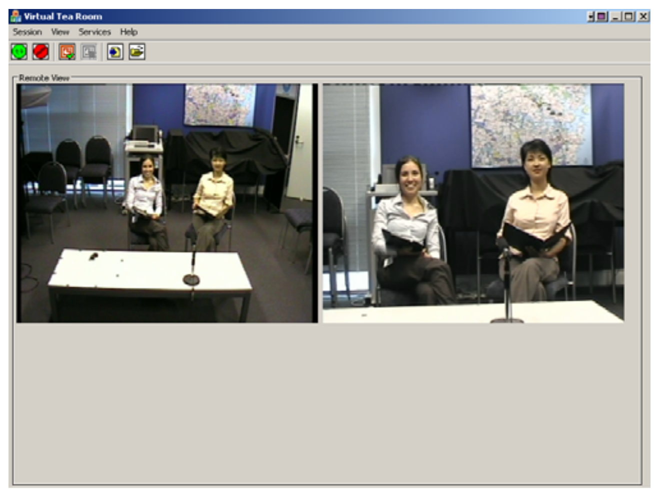


Figure 1: Remote Interface

Groups of three completed the scenario "Desert Survival Game"; used in previous studies of social presence [2;5;7]. Verbal tasks involving negotiation benefit from video support and are therefore design appropriate [8]. Two activities characterize the activity: individual rank-ordering of items critical to survival (e.g. compass, sunglasses, overcoat), and a group activity negotiating about items' importance to group survival. The Networked Mind tool was administered to participants in both conditions post-task [9]. Co-located participants rated both the physical present and telepresent participant. The remote participant rated the co-located pair together.

Post-trial interviews and discussions captured recall of task activity and impressions of the mediated encounter. Some remotely located participants regarded the interaction "like watching TV," which could infer a lower sense of social presence than co-location. The remote participant's visual behaviour is captured with gaze tracking technology (See Stitzlein et al. (submitted) for preliminary results [10]).

3. Results of Study

This pilot consisted of 24 participants (10 females and 14 males), between the age of 20 and 44 (mean = 26 years, S.D. = 6.69). Cronbach alpha for the six social presence subscales ranged from 0.83 to 0.94, satisfying reliability requirements [11]. Three groups of questionnaire responses were statistically analysed in one-way ANOVAs, revealing no significant differences between conditions with respect to social presence ratings (See Table 1). In interview responses, participants reported satisfactory physical descriptions of others and recalled task artefacts like first item of consensus and an item of debate. Their impressions of groupware technology and activity context indicate the most salient aspects of the collaboration.

	Co-located Pair		Remote Person
Social Presence Factor	**F2F**	**Remotely located**	**Remotely located**
Co-presence	6.17 (.71)	5.71 (.92)	6.04 (.56)
Attentional Allocation	4.90 (1.10)	4.82 (1.25)	5.29 (.90)
Perceived Message Understanding	5.90 (.85)	5.56 (1.06)	5.56 (.62)
Perceived Affective Understanding	4.74 (.90)	4.22 (1.46)	4.83 (.93)
Perceived Affective Interdependence	4.33 (1.03)	4.18 (1.04)	4.13 (1.40)
Perceived Behavioural Interdependence	4.24 (1.07)	3.89 (1.23)	4.60 (.69)

Table 1: Mean Scores by Media Condition & Social Presence Factors
Reported: mean ratings (standard deviation)

4. Discussion

We administered data capture techniques in a mixed presence setting where individuals engaged in a collaborative scenario. The main objectives were to measure the degree of social presence in this configuration, subjecting the questionnaire to a "reality test" of telepresence [12]. Administering a questionnaire in a mixed presence group contributes to the validity criteria of this particular social presence tool [5]. Results indicate impressions of social presence for someone physically present compared to someone telepresent are statistically indistinguishable in this setting. Additional data forms suggest that use of a multi-method approach more completely captures social presence.

Of course, there are some experimental limitations: small sample size and possible confound of sequential completion of questionnaires on a single experimental event. Null differences could also imply a lack of variance in presence levels between the simultaneous mediated and F2F encounters.

Future work will reflect interest in a multi-measurement approach, blending conventional questionnaires with behavioural observations and quantitative data. Such analyses in real-world videoconference settings benefit the research community as well as designers of telepresence technology.

References

[1] Ellis, C. A., Gibbs, S. J., and Rein, G. L. (1991). Groupware: some issues and experiences. *Communications of the ACM 34(1)*, 39-58.

[2] Hauber, J., Regenbrecht, H., Hills, A., Cockburn, A., and Billinghurst, M. Social presence in two- and three-dimensional videoconferencing. In *Proceedings of ISPR 2005,* 2005.

[3] Yamaashi, K, Cooperstock, J. R, Narine, T, and Buxton, W. Beating the limitations of camera-monitor mediated telepresence with extra eyes. In *Proceedings of CHI '96*, 50-57. 1996.

[4] Presence-research.org URL: http://www.presence-research.org/

[5] Biocca, F., Harms, C., and Gregg, J. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *Proceedings of ISPR 200*1, 2001.

[6] Ogawa, A., Kobayashi, K., Sugiura, K., Nakamura, O., and Murai, J. Design and implementation of DV based video over RTP. In *IEEE Packet Video Workshop, 2000.*

[7] RogerKnapp URL: http://www.rogerknapp.com/download/games.htm

[8] Veinott, E. S., Olson, J., Olson, G. M., and Fu, X. Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In *Proceedings of CHI '99,* 302-309. 1999.

[9] Harms, C. M., Levine, T., and Biocca, F. The effects of media type and personal relationship on perceptions of social presence. Thesis/Dissertation. Michigan State University, East Lansing, 2004.

[10] Stitzlein, C. A, Li, J., and Alem, L. A study of visual behaviour in video mediated negotiation using gaze tracking. Submitted poster for CSCW 2006.

[11] Cohen, J. (1992). A Power Primer. Psychological Bulletin 12(1), 155-159.

[12] Usoh, M., Catena, E., Arman, S., and Slater, M. (2000). Using presence questionnaires in reality. Presence 9(5), 497-503.

Presence Considerations In Music Production

Jack Klotz, Matthew Lombard
Temple University
{jklotz@temple.edu, lombard@temple.edu}

Abstract

This demonstration by a music industry professional illustrates a series of considerations in the producing and mixing of commercial music related to presence. The goal is to stimulate further research and to work toward the creation of recommendations for music industry professionals so that they can increase audience members' sense of connection with music and performers.

1. Introduction

Music recording and mix engineers, by and large, do not think about presence in their work, at least not in the sense of the term that scholars use in relation to mediated communication.

When it comes to aesthetic considerations in music production, engineers and producers may think along the lines of establishing the listeners' perspective relative to the music; creating or recreating a space in which the music seems to be performed; capturing a rough live, or creating a refined layered, performance; and using musicians or recreating instrument sounds with technology. Each of these will impact the sense of presence, defined here as a sense of connection with music and performers that includes perceptions of physical space, social interaction, and real-ness or genuine-ness – that is felt by the listener. There are several techniques that are employed by sound engineers during the recording process (the initial "gathering" of the various elements), and more importantly during the mixing process (the final "assembly" of those various elements) to accomplish these goals.

2. Establishing Perspective

In the mix engineer's lexicon, establishing the listeners' perspective is largely a matter of determining where to place the listener relative to the "performance" to which they're listening. The recording engineer will establish the listeners' perspective by their microphone placement and, in instances where multiple microphones are used, the mix engineer determines the listeners' virtual position by manipulating the various elements utilizing three primary tools: panning, reverb and equalization.

The mix engineer uses *panning* controls to position the various sonic elements in the sound field. In the case of stereo, the pan control works like the balance control on a stereo, shifting volume from left to right in a continuous panorama. In panning, there are two primary considerations: listener perspective and width of the sound "image". In a typical pop music mix, the mix engineer will position the individual pieces of the drum kit (bass drum, snare, tom-toms, etc.) in a specific location in the panorama based on whether the engineer wants to position the listener in the virtual audience, looking at the drummer, or behind the drum kit, in the drummer's seat – placing the listener, if only subconsciously, inside the performance. Whatever the perspective of the listener, the image can be very wide and dramatic or more narrow and natural – a roll across the tom-toms, for example, could seem to fly from one side of the room to the other, or be made to sound as if the drums are set up in a single location in a large room.

A good deal of recording is done in acoustically controlled studios. Some are designed to have a desirable ring or reverberant tone (referred to as *live* rooms), but many are designed to avoid that reverb and are acoustically *dead*. The engineer chooses the style of reverb on a continuum that ranges from natural-sounding, realistic emulation to polished, dramatic, artificial sounds. Reverb defines not only the size of the virtual room (more reverb suggests a larger space), but the performers' locations within that space (in conjunction with panning, above) and the listeners' apparent relative distance from the performers (in conjunction with relative volume in the mix and use of equalization, below).

Equalization (EQ) allows recording and mix engineers to alter the tonal quality of sounds by boosting or cutting the volume of a given frequency. Engineers tend to think in terms of increasing "warmth," adding "punch," making a sound "fatter," and other descriptive terms. EQ can also be one part of creating a sense of closeness or distance. The further away from a sound one is, the less bright (dominated by high frequencies) it will sound; and the closer one is, the more bright it will sound. EQ can thereby be used to manipulate brightness, increasing it to create a feeling of closeness, reducing it to create a feeling of distance.

3. Intimacy and Perception of Content

A comparison of two recordings of the pop song "…Baby One More Time" illustrates the considerations above and their effects on listeners, along with other manipulations discussed below. The 1999 hit version by Britney Spears is a studio production full of electric instruments, synthesizers, drum machines and a wealth of overdubs and studio "tricks"; and the mix features wide panning and lots of big, lush reverbs. A later recording by the British alternative rock group Travis is performed live, in front of an audience, with simple, non-electronic instrumentation; no overdubbing or added reverb; and no panning, and sounds rough, unpolished, and natural. The Travis version is often described as more "real" or "authentic" than Spears' more "manufactured" or "artificial" version. In presence terms, the Spears version seems less intimate, more distant, less connected.

3.1. Studio vs. Location

In presence terms, the difference between a studio and location recording is not just the (potential) difference in the apparent size of the space and locations of the listener and performer(s), but the degree to which the performance seems "alive," "real," and "heartfelt," all of which impact the degree of connection the listener feels with the music and performer.

A studio recording of "Rock n' Roll All Night" by the 1970s rock group Kiss had been unsuccessfully marketed as a single, but one year afterward, the "live" version proved to be their breakthrough hit. The performers' energy is significantly improved compared to the studio version, making the overall sound more compelling, and creating a sense of being present at the concert and a heightened feeling of social presence (parasocial interaction [1]) for the listener.

3.2. Tracking & Overdubbing vs. Recording "Live"

In music production parlance, recording performers "live" does not necessarily refer to capturing a concert performance. It also refers to studio recordings where musicians perform together. The other approach to studio recording is called overdubbing and involves performing each instrument separately and layering them together.

According to legendary record producer Sir George Martin, "In order to make a good record, you have to have fun doing it – that feeling is passed directly to the audience. It can't be faked or inserted after the fact with any piece of technology. It's either there or it's not" [2]. Excessive overdubbing is likely to reduce the sense of fun among the musicians and thereby reduce presence for listeners.

Fun is evident in the Beatles' recording of "Twist And Shout" from early 1963, which was recorded in one take with all members of the band playing and singing together in the studio – a performance given at the end of a 12-hour day of recording with the lead vocalist, John Lennon, suffering from a bad head cold. Conversely, Steely Dan's 1976 hit "Peg" reveals a "perfect" performance, instrumentally and vocally. While still a very good record, it lacks the energy, the "fun" and the level of connection of "Twist And Shout." The production is "slick" and the recording is "clean," but this lessens the sense of genuineness.

3.3. MIDI vs. Musicians

Removing human musicians from the process reduces even further the possibility for energy in the performance. MIDI makes it not only possible, but relatively easy for a composer to perform all of the parts of a composition him or herself, exactly as they hear them in their mind's ear. The trade-off between control and energy can be heard in two Stevie Wonder songs, 1976's "I Wish" which was done "live" in the studio by musicians playing instruments, capturing the social interaction of the musicians, and "That Girl" from 1982 performed by Wonder alone, using MIDI.

Conclusions

Professionals in the recording industry use a variety of sophisticated tools and techniques that create for listeners a sense of 'being there' in real or created recording spaces in which they react and even seem to interact with genuine or 'real' music and performers. These recordings strengthen the sense of connection listeners feel with performers as well as their music. In commercial terms, when fans can connect with these artists by listening to their recordings they should be more loyal customers, which can lead to greater commercial success and longevity for such performers.

There are intriguing opportunities for research in this area. One possibility would be to test market various versions of the same song, each using different combinations of the approaches described above, with listeners in various demographic groups, and measure different aspects of presence, emotional reactions to the music and performers, and purchase intentions.

If presence scholars were to develop guidelines for professionals in the recording industry, the suggestion could be to record on location rather than in studio, have the performers all perform together rather than using overdubbing, and reduce dependence on computer-controlled synthesis whenever possible and instead make use of performing musicians. When mixing, engineers should lean away from the "larger than life" approach of exaggerated panning, bright EQ and large, bright reverbs, and more toward the "intimate" approach of narrower panning, warm EQ and smaller, warmer reverbs. In short, by either capturing music in or more accurately replicating real spaces and by embracing the imperfections that make a musical performance human, one increases the possibility for one's recording to make human-to-human contact.

References

[1] Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: *Observations on intimacy at a distance. Psychiatry, 19*(3), 215-229.

[2] Martin, George. (2000). The Making of Sergeant Pepper. Lecture given at the Sanders Theater, Harvard University on October 12. Cambridge, Massachusetts, USA.

Building Tele-Presence Framework for Performing Robotic Surgical Procedures

Peter B. Panfilov, Frank M. Cardullo, and Harold W. Lewis, III
Moscow State Institute of Electronics and Mathematics, SUNY at Binghamton, SUNY at Binghamton
{panfilov@miem.edu.ru, cardullo@binghamton.edu, hlewis@binghamton.edu}

Abstract

A research project is being presented aimed at the development of a system of methods that in the future will make it both safe and practical to perform telerobotic surgery. To mitigate the effects of the physically unavoidable delay we propose a telepresence system framework, which is to have the surgeon operate through a simulator running in real-time enhanced with an intelligent controller to provide the safety and efficiency of an operation. Three major research areas must be explored in order to ensure achieving the objectives of our project. They are: simulator as a predictor, image processing, and intelligent control. Each is equally necessary for success of the project. These are diverse, interdisciplinary areas of investigation, thereby requiring a highly coordinated effort by all the members of development team, to ensure an integrated system.

1. Introduction

Robotics is one of key technologies that have a strong potential to change how we live in the 21st century. We have already seen robots exploring surfaces of distant planets and the depths of the ocean, streamlining and speeding up the assembly lines in manufacturing industry. Robotic vacuum cleaners, lawn movers and even pets found their ways to our houses. Among the medical applications of robotics the minimally invasive surgery was the first to demonstrate a real advantages and benefits of introducing robotic devices into operating room over conventional surgical methods. So far, these machines have been used to position an endoscope, perform gallbladder surgery and correct gastroesophogeal reflux and heartburn.

It is interesting to note that in the late 1980s, after its inception the utilization of laparoscopic cholecystectomy grew rapidly. However, minimally invasive surgery (MIS) for other operations has not experienced the same pattern of growth. According to Ballantyne [1], the reason is that in general laparoscopic procedures are hard to learn, perform and master. This is a consequence of the fact that the camera platform is unstable, the instruments have a restrictive number of degrees of freedom and the imagery presented to the surgeon does not offer sufficient depth information. The solution seems to be at hand with the significant growth of robotic surgery. This is surgery where-in the surgeon operates through a robot. In a sense this robot is a telemanipulator under the control of the surgeon. The robotic system provides a stable video platform, added dexterity and in some cases a stereoscopic view of the surgical field.

The surgical robots use technology that allows the human surgeon to get closer to the surgical site than human vision will allow, and work at a smaller scale than conventional surgery permits. The robotic surgery system (e.g., the daVinci robot) consists of two primary components:

- A viewing and control console (operating station)
- A surgical (robotic) arm unit

Sitting at the control console, a few feet from the operating table, the surgeon looks into a viewfinder to examine the 3D images being sent by the camera inside the patient. The images show the surgical site and the two surgical instruments mounted on the tips of the rods. Joystick-like controls, located just underneath the screen, are used by the surgeon to manipulate the surgical instruments. Each time one of the joysticks is moved, a computer sends an electronic signal to one of the instruments, which moves in sync with the movements of the surgeon's hands.

Since proximal robotic surgery seems to be maturing the next logical step in surgical care is to extend to remote applications of robotic surgery. That is to say, the surgeon and the operating console are at one location and the robot and patient at another. The idea of remote robotic surgery, or as some refer to it, telesurgery, has been an objective for some time, especially in the military. This advancement is seen by the military as the means by which the next major improvement in battlefield survivability [2]. In addition to the military application, the technology could be useful if an astronaut were to require emergency surgery while on the space station. Furthermore, perhaps the most ubiquitous application will be in civilian medicine. Patients in medically remote areas would have the option of receiving an operation performed by a renowned surgeon even though the surgeon and patient may be thousands of miles apart.

The long term objective of this research is to develop a framework for remote robotic surgery application which will permit surgery between any two places on earth with a patient in one location and the surgeon in another. The major impediment to remote surgery is the effect of telecommunication delay on the surgeon's performance. It has been shown in a myriad of studies of human-in-the-loop systems that system delays lead to degraded operator performance and ultimately unstable systems. Since the delay cannot be eliminated, in order to accomplish this objective, the only solution is to mitigate the effects of the delay on the surgeon performing the operation. Relying on many years of experience in human in the loop simulators, including a simulator to train surgeons in laparoscopic cholecystectomy, we propose a method whereby the surgeon operates through a simulator in a virtual environment free from the impediments of telecommunication delay. Since, in robotic surgery, the surgeon is already in a synthetic environment the introduction of a simulator does not significantly alter the physician's perceptual stimuli. The operating station containing the control inceptors and the visual displays is the same as that used to control the surgical robot in the conventional configuration. For robotic surgery systems, one can often use computing power to create novel interfaces such as virtual environment interfaces that enable new paradigms of instrument operation and data visualization.

These systems are inherently distributed systems, often with the robot, an interface generation computer, and special purpose interface devices such as haptic force-feedback devices and head-mounted displays distributed on a local area network in a single room or building. This simple distribution naturally motivates a larger scale distribution of system components across the wide area and/or global networks such as the Internet to enable telepresence experience and teleoperation of the special purpose medical equipment. The result is a distributed virtual operating room consisting of a geographically separated collection of medical instruments, computers, and physicians all interconnected via a computer network.

Most researchers in robotics area now accept the definition of telepresence put forward by Thomas Sheridan [3] as the sense of actually being at a remote or synthetic workplace which users of telerobot or virtual environment systems developed during operation of the system's human interface. Since users of a virtual environment interface are in the same position with respect to simulated effectors in the virtual environment as that of human telerobotics controllers with respect to a remote robot, the two control situations also may be generalized under the term virtual environment (VE). This generalization is reasonable since with respect to the various users' physical viewpoints in either case user experiences *presence in an environment by means of a communication medium*.

To qualify the operators' sense of remote presence during teleoperation or use of virtual environment interfaces as an explanatory scientific concept, Prof. Sheridan identified three determinants of presence [3]: 1) the extent and fidelity of the sensory information that may be displayed to the users, 2) the extent and fidelity of the users' control over the sensory information and 3) the extent and fidelity of the users control of effectors in the environment. These three components suggest measurement by well-established continuous variables associated with physical sensors and effectors, e.g. bandwidth, latency, dynamic range etc.

As emphasized by Stephen Ellis from NASA Ames in [4], presence or equivalently simulation realism is particularly useful as it identifies reasonably independent, easily measurable characteristics of communication channels. In terms of Sheridan's analysis of presence, for example, the fidelity of sensory display could be instantiated as a bandwidth, x_1, or update rate of a visual display. Similarly, the bandwidth or update rate of the control of a virtual effector could provide another potentially continuous variable, x_2, influencing a measure of presence. These signal characteristics are especially useful since they can be considered independently of the users' ability to modify them through control of sensor positions. One may additionally easily imagine that improvements in any of them would lead to improved operator performance, improved realism, and an increased sense of presence within the virtual environment. In fact, with respect to manual control the dynamic requirements to communication channel are classically known to determine the fidelity of control [4].

The proposed surgical telerobotics framework challenges the problem of communication channels' characteristic improvement through a complex combination of systems science, optimal control, and intelligent systems approaches. In our embodiment the simulator acts as a predictor, illustrating for the surgeon what would be the case if there were no delay, and – somewhat simplistically – the image preprocessor at the robot side can be thought of as a corrector. The intelligent controller is designed as an optimizer. Fig. 1 is a simplified view of the proposed framework and the general research areas. The following paragraphs explain the architecture.

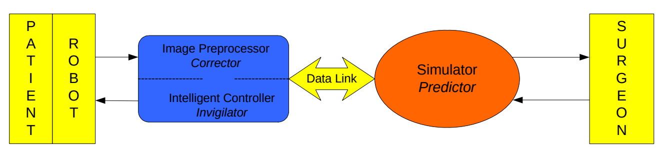


Figure 8 Surgical telerobotics as a tele-presence system

Lombard and his colleagues define presence as "the perceptional illusion of nonmediation" [5], which is the consequence of transparent mediating systems. By "illusion of nonmediation," they refer to a phenomenon in which "… a person fails to perceive or acknowledge the existence of a medium in his or her communication environment and responds as he or she would if the medium were not there" [5]. The three components of the framework working in concert should ultimately constitute a communication medium which is transparent enough to bring the human operator as close as possible to the telepresence "**ideal** of sensing sufficient information about the teleoperator and task environment, and communicating this to the human operator in a sufficiently natural way, that the operator feels physically present at the remote site."[6]

2. The real-time simulator as predictor

Modern simulators tend to be very complex systems in their own right, but one particular aspect of how the simulator will function in this case is where we place the emphasis in calling our simulator a predictor. That is, in order for our time-delay-mitigation scheme to work, the simulator must predict what will happen in the surgical field before it happens. Furthermore, the predictive mechanisms in this case are based on dynamic modeling of a far from trivial sort. Adding more complexity to the task, the system must be designed to allow the models to be updated in real time as the delayed information from the surgical field becomes available. Clearly, dynamics models both for the robot dynamics and organ dynamics are necessary for the simulator to function in this way. Though both are challenging, the organ dynamics modeling is known in medical research circles to be extraordinarily difficult, particularly in the case of soft tissue. For this, we intend to experiment with a variety of approaches, which will include both finite element analysis and continuum analytical models.

The simulator as used here is clearly a perfect example of an anticipatory system. Note that anticipatory behavior is often viewed in the literature [8, 9] as a primary characteristic in intelligent systems.

2.1. The time delay problem in tele-robotic surgery

It is well known that system delays will cause a deterioration of the human-machine system performance. As a matter of fact this is true for any control system, not only a human-in-the-loop control system. Fig. 2 illustrates the time domain effect of delays of 0, 200, 400, and 800 ms where the input is a unit step [7].

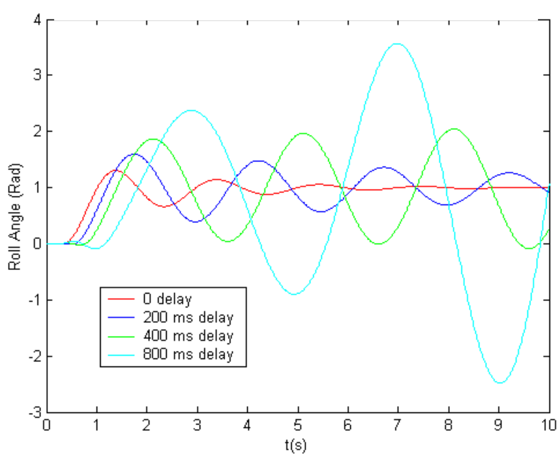


Figure 2 Step responses of a closed loop system with different delays

The graph indicates that as the delay increases, the response lags the input by a greater amount. In addition the 400 ms delay case seems to display limited stability, while the 800 ms delay case clearly exhibits unstable response. The system analyzed includes a fourth order plant and a human operator model, to which the delays are added.

Fig. 3 presents the results of frequency domain analysis of the same system. Here, one observes that the 400 ms delay case yields a phase margin of approximately zero, while the 800 ms case has a negative phase margin. We can then examine human operator performance data in a system with and without delays. There are many such examples in the literature. It is observed that when delays become long, human operators will adopt a move and wait strategy. This allows the operator to observe the results of his/her action before committing to another action. The move and wait strategy may be acceptable for controlling a lunar or Mars rover but it is unacceptable in tightly closed loop applications, robotic surgery being one of them.

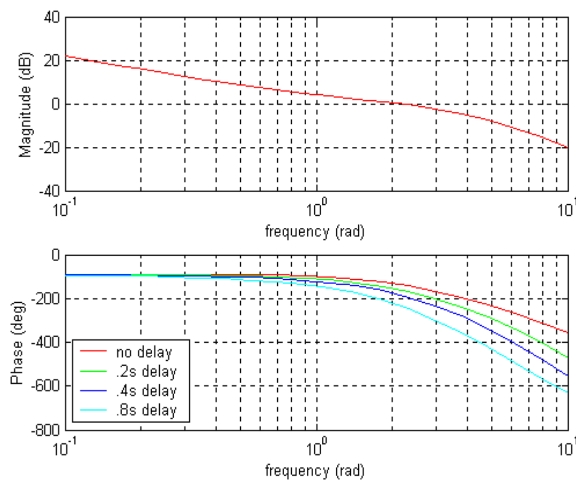


Figure 3 Frequency responses (Bode diagrams) of a closed loop system with different delays

When the delays in an aircraft flight control system become too long the control loop becomes unstable and the aircraft is said to display pilot induced oscillation. This is another case where the move and wait strategy will not work.

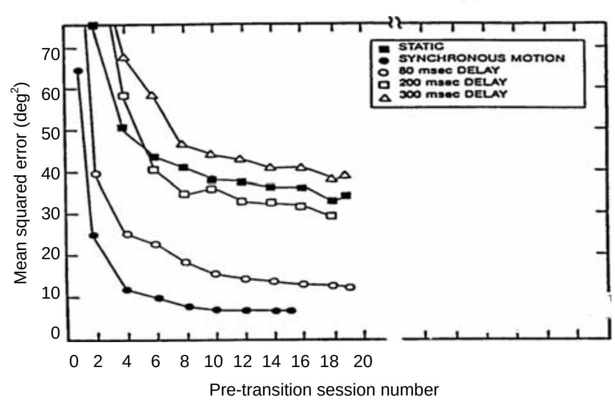


Figure 4 Effect of delay on a system operator

performing a tracking task with/without force feedback

Fig. 4 illustrates the effect of delay on a system operator performing a tracking task with and without force feedback. There were several cases of delay (0, 80, 200 and 300 ms) in the force feedback. In all cases the subjects had a narrow field of view visual presentation. The graph shows that at 200 and 300 ms delay in the force feedback the operator's performance is essentially as bad or worse as with no force feedback [10]. Whereas, at a delay of 80 ms his/her performance is much improved and almost as good as a fully synchronous feedback.

Preliminary studies were conducted using experienced laparoscopists in a suture knot tying task. The task was performed using a laparoscopic training device with delays introduced, in 25 ms intervals, into the video monitor via an analog delay device from Prime Image. The performance metric used in this study was the time it took the subjects to complete the knot. Fig. 5 illustrates the results. For delays up to about 100 ms the execution time remained relatively constant at about 13 seconds. Above the 100 ms point the time increases substantially. Preliminary results seem to indicate that by the time the transport delay approaches 500 ms the time to complete the knot is about 90 seconds. One of the interesting results is that subjects began to experience nausea at delays approaching one second. This result was quite unexpected based on our considerable experience with simulator sickness.

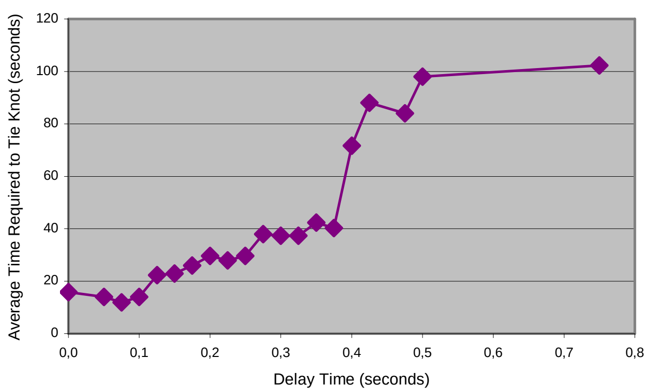


Figure 5 Knot tie time as a function of delay time

2.2. Organ dynamics modeling

Surgery simulation involves biological tissues, it is therefore essential to determine the deformation of the tissue while in contact with a surgical instrument. This is important to transfer a realistic picture of the tissue to the surgeon in real time. The problem is how to determine the interaction between the tool and the tissue. This requires a rigorous model of the material properties, an accurate simulation method to reflect the actual behavior of the tissue, and fast simulation results to enable real-time interactive simulation.

The organ dynamics model must be accurate, timely, and responsive. For the success of this telerobotic system, realistic visual deformation and haptic output of the simulator is essential. The surgeon must see and – it is desirable to feel the organs presented by the simulator as if they were the actual organs of the patient. As an aside it should be noted that the daVinci robot does not provide significant, haptic feel of the tissue. However, as part of this project we will add this capability, since haptic feedback has been demonstrated to mitigate some of the effects of delay by virtue of the fact that the proprioceptive system is faster responding than vision. These simulator outputs depend on the organ dynamics models. The organ dynamics models are based on solid models of the components of the biomechanical system upon which the surgery is being performed, - in this case soft tissues.

The first requirement of the organ dynamics model of the biological system is accuracy: the deformation and force feedback of the simulated organ on which the surgeon is operating must be as close as possible to the actual patient organ. Secondly, the computation that is required to determine the simulated deformation and forces must be performed in a time interval that appears to the surgeon to be real-time. Also, the organ dynamics model must be capable of being quickly updated using the decompressed video feedback that provides corrections without creating visual or haptic discontinuities, "jumps", that could disturb the surgeon. In other words, the organ dynamics model must be responsive to the correcting feedback.

The organ dynamics model employed by the simulator will respond with force and displacement feedback to the surgeon to mimic the actual deformation of the structure, e.g., the organ, at the remote location. The simulation of the deformation allows the surgeon to pursue the surgery on the deformed geometry. This overcomes the delay associated with the transmission from the actual remote operating site. The response of the structure to the actual robotic surgery at the patient end will be used to correct the model, but the corrections must be kept small. A computational model of the dynamic response of the structure to the surgery is needed. The final trade-offs in the design of the model will be between these three constraints: accuracy, timeliness, and responsiveness.

2.2.1. Elements of organ dynamics model. There are three elements of the organ dynamics model: (1) the structure (geometry) of the organ, (2) constitutive material properties of the soft tissues, and (3) the applied forces.

Although a generic base-model of human organs can be a starting point for use in the simulator, the geometry of the organ that is used by the simulator during surgery cannot be a generic model. It must be patient-specific. The geometry of the patient-specific organ is needed to develop an accurate organ model.

There are several techniques that can be used to determine the geometry of an individual organ. Among these are x-ray, magnetic resonance (MR), computed tomography (CT). Chen, et al. [11], have used MR images to model knees. These were analyzed and mapped to a known geometric (cylindrical) kinematic model. Subsequently, a 3D virtual model of the knee was constructed [12]. Gibson, et al. [13], have also taken this approach. In Gibson's research, construction from MRI involves an analyst's visual examination of the segmented images to distinguish tissue intensities. The creation of the geometric model of the organ will require imaging of the patient prior to surgery and could require manual or programmed image analysis.

In addition to the geometry of the structure, constitutive material properties are needed for each type of tissue present: ligament, cartilage, other soft tissues and adjacent bone. To simulate the results from the surgical procedure, the following are needed: (1) the amount of deformation, (2) the force exerted on the surgical instrument, and (3) the resilience of the material. The mechanical properties of biological material are inherently non-linear, inhomogeneous, and anisotropic, for example, soft tissues, such as ligaments, have viscoelastic properties. Bulk, ex vivo material properties are available from the general literature [14], but these can vary significantly from in vivo properties.

Indentation techniques have been used to determine in vivo material properties. Ottensmeyer [15] describes an instrument (TeMPeST 1-D) that was used to determine the viscoelastic properties of porcine liver. The device generates small amplitude oscillations and the force response is recorded. An elastic modulus and viscoelastic stress-strain equation were determined. The indentation technique has been also used by Tay, et al. [16] to determine the force response of pig liver and esophagus tissue. A force vs. displacement plot for in vivo pig esophagus tissue is presented. Material property data can be determined using this force-displacement curve. The apparatus that was used consisted of an indenter with force sensor that was fixed to a programmable force feedback haptic device (Phantom Premium-T 1.0). In addition, the indentation method is employed by Samosky, Grimson and Gray [17] to determine cartilage stiffness.

Vuskovi, et al. [18] used a tissue aspiration method that is a non-contact technique. "The measurement method consists in leaning a tube against the target tissue and gradually reducing the pressure in the tube." The profile of the tissue in the tube is measured. The instrument consists of a tube, mirror and lens with light provided by an optical fiber, and pressure sensor. The acquired data is then fit to a finite element model to determine the material constants. A technique for directly measuring in vivo material properties has been developed by Brouwer, et al. [19]. A device grasps a portion of tissue and stretches and compresses the tissue recording force and displacement. They have successfully applied the technique to determine porcine tissue properties, but this technique does not appear to be transferable to use on humans.

For the simulation, a data acquisition system is needed to determine the forces that are applied to the structure of interest. Forces and deformations have been measured using surgical instruments that are fitted with six-axis force transducers [19]. These transducers have been used to examine tissue by mounting them to a probe. As the probe is moved to selected points on the surface of the tissue (e.g., cartilage), probe tip position and force are recorded [17].

2.2.2. Objectives and scope of organ dynamics modeling research in the current project. Four major goals for the tissue modeling in the project define the scope of the research.

First, the creation of the geometric dynamic model of the organ will require determining material properties for the various types of tissue involved and the loads applied to the tissue due to the surgical instruments. This includes 1) investigation and analysis of the indentation and aspiration techniques, 2) development of a method to determine the required material properties, and 3) conducting experiments to measure forces applied during surgery.

Second, the major objective of the organ model is to determine the change in topology of the organ when it is impacted by a surgical instrument. The topology of the organ is defined by the finite element (FE) model meshes. The location of each node in space is calculated and made available to the image generator. The image generator will then produce the graphical primitives based on the location of the nodes. It is common for collision detection to be carried out in the image generator and it is anticipated that this will be the case here since the image generator is rendering not only the tissue but the surgical implement from the perspective of the video camera.

Third, to determine in detail how to integrate the organ dynamics model most effectively into the overall system, it is necessary to determine and to demonstrate methods for updating the organ dynamics model. As shown on the system block diagram in the figure 6 there are two inputs to the organ dynamics model: surgeon input and image feedback. The direct surgeon input is the motion of the inceptors. This is translated into the motion of the robot effectors through the robot dynamic model resulting in the actions that are applied to the organ dynamics model. The second input, the image feedback, is the decompressed and preprocessed image from the patient end. This image is used to update the state of the organ dynamics model so that it does not deviate from the actual organ.

Fourth, to optimize computational efficiency, several techniques will be employed. A reduced order FE models will be developed. This factor alone will save substantial computation time. We will also create zones around the surgical probe, once it is in contact with tissue. The zone closest to the probe will require the finest mesh for the finite element model and the highest computation rate. The more remote zones will have coarser meshes and slower computation rate because they will be less impacted by probe contact. We anticipate at least three zones. It may also be possible to use lower order models in these outer zones.

2.3. Real-time realistic visual simulation

Visual simulation has become a commonly employed method to provide the environment in which to train surgeons to perform common, yet risky, medical procedures. Typical examples are laparoscopic cholecystectomy and gynecological sterilization. Today the virtual environment comprises a 3D photo-textured polygonal database representing the anatomy of the critical internal organs involved in the procedure.

2.3.1. Dynamic database alteration. The majority of any given surgery is often spent cutting, cauterizing, and removing the fatty and connective tissue that obscures and interferes with the areas of interest. Often the areas of interest can be difficult to differentiate from the surrounding tissue, especially when considering the challenge of spatial cognition. For example, in a cholecystectomy (removal of the gall bladder), the main challenge is in distinguishing the cystic duct from the common bile duct and the hepatic artery. Since the purpose of this research is to ultimately perform live surgery, our objective for free-form volumetric dynamics is to realistically simulate the physical reaction of tissue due to actions such as cutting and tearing as would always be found in a live situation. This includes the simulation of different kinds of tissue ranging from organ tissue to fatty and connective tissue, as well as striations in the tissue. The internal representation used to model tissue must allow for free-form alterations including cutting and tearing, and will provide realistic force-feedback. To simulate tissue with this level of fidelity requires a new approach that goes beyond the limitations of current approaches.

Our initial research is based on the use of free-form iso-spheres or a high-order surface (HOS), with the use of a space-filling Union of Spheres (UoS) as the underlying physical model for tissue. With UoS, objects are modeled using collections of packed spheres. Each sphere can be seen as an independent object with forces acting on them computed by the organ dynamics model.

A major advantage to using UoS over traditional approaches is that the generation of UoS models from medical imagery sources such as MRI or CAT scan data is straightforward. This allows for the possibility of "mission rehearsal" on actual patient data prior to the surgery, and also reduces the challenges and costs of generating realistic anatomical models, and is a perfect fit to include patient-specific data such as MRI or CAT scan information for the actual surgery. Another advantage of a UoS approach is that collision detection becomes a trivial problem. The intersection of an object (such as a surgical instrument) with a sphere is a simple calculation because inside/outside information is readily available. Also, because the spheres have a known spatial location and extent, only a small subset of spheres need be compared to any given object.

The dynamics among neighboring spheres will be governed by the state vector of each sphere calculated by the organ dynamics model.

Associated with each sphere will be numerous properties including a flag indicating if a sphere is an interior or surface sphere. Initially, the minimum number of spheres needed to represent the anatomy will be used. This will reduce both the memory and computational burdens of the force model. At runtime, spheres will have the ability to subdivide and merge, depending on the complexity required to represent a given modification. This will allow the force model to be scalable, capable of automatically taking advantage of future hardware and software improvements. The maximum level of subdivision will be limited by the available computational resources, ensuring the highest possible fidelity at interactive rates.

2.3.2. Spatial database structure. Many of the algorithms that will be used in our approach require data to be efficiently accessible based on spatial location. The algorithms that benefit from exploiting the spatial organization of data include nearest neighbor searching for the propagation of force through the union of spheres and collision detection between spheres and other objects. The visualization system will also require spatial efficiency in order to perform visibility determination to reduce the rendering burden as well as texture manipulation and synthesis for additional realism.

The challenge that arises when using an UoS approach is how to derive the structure of the control points and the texture coordinate values for the generated surface points. One possibility for deriving control point values is that the number and relationships of control points is fixed, but their values can vary. This is similar to how traditional SMD systems represent the underlying physical model of an object. Our goal, however, is to visualize free-form deformation of an isosurface physical model. This requires a more flexible approach to rendering dynamic high order surfaces.

2.3.3. Texture synthesis and mapping. The mapping of textures onto high order surfaces has been a long-standing challenge to computer graphics. Because the surface is procedurally generated, texture coordinates are not easily defined for areas between control points. To address this challenge, a second high order surface can be generated specifically for the computation of texture coordinates. Traditionally, this method incurs substantial overhead. Recent advancements in hardware-accelerated texture techniques make the texturing of high order surfaces feasible for real-time. These advancements include procedural texture coordinate generation and support for 3D textures. 3D textures provide an additional depth component that allows for effects such as volumetric obscuration of the image, impostering, functional lookup, and procedural texture generation and noise. With a 3D texture, the internal volume of an object can have associated texture information. This will allow for the accurate texture mapping of newly introduced surfaces resulting from cutting, tearing, cauterizing, etc.

Texture synthesis will be used to take source imagery from surgical videos and photographs and generate realistic textures for applying to the simulated tissue in the visual scene. Without texture synthesis, objects look artificial because of the noticeable repetition or stretching in the image. Texture synthesis provides a more natural looking result and can be generated at run-time. It will also be used to manipulate textures for representing tissue discoloration due to tearing, cauterizing, pathologies, etc. Texture synthesis will be used to insert the revised texture information from the patient during information transfer. This process will be continued in the real-time process, using the extracted texture from the ongoing patient video stream.

3. Image preprocessor

An image preprocessor module on the patient side of the communications link is essential to the functioning of the overall system. The purpose of this component is to recognize the organs and various other objects in the surgical field from the video imagery coming from the surgical camera. It will feed this information on the location and states of these objects both to the intelligent controller (on the patient side) and to the simulator (on the surgeon side). This image understanding task encompasses the processing, encoding, and recognition tasks that will be accomplished on the patient side, using the video from the surgical camera(s). The problem of image identification has been a focus of research in the community of image understanding for a long time, and the results still seem to indicate that the most effective approach depends on the context of the application.

A new approach to object recognition and categorization in image is under development by members of our team. A camera image is basically a 2D projection of a 3D world in the field of view of the camera. The approach is to reverse this process by determining the 3D world that generated the 2D image. This approach has been successful in extracting 3D objects, including buildings and trees. Although buildings are generally rectilinear, the objects extracted can be of an arbitrary shape. The software to do this in real time has been developed to enhance 2D images, detect specified objects, extract 3D objects, and remove others from the 2D image.

The extraction process is done in two steps. First, the object is detected or recognized. Second, the 3D parameters are extracted. The method being proposed is geometry-based in that it will use a library of 3D models that can describe what is being seen in the camera (organs, veins, etc.). These models are not rigid models, but rather are models that can change shape depending upon specified parameters. These models are likely to be patient specific, having been generated by an offline process prior to the surgery. The method decomposes the 2D image into regions. Regions are contiguous sets of pixels that are determined to belong to a specified set based on any arbitrary criteria. Examples of these criteria are color or texture. All objects in the 2D scene are composed of one or more of these regions. The regions in the image are then analyzed to determine the boundaries of the 2D projected objects, with the result being input to algorithms for determining the depth component, defined as the axis normal to the plane of the camera. This component for any surface is determined using calculations based on color, texture, and diffuse and specular reflection from the surface. The edges of the 2D objects are found with pattern searches and correlated with the library model data to ascertain the actual 3D structure at the time the 2D image was obtained. The accuracy of the resulting 3D model will depend upon a number of factors, including the resolution of the camera, whether the image is in color or not, the quality of illumination, and the availability of accurate models of the objects being displayed.

4. The intelligent controller

This device, located on the robot/patient side of the communication link, performs in two critical roles. In the ultimate system for use on actual patients, the intelligent controller will be necessary to provide both an added measure of safety and an improved level of efficiency in the presence of time delay. Both the safety role and the efficiency-enhancement role require intelligent behavior.

The need for an added element of safety in the presence of time delay is quite obvious. For a variety of reasons, even when the surgeon as well as the various other components of the system are performing perfectly, the existence of time delay prevents the possibility of 100% certainty as to where various tissue will be in relation to the surgical instruments at any given instant in the future. Because the intelligent controller will be proximate to (and linked directly to) the robot, it will interact with the robot without significant delay, and thereby has the potential to control all robot movements instantaneously. Thus, as a last line of defense against the possibility of accidental collisions between surgical instruments and the patient's vital organs, the intelligent controller will ultimately play a critical role.

The need for improving the level of efficiency over what it would otherwise be in the presence of time delay is also clear. Finishing surgery in a timely manner and preventing unnecessary frustration for the surgeon are always important goals. While it may be true that the delays associated with telerobotic surgery will never allow it to be quite as efficient as proximate robotic surgery, the goal at least must be to make it ultimately as efficient as possible.

Although in the course of the research, we will attempt to apply a variety of advanced approaches to machine intelligence in designing effective intelligent controller prototypes, some basic aspects need not be particularly complex. For the safety aspect, a fairly effective controller could be based on nothing more than a three-dimensional geometric model of the surgical field combined with a simple type of production rule system. A typical rule for the case of gall bladder surgery might look roughly like the following:

IF left end-effector holds sharp instrument AND instrument is within 5 mm of common bile duct AND an override command has not just been submitted by the surgeon
THEN stop movement of left end-effector immediately, send safety alarm signal to surgeon, wait for reset by surgeon

The production rule system for the case of gall bladder surgery may be comprised of perhaps a few dozen such rules. This is a very basic form of the traditional approach to artificial intelligence. Using this as a starting point, we can readily add more sophisticated machine intelligence approaches.

One fairly straightforward addition would be the use of fuzzy logic (FL) in the production rules. FL is simply a calculus for representing mathematically the way humans use somewhat vague concepts, such as “very close” or “rapidly”, when reasoning about complex systems. For example the rule above could be made more sophisticated by changing the antecedent to the following:

IF left end-effector holds sharp instrument AND (instrument is very close to common bile duct OR (instrument is somewhat close to common bile duct AND left end-effector is moving rapidly)) AND an override command has not just been submitted by the surgeon

Naturally, when using fuzzy logic, it will also be necessary to represent in the system those numerical values associated with such terms as “very close”, “somewhat close”, and “rapidly”. It is quite simple to represent such terms in a particular context for relative locations, velocities, accelerations, and any other types of variables we may use.

Fuzzy logic is one component of what is now known widely as “the soft computing approach”. Soft computing (SC) is a term coined by Lotfi Zadeh around 1990 to represent the emerging trend to design complex systems based on hybrids of four component methodologies, each of which had been evolving over the previous three or four decades [20, 21]. These component methodologies of SC are referred to most generically as fuzzy logic (FL), artificial neural networks (ANN), evolutionary computing (EC), and probabilistic reasoning (PR). The key concept of this approach is not just that these four methodologies tend to be powerful in and of themselves, but rather that there tends to be a synergistic effect when two or more of them are combined in appropriate hybrids. The SC approach, also referred to as computational intelligence, may seem to the layman a bit like science fiction, but it is a successful and well established amalgam of methodologies in some fields of engineering and is based ultimately on decades of advanced research.

The success of SC has been demonstrated most graphically in the context of feedback control, particularly in inherently complex control applications. There have been very many citations in the literature of the successful application of SC hybrids, including for example in [22]. Certainly we will experiment with applying them here as well, and we can already state with confidence that they will be useful in the context of the intelligent controller.

We will also experiment with another approach known in general as optimal control techniques, which use a model reference approach. In this case a model of the entire system; patient, robot and surgeon is employed along with a cost function which will be minimized to determine the coefficients of the parameters in the control laws. This need not be viewed as an entirely separate approach in the sense that optimal control concepts are often part of the SC methodology.

5. The integrated tele-robotic surgery system

The total system behaves as the human surgeon would if there were not a performance encumbering delay. Because the simulator through which the surgeon operates is running in real time the surgeon sees reaction to inceptor movements much more quickly than would be the case if he/she were required to wait while the signals made a complete round trip over the long haul network. Fig. 6 is a detailed representation of the proposed system and the general research areas. The following paragraphs briefly explain the architecture.

The signal path from the surgeon’s inceptor movement proceeds to the simulator and simultaneously to the intelligent controller, which commands the robot movement. This simulator is like any other in that it calculates all of the system dynamics in real time and from these computations come changes to the system states, which alter the visual scene observed by the physician. The visual scene is generated by high speed computer graphics engines not unlike those employed by modern flight simulators. However, a unique aspect of the proposed embodiment is that the graphics image is periodically updated by the video image transmitted over the long haul network. This approach ensures that the visual scenes at the patient and at the simulator are never allowed to deviate perceptibly. This update is generated by a complex scheme of image decoding, texture extraction and image format transformation.

The intelligent controller performs the dual role of optimizing robot performance and preventing inadvertent incisions. The research will investigate two general approaches to the design. One approach will use optimal control theory and the other will utilize a hybrid of soft computing techniques (fuzzy control, neural networks and genetic algorithms). Both of these techniques have been used successfully to control autonomous aircraft.

The simulator also calculates appropriate inceptor forces. In the near term, the drive signal math model for the haptic stimuli will be essentially the same as that in the actual robot although it will rely on a sophisticated organ dynamics model to compute the appropriate organ forces interacting with the robot end effectors. Eventually haptic feedback will be applied to enhance the environment for the surgeon. It has been shown in other applications that haptic stimuli, even though artificial, provide information to the operator that improves human performance and dramatically improves the "feeling of presence".

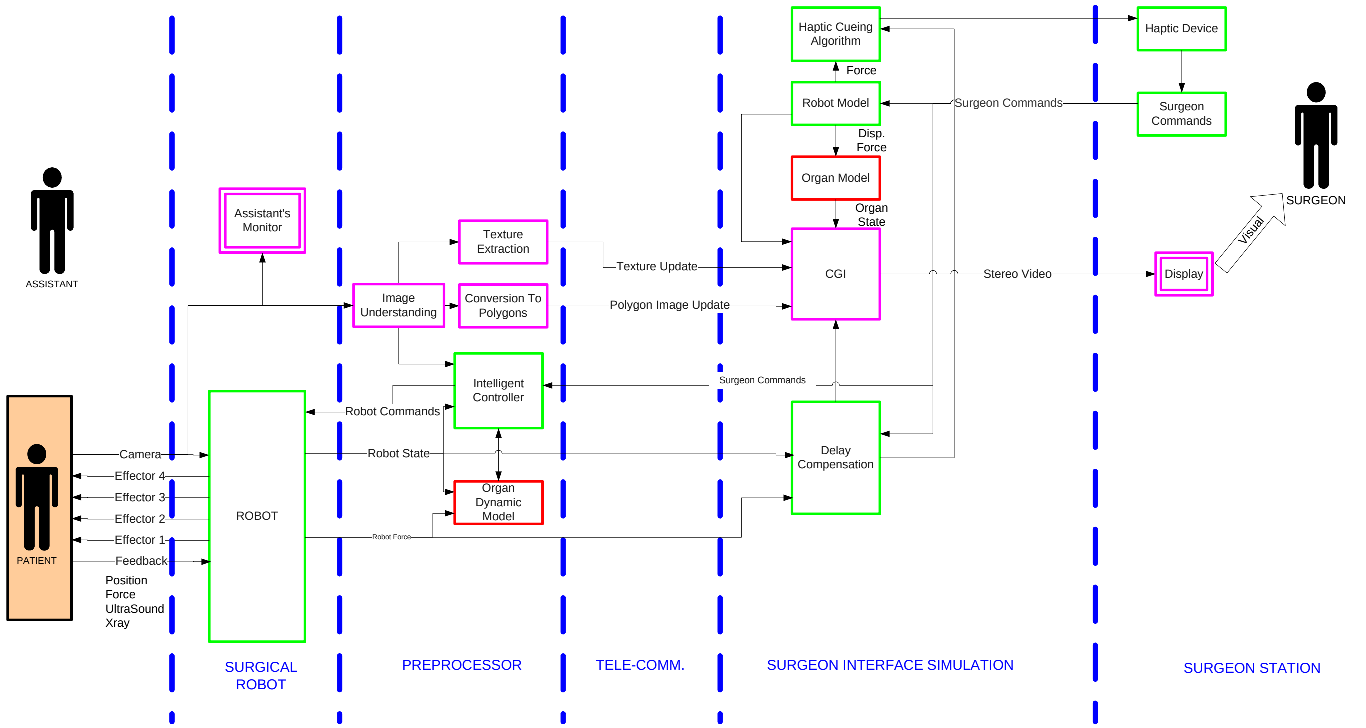


Figure 6 Block diagram of the surgical telerobotics system

Conclusions

The major impediment to remote robotic surgery is the effect of telecommunications delay on the surgeon's performance. An innovative approach to time delay effect mitigation has been presented, which is to have the surgeon operate through a simulator running in real time.

In the proposed telepresence framework the simulator acts as a predictor, providing information to the surgeon consistent with the no delay situation. Clearly, dynamics models both for the robot dynamics and organ dynamics are necessary for the simulator to function in this way. The image preprocessor portion is the essential corrector. The intelligent controller is designed as an invigilator. The total integrated surgical telerobotics system is to behave as the human surgeon would if there were not a performance encumbering delay.

Four major research areas must be explored in order to ensure achieving the objectives of our project. The paper presents a discussion of those key areas as summarized below.

1) Simulator as predictor (Delay analysis and compensation). The delays encountered in remote robotic surgery will be greater than any encountered in human-machine systems analysis, with the possible exception of remote operations in space. Therefore, tests will be conducted to determine the maximum "tolerable" delay. Subsequently novel compensation techniques will be developed. Included will be the development of the real-time simulator, which is at the heart of our approach. The simulator will present real-time, stereoscopic images and artificial haptic stimuli to the surgeon.

2) Organ Modeling. This aspect of the research will be challenging. Our approach is to investigate the relative merits of two different approaches, continuum mechanics and the finite element method. Since the tissue model runs in real time, computational parsimony is imperative.

3) Image Processing. Because of the delay and the possibility of insufficient bandwidth a high level of novel image processing is necessary. This image processing will include several innovative aspects, including; image interpretation, video to graphical conversion, texture extraction, geometric processing, image compression and image generation at the surgeon station.

4) Intelligent Control. Since the approach we propose is in a sense "predictor" based, albeit a very sophisticated predictor, a controller, which not only optimizes end effector

trajectory but also avoids error, is essential. We propose to investigate two different approaches to the controller design. One approach is based on modern control theory; the other one involves soft computing techniques.

References

[1] E. Ballantyne, H. Garth. The Pitfalls of Laparoscopic Surgery: Challenges for Robotics and Telerobotic Surgery. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*. Vol.12, No.1, 1-5, 2002.

[2] R.M. Satava. Surgical Robotics: The Early Chronicles. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*. Vol.12, No.1, 6-16, 2002.

[3] T.B. Sheridan. Musings on telepresence and virtual presence. *Presence*. 1(1), 120-125, 1992.

[4] S.R. Ellis. Presence of Mind: A Reaction to Thomas Sheridan's "Further Musings on the Psychophysics of Presence", *Presence: Teleoperators and Virtual Environments, MIT Press*, 5(2), 247-259, 1996.

[5] M. Lombard, T. Ditton. Measuring Presence: A literature-based approach to the development of a standardized paper-and-pencil instrument. Paper presented at *Presence 2000: the 3rd International Workshop on Presence* (Delft, Netherlands), 2000.

[6] R. Stone. Haptic feedback: A potted history, from telepresence to virtual reality. In *The First International Workshop on Haptic Human-Computer Interaction* (Glasgow, UK) Springer-Verlag Lecture Notes in Computer Science, 1-7, 2000.

[7] L. Guo, F.M. Cardullo, J.A. Houck, L.C. Kelly, T. Wolters. New Predictive Filters for Compensating the Transport Delay on a Flight Simulator. In *Proc. the AIAA Modelling and Simulation Technologies Conference, Providence (RI), August 17-19* (paper number AIAA 2004-5441), 2004.

[8] G.J. Klir. The Role of Anticipation in Intelligent Systems. Computing Anticipatory Systems, ed. D.M. Dubois, American Institute of Physics, 37-46, 2002.

[9] R. Rosen. Anticipatory Systems in Retrospect and Prospect. *General Systems Yearbook*, Vol. 24, No. 11, 1979.

[10] W.H. Levison, R.E. Lancraft, A.M. Junker. Effects of Simulator Delays on Performance and Learning in a Roll-Axis Tracking Task. In *Proc. 15th Annual Conference on Manual Control (AFFDL-TR-79-3134).* Wright Patterson Air Force Base: Air Force Flight Dynamics Laboratory.

[11] J.X. Chen, et al. Knee Surgery Assistance: Patient Model Construction, Motion Simulation, and Biomechanical Visualization. In *IEEE Transactions on Biomedical Engineering*. 48:9. September 2001.

[12] J.X. Chen, S.-M. Kang. Model-order reduction of nonlinear MEMS devices through arclength-based Karhunen-Loeve decomposition. In *Proc. IEEE International Symposium on Circuits and Systems: ISCAS 2001*, Vol. 3, 457-460.

[13] S. Gibson, et al. Simulating Knee Surgery Using Volumetric Object Representations. *CVR Med 1997*, 369-378.

[14] Y.C. Fung. Biomechanics Mechanical Properties of Living Tissue, Second edition. Springer Verlag, New York, 1993.

[15] M.P. Ottensmeyer. In Vivo Measurement of Solid Organ Visco-Elastic Properties. In *Proc. MMVR 2002 - Medicine Meets Virtual Reality 02/10,* 328-333. January 2002.

[16] B.K. Tay, N. Stylopoulus, S. De, D.W. Rattner, M.A. Srinivasan. Measurement of In-vivo Force Response of Intra-abdominal Soft Tissues for Surgical Simulation. In *Proc. MMVR 2002 - Medicine Meets Virtual Reality 02/10,* 514-519. January 2002.

[17] J.T. Samosky, W.E.L Grimson, M.L. Gray. Measurement of Surgical Instrument Force Trajectories and Cartilage Stiffness for Simulation of Arthroscopic Knee Surgery. (http://www.ai.mit.edu/research/abstracts/abstracts2000/pdf/z-samosky.pdf)

[18] V. Vuskovic, M. Kauer, G. Szekely, M. Reidy. Realistic Force Feedback for Virtual Reality Based Diagnostic Surgery Simulators. In *Proc. IEEE ICRA 2000*, 1592-1598.

[19] I. Brouwer, et al. Measuring In Vivo Soft Tissue Properties for Haptic Modeling in Surgical Simulation. In *Proc. MMVR 2001 - Medicine Meets Virtual Reality 01/9*, 69-74. January 2001.

[20] L.A. Zadeh. From Computing with Numbers to Computing with Words—From Manipulation of Measurements to Manipulation of Perceptions. In *IEEE Trans. Circuits and Systems –Fundamental Theory and Applications*, vol. 45, No. 1, 105-119. January, 1999.

[21] J.-S.R. Jang, C.-T. Sun, E. Mizutani. Neuro-Fuzzy and Soft Computing, Pearson Education, 1996.

[22] H.W. Lewis, III. The Foundations of Fuzzy Control, Plenum, 1997.

VICC: Virtual Incident Command Center

Julius S. Gyorfi, Eric R. Buhrke, Mark A. Tarlton, Juan M. Lopez, George T. Valliath
Motorola Labs, Physical Realization Research Center
{julius.gyorfi, buhrke, mark.tarlton, juan.lopez, george.valliath}@motorola.com

Abstract

This paper describes an application of telepresence technology to the incident management domain. The system combines national guidelines for incident management with many aspects of collaborative virtual environments to enable effective communication between first responders in the field and remotely located command personnel. A brief overview of existing incident management systems is given followed by a set of requirements for future systems. We then describe our virtual incident command center (VICC) prototype, explain how it addresses the requirements, and outline our future plans. Finally, we report feedback from ongoing demonstrations of the prototype system that supports our contention that VICC represents a unique solution to the incident management problem.

Keywords--- **incident management, incident command, virtual incident command center, VICC**

1. Introduction

Recent world events have highlighted the need for standardized processes for incident management that span organizational, geographic, and political boundaries. The terrorist attacks on the World Trade Center and the Pentagon, the hurricane Katrina disaster, and the Indian Ocean tsunami of 2004 are just a few examples of large-scale catastrophes that underscore the need for technologies and procedures to enable on-the-spot decision-making, interagency coordination, and the ability to rapidly adapt to changing circumstances. Incident management systems are used by emergency responders to address these problems by establishing command structures, supporting approved operational procedures, and managing resources during an emergency. Communication and coordination between responders is a critical element of incident management. Interactive communication typically takes the form of face-to-face meetings, phone calls, or radio exchanges. Additionally, incident management requires persistent documentation and tracking of resources, organizational structure, and task assignments. These forms of communications include whiteboards, paper documents, and more recently online database systems for data management. Incident management operations become more complicated if: (i) the personnel involved are geographically dispersed, (ii) the topics of discussion require visual information such as maps, diagrams, photographs, and documents, or (iii) the results of discussions need to be documented and communicated.

This paper describes the *Virtual Incident Command Center* (VICC), a prototype application that supports incident command by integrating interpersonal communications with collaboration capabilities involving shared materials like documents, images, and 3D location data. VICC uses a 3D room metaphor to give geographically dispersed users the experience of being present at the same location and engaging in face-to-face interactions. The VICC user interface utilizes three-dimensional maps and interactive displays that can enhance users' awareness of an incident site and give them more control over their information management and decision-making processes. Participants can join or leave conversations as needed, allowing the summoning of remote expertise when necessary.

We begin with an overview of incident command systems and requirements; then, we describe the VICC system and its features.

2. Incident Management

Incident commanders have unique needs for information from the incident site, for additional information from multiple agency and emergency operations center databases, and for sharing information about the incident with key federal, state, and local government agencies. These needs present a complicated technology challenge. The technology needs to be able to provide real-time communications, data transfer, heterogeneous database access, and interoperability for voice and data communications.

2.1. Incident Command System

The Occupational Health and Safety Administration of the Department of Labor maintains documentation and training information for an Incident Command System (ICS) [1] in support of a National Response System. ICS had its origin during the California wild fires of the 1970's. These fires quickly spread across several counties involving many jurisdictions and many different agencies. A common system was needed to manage administration, planning, logistics, finance, and command operations for all of the agencies involved. As a result, the California fire service created ICS to be an emergency management system standard. Its goal was to eliminate problems when supervising people and organizational structures, site information, communications, coordination, planning, terminologies, technologies, and the conflicting objectives of different agencies by specifying rules of hierarchy and interaction between these parties.

ICS was created to provide a management structure as well as a system of best practices for conducting on-site operations. It is applicable to small-scale daily operational activities as well as major mobilizations. ICS provides command center and operational staff a standardized operational structure and common terminology. Hence, ICS is a useful and flexible management system that is particularly adaptable to incidents requiring multi-jurisdictional or multi-disciplinary responses. ICS provides the flexibility needed to rapidly activate and establish an organizational format around the functions that need to be performed during an incident.

There are several disadvantages with this system, however. One is the requirement that all management personnel be collocated in order to work together. This not only increases the time to create an ICS but also makes it vulnerable to failure or attack. On September 11, an ICS was setup at the incident site itself inside the World Trade Center. It was lost when the building collapsed. Another disadvantage is that ICS does not specify ways to leverage technology in order to make incident management processes more effective. These disadvantages led to the development of Virtual ICS.

2.2. Virtual ICS

Virtual ICS is a concept wherein command center participants can share information, make decisions, and deploy resources without being physically present in the command center. Computer terminals and computer networks are used to communicate and share data. This shortens the time needed to setup a command center while also making the command center more resilient by allowing it to continue operating even if one part becomes damaged. These systems rely heavily on web-enabled software. Virtual ICS allows participants to work from their normal workstations, from home, or in the field. Emergency plans and reports are available at any location. All information can be maintained in a central database that is available to command center participants from anywhere in the world. There are many commercial virtual ICS software products available.

These systems are a definite improvement on the basic ICS approach but they give up some of the intangible benefits of the original. Because participants can be remotely located, their ability to collaborate is hampered. Communication is limited to video conferencing at best, but typically uses phone lines. The ability to read body language and gestures is lost. Information sharing is also hampered by this approach. In the classic ICS methodology, the participants can open up a map, lay it on the table, draw on it, point to features, and plan with it. This interaction is lost in virtual ICS. These key sources of information and methods of planning must be communicated by low-technology means. Each participant's information is limited to the size of his display, providing a small portal into the world. This may require the use of multiple displays or it may force the participant to supplement what is on his display by other means (e.g. sticky notes) to organize information that cannot otherwise be displayed.

2.3. National Incident Management System

In 2003, the presidential directive HSPD-5 called for the creation of a national incident management system (NIMS) [2] to provide a consistent approach to federal, state, and local operations when responding to domestic incidents of any size and nature. HSPD-5 required all federal government agencies to adopt NIMS as part of the National Response Plan. The NIMS plan goes beyond ICS by addressing multi-agency coordination issues (i.e. unified command) organized by functional discipline (e.g. fire, law enforcement, medical, etc.) NIMS also called for research and development to solve operational problems and develop new incident management capabilities.

2.4. Unified Incident Command and Decision Support

The National Memorial Institute for the Prevention of Terrorism (MIPT) published a plan, "Project Responder: National Technology Plan for Emergency Response to Catastrophic Terrorism," [3] recommending technology investments in specific areas to improve the twelve National Terrorism Response Objectives it identified; chapter IV in particular focuses on unified incident command, decision support, and interoperable communications. Multimedia Supported Telepresence (recommendation UIC.5) is expressly identified as one of the Needed Functional Capabilities and Priorities for unified incident command. This report assessed the current and near-term telepresence capabilities for incident management to be marginal.

Building upon the recommendations in Project Responder and the NIMS research recommendations, the Department of Homeland Security (DHS), through its Advanced Research Project Agency (HSARPA), solicited proposals for an information management and sharing architecture to support emergency responders and commanders in their tasks [4]. Section 3 of this solicitation notice contains the requirements for a unified incident command and decision support system, henceforth referred to as the UICDS requirements. These requirements subsume the ICS/virtual ICS features, Project Responder recommendations, and NIMS requirements. The UICDS requirements are divided into five broad categories:

1. Seamless Integration
2. Information Integrity
3. Comprehensive Information Management
4. Information Sharing

5. Multimedia/Multimodal Communication

3. Virtual Incident Command Center

We developed the Virtual Incident Command Center (VICC) with the UICDS requirements in mind. VICC brings together many different technologies and integrates them into a powerful new tool for first responders. VICC uses 3D graphics technology to bring people and multimedia data together for easy and effective collaboration in a virtual meeting space. Issues of information overload are addressed by organizing the space in layers where the individual's specific tasks are displayed in the foreground, an incident scene map annotated with responder information is shown in the center of the space, meeting collaborators are arranged around the scene map, and streaming media resources are presented in the background. VICC is intended to permit a geographically dispersed command team to efficiently communicate about an incident and also to enable remotely located subject matter experts to quickly join and participate in the incident management process.

3.1. Architecture

3.1.1. Flexible Design. For a command center to be effective it must be easy, simple, and quick to establish. VICC is a TCP/IP-based client-server system in which the individual clients (i.e. participants) only need to know the IP address of the server to establish a connection. Clients connect via a fast, automated process through which the server configures them based on their access privileges and needs. This provides a flexible design in which new clients and devices can be integrated into the system simply by assigning them (or identifying their existing) IP addresses. This has the added benefit of being able to work with any national incident management infrastructure and eliminates the need to have on-site hardware at an incident location.

Clients can access the system through different devices. There are only a few requirements for these devices. They should have Java virtual machines, 3D graphics capabilities, and minimal amounts of memory to run the client application. They should also be easily configurable or pre-configured to work with simple intuitive user interfaces so that participants can focus on the incident instead of the devices. Once a client is activated, it connects to one or more servers to retrieve lists of available users and organizations. The participant can then immediately begin to interact and collaborate with these groups.

3.1.2. Scalability. Incidents may start small and then grow larger over time. VICC was designed to be scalable to handle large numbers of participants who simultaneously log into the system, organize themselves, and share information. The maximum number of participants in VICC is only limited by the available network bandwidth; the number of participants that the system can support decreases as the amount of data they share increases. As more participants are added, the data transfer rates between the clients and servers will also be affected.

3.1.3. Multiple Platform Support. The client application was developed using Java to be platform-independent. The graphics are based on an OpenGL implementation of the Java Mobile 3D Graphics (M3G) API for J2ME [5]. This API has a small footprint so that it can be used on small handheld devices as well as large workstations, supporting our goal of enabling many different devices to access the system.

3.1.4. Fault Tolerance. VICC was designed to be robust in the face of connectivity problems. If network connections are temporarily lost, both client and server attempt to reestablish the connection. If a server was temporarily disconnected from the clients, the clients will update the server with their current states upon reconnecting. If a client was temporarily disconnected, the server will reset the client's state after reconnecting.

3.2. Collaboration

Enabling collaboration is the fundamental reason VICC was developed. Participants should be able to customize their environments to display data that is relevant and appropriate for their tasks. They should also be able to organize and control the flow of data so they aren't overwhelmed but have access to the critical information they need for decision-making. They should be able to draw on, point to, and manipulate a 3D object like a map of an incident site in real time in conjunction with other remotely located participants. Their interactions should be as natural as if they were seated around the same table. All body language and gesture communication should be preserved while still leveraging all the qualities of a virtual ICS.

3.2.1. Virtual Environment. VICC is a fully interactive 3D virtual collaborative environment designed to meet incident command requirements. VICC can be viewed on a variety of displays: head-mounted displays (HMDs), computer monitors, wall projectors, and handheld device displays. Computer monitors and wall projectors offer the best views of the virtual environment as far as clarity and size are concerned as well as simple interaction capabilities via keyboard and mouse connections whereas HMDs create a more immersive experience but make interaction more difficult. Figure 1 shows a screenshot displayed on a computer monitor of an Operations Center room within VICC from the point of view of one particular participant. This room contains several big screens on the walls for displaying live video feeds, still images, documents, or applications. There is a virtual table around which other participants may be "seated" as they join the room. There are panels around the room with various customizable buttons and displays. The buttons can be set to trigger applications or events and the

Figure 9: Virtual Incident Command Center Operations Center Screenshot

displays can be configured to show any content source associated with the room. This room is only one possible configuration among many; by modifying the underlying 3D model, we can change the appearance and layout of the room.

A key aspect of 3D environments is perspective. VICC incorporates this concept to give participants a common context when discussing elements of their shared environment. For example, if participant A is to the left of participant B, participant C will see the same relationship between A and B from his viewpoint. This gives consistency to verbal descriptions given to other participants, e.g. "look at the image over participant D's right shoulder…" This is an example of the quantitative awareness described by Greenhalgh and Benford [6] and also an application of the "What You See is What I See" and "Peripheral Awareness" principles highlighted by Redfern and Naughton [7] as crucial to computer-supported collaborative work environments.

Another important capability of 3D environments is the incorporation of visualization tools. VICC supports both 2D and 3D visualization tools as well as multimedia applications. These will be discussed later.

3.2.2. User Interface. The VICC user interface was designed to be very user friendly. Originally there was no graphical user interface (GUI) beyond a client window into the 3D environment that allowed interaction through keyboard and mouse events. However, as more features were added to the system, a minimal GUI was added to manage group affiliations, active user lists, and a few client applications that were integrated with VICC. The latest GUI is still simple, but intuitive to use and thus easy to deploy. New users will not be overwhelmed by the complexity of the GUI nor will they need to read through extensive documentation to begin using the system.

3.2.3. Levels of Interaction. There are various levels of interaction in VICC. VICC participants can easily interact with each other through their avatars and voice communications. Participants can interact with the environment through the GUI and client applications such as a 3D pointer to share and exchange information.

Each VICC participant is represented by an avatar. The avatar is the virtual representation of the participant in the virtual environment. It handles all communication and interpersonal interactions. The avatar can be something as simple as an image or an icon or it can be something as complex as a realistic three-dimensional full-body model of the person that is animated by parameters calculated through motion and position tracking algorithms. Three-dimensional avatars are preferred because they can be manipulated to show viewpoint and attention, although it is understood that not every client will have the necessary hardware or capacity to use utilize such an avatar. Such participants can, however, still be represented by a static full-body avatar with audio capabilities. Alternatively, they can participate via an audio/video connection as described by Sun et al. [8].

Participants will always convey their presence in the conversation regardless of how they connect to the system. For example, if a participant is in a fully immersive setup, his avatar will appear in the conference fully animated and fully expressive. But another participant who is connected via a desktop PC may only have an avatar that conveys audio.

In support of the interactive nature of the environment, each participant's avatar will have a mouse-controlled 3D pointer to indicate or manipulate objects in the virtual environment. The pointer's orientation, length, and appearance can be modified by the participant or hidden if not needed. Figure 2 shows a participant viewing an object being indicated by another participant's pointer.

Every object in the virtual environment can be selected and manipulated. For example, the material properties of any virtual object's surface can be changed by first selecting it with the pointer and then changing the desired property via a popup dialog window. Objects can be moved and rotated to enable viewing fine details if necessary. As an example, a 2D map may be loaded into the room and rotated or scaled to enable the participants to identify important sites or routes.

3.2.4. External Data Sources. VICC currently maintains a database of virtual objects and their states to synchronize the views of all the active clients. However, the Seamless Integration and Comprehensive Information Management portions of the UICDS requirements call for the capability to access predefined data sources and use easy-to-construct/integrate databases. VICC does not yet have the capability to integrate existing databases, but it does have the capability to access predefined data sources. For example, VICC can receive streaming biometric data from first responders and then display that data to the incident commander.

3.2.5. Standard Terminology. VICC is intended to be a solution satisfying the UICDS requirements subject to the common terminology and nomenclature specified in the NIMS standard.

3.3. Information Management

3.3.1. Multimedia Support. VICC supports the simultaneous display and presentation of many different data sources. The virtual environment can contain multiple display screens, each of which can display a separate video stream. For example, multiple cameras may be set up around an incident site to provide different views of the scene that can simultaneously be viewed by the incident commander. A participant in such a scenario may need to selectively tune out distracting elements and focus on a single, important video source if necessary (Goebbels and Lalioti [9] investigated factors affecting collaborative virtual environments, including video and audio.)

The display screens can also be used for other content, e.g. still images or slide shows that can be controlled by the participants. Alternately, text documents, web pages, and streaming data can also be displayed. Software applications could be incorporated into VICC if desired, although VICC was not built upon groupware foundations as described by García et al. [10] or embodied in Tixeo's Workspace3D product [11].

The display screens do not have to be fixed to the room. They can be free-floating "billboards" that can be moved around, made translucent, or completely hidden to give the participant leeway to organize the information and layer it in a useful and appropriate way to prevent information overload.

3.3.2. Selective Information Sharing. VICC was developed around a multiple work context paradigm to organize and selectively share information. Different work contexts are represented by virtual rooms in the shared environment. A virtual room is the space where all dialogs and interactions take place, similar to Huang and Ma's concept [12], except in 3D. Virtual rooms can contain tables, chairs, screens, and other objects in the same way that physical rooms do. Virtual rooms can change in size and appearance and can be customized to meet the demands of any situation.

Participants can be invited to a room to join a conversation or collaborative session. Given appropriate permissions, participants can be present in many different rooms at once, although only one room will have the participant's focus at a time.

Participants can alternately be prevented from entering a room based on access privileges and security protocols. Private rooms can be created for selected individuals to discuss confidential information in parallel to public meetings taking place in other rooms.

In support of NIMS and the UICDS requirements, participants may be grouped together by functional areas, security levels, and/or areas of expertise. Each of these groups can have its own room.

As an incident grows from small to large, new rooms can be added. For example, the finance department of the incident command team may require the addition of several people to its staff. A new finance room can be created to contain all the people needed to manage the budget of the incident. The lead finance team member can maintain a presence in both rooms to manage the finance operations and report to the incident commander.

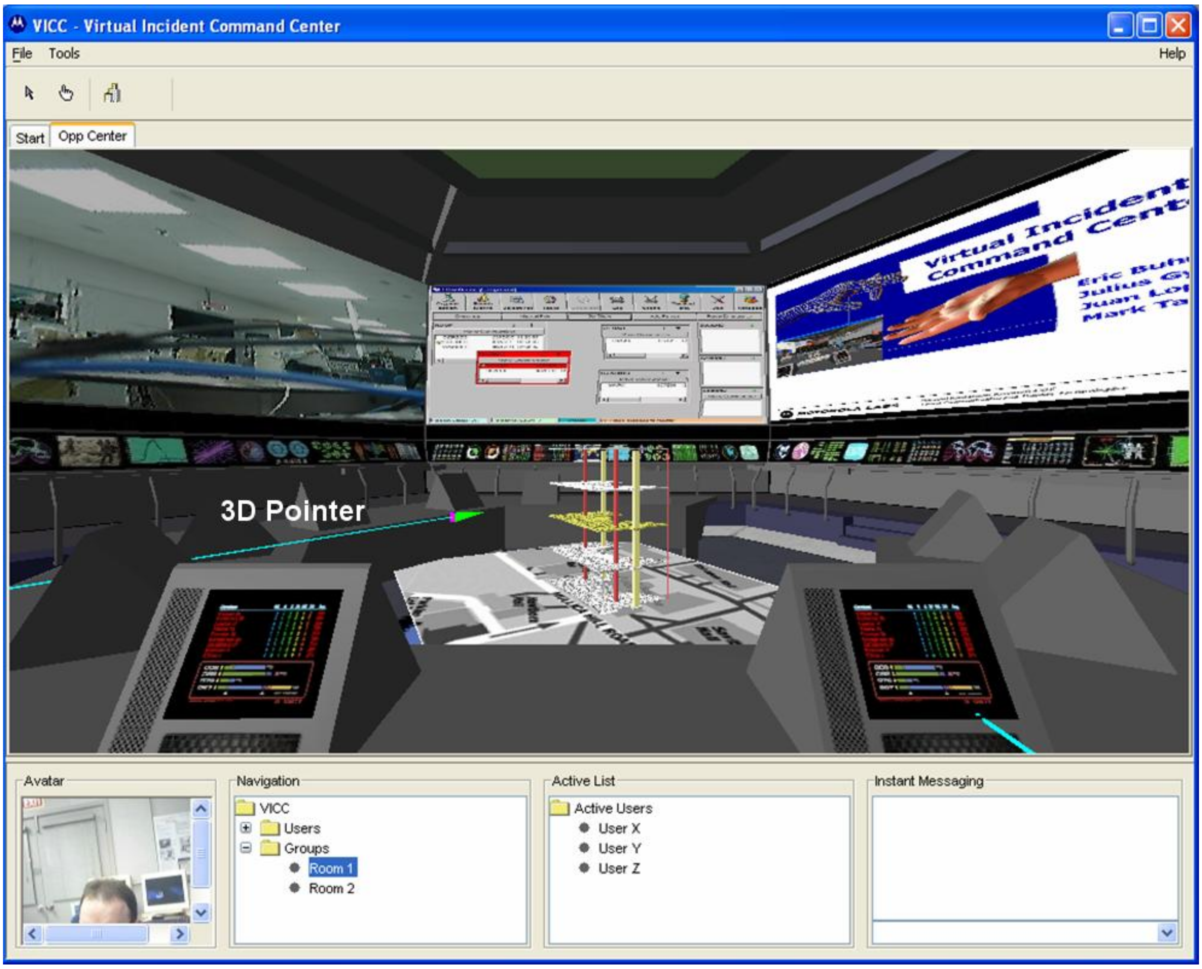


Figure 10: Remote Participant using 3D Pointer to Indicate Model of Incident Site

3.3.3. Graphic Location Displays. Yet another benefit of the 3D world is the ability to include three-dimensional objects in a discussion. These objects can be analyzed, manipulated, and animated (see Pinho et al. [13] for a cooperative object manipulation framework and Hindmarsh et al. [14] for an object-focused interaction study.) One example of such an object is a building model. A 3D model of an incident can be loaded, if available, and manipulated, viewed, and annotated by the participants to examine floor plans, evacuation routes, etc. Additionally, if streaming location data is available for the responders on the scene, this data can be mapped in real-time onto the build model to help keep track of everyone.

3.3.4. Logging Incident Information. VICC currently maintains logs of system calls, object state changes, and client-server messages. However, in order to be fully compliant with the information logging requirements of the Comprehensive Information Management portion of the UICDS requirements, we will need to add the capability to record all the actions that take place in the virtual environment to provide source material for post-mortem evaluations of incident responses and lessons-learned documentation. It will also be necessary to add the capability to index and cross-reference this information with other data. These last two capabilities are currently lacking in our system but will be dealt with in the context of decision support.

3.3.5. Decision Support Capability. Decision-making is a key duty of an incident commander. Decision-making requires timely access to information, logging the processes involved, gaining rapid access to subject matter experts, and getting rapid approvals. Decision support capabilities in VICC will augment the incident commander's own competencies with knowledge management applications to assist these activities. We are currently investigating this area, including existing work in the area of situation awareness in command in control settings such as the TADMUS project [15].

Table 1: VICC Features Mapped to the UICDS Requirements.

		DHS UICDS Reqs				
		Seamless Integration	Information Integrity	Comprehensive Information Mgmt	Information Sharing	Multimedia/Multimodal Communication
Architecture	Flexible Design	✓	✓			
	Scalability					✓
	Multiple Platform Support					✓
	Fault Tolerance		✓			
Collaboration	Virtual Environment			✓		✓
	User Interface			✓		✓
	Levels of Interaction	✓		✓		✓
	External Data Sources	✓	✓			
	Standard Terminology	✓				
Information Mgmt	Multimedia Support			✓		✓
	Selective Information Sharing			✓	✓	
	Graphic Location Displays			✓		
	Logging Incident Information			✓		
	Decision Support Capability			✓		

Conclusions

We have developed a functional prototype of a virtual incident command center that addresses many of the UICDS requirements (see table 1 for a mapping of VICC's features to these requirements.) We described the current capabilities of the system as well as planned future enhancements. We are also pursuing the possibility of commercializing the system by transferring this technology to a product development organization that would then perform systematic user testing and trials before creating a finished product. Beyond this, we plan to conduct more research into virtual collaboration and develop additional use cases for VICC outside of incident command.

VICC has been demonstrated numerous times to many different individuals and organizations, almost always resulting in great enthusiasm from those who have viewed it. We have demonstrated the system to local law enforcement officials and fire officials who have expressed interest in using it as soon as it is available. In addition, the system has been shown to governmental officials at the local, state, and national levels. The system has also garnered international exposure through demonstrations to officials from the governments of several foreign countries. These demonstrations also stimulated discussions regarding user interfaces to unified incident command, maintaining situational awareness, and dealing with information overload. Based on the amount of positive feedback we have received, we are confident that VICC represents a step forward in incident command and that it is a novel application of telepresence technology to this domain.

Acknowledgements

The authors would like to thank Iwona Turlik, director of the Physical Realization Research Center of Motorola Labs, and Tom Babin, manager of the Knowledge Discovery and Applications group, for their support of this project. The authors would also like to thank the members of Advanced Technology Strategy group of Motorola's Networks and Enterprise business for their interest in VICC and their input and assistance during its development.

References

[1] Incident Command System, Occupational Safety & Health Admin.: http://www.osha.gov/SLTC/etools/ics/index.html

[2] National Incident Management System, Dept. of Homeland Security, http://www.fema.gov/pdf/nims/nims_doc_full.pdf

[3] "Project Responder: National Technology Plan for Emergency Response to Catastrophic Terrorism." Edited by Thomas M. Garwin, Neal A. Pollard, and Robert V. Tuohy, published by Hicks and Associates, Inc. for The National Memorial Institute for the Prevention of Terrorism and the United States Department of Homeland Security, April 2004.

[4] Department of Homeland Security, Homeland Security Advanced Research Programs Agency (HSARPA), Broad Agency Announcement 04-14 (BAA 04-14), "Innovative Architectures for Unified Incident Command and Decision Support (UICDS)." October 4, 2004

[5] JSR 184: Mobile 3D Graphics API for J2ME: http://www.jcp.org/en/jsr/detail?id=184

[6] C. Greenhalgh, S. Benford. MASSIVE: a collaborative virtual environment for teleconferencing. *ACM Trans. on Computer-Human Interaction*, vol. 2, no. 3, pp. 239-261, 1995.

[7] S. Redfern, N. Naughton. Collaborative virtual environments to support communication and community in internet-based distance education. *J. of Information Technology Education*, vol. 1, no. 3, 2002.

[8] L. Sun, Y. Zhong, Z. Zhong. Natural interaction synthesizing in virtual teleconferencing, In *Proc. of the Intl. Conf. on Image Processing*, vol. 2, pp. 405-408, 2002.

[9] G. Goebbels, V. Lalioti. Co-presence and co-working in distributed collaborative virtual environments. In *Proc. of the 1st Intl. Conf. on Computer Graphics*, Virtual Reality and Visualisation, pp.109-114, 2001.

[10] P. García, O. Montalà, C. Pairot, R. Rallo, A. G. Skarmeta. MOVE: component groupware foundations for collaborative virtual environments. In *Proc. of the 4th Intl. Conf. on Collaborative Virtual Environments*, pp. 55-62, 2002.

[11] Workspace3D: http://www.tixeo.com/Products.htm

[12] R. Huang, J. Ma. A general purpose virtual collaboration room. In *Proc. of the 5th Intl. Conf. on Engineering of Complex Computer Systems*, pp 21-29, 1999.

[13] M. S. Pinho, D. A. Bowman, C. M.D.S. Freitas. Cooperative object manipulation in immersive virtual environments: framework and techniques. In *Proc. of the ACM Symp. on Virtual Reality Software and Technology*, pp. 171-178, 2002.

[14] J. Hindmarsh, M. Fraser, C. Heath, S. Benford, C. Greenhalgh. Object-focused interaction in collaborative virtual environments. *ACM Trans. on Computer-Human Interaction*, vol. 7, no. 4, pp. 277-509, 2000.

[15] J. A. Cannon-Bowers and E. Salas. Individual and team decision making under stress: Theoretical underpinnings. In *Making decisions under stress: Implications for individual and team training*, J. A. Cannon-Bowers and E. Salas, Eds. Washington, DC: APA Press, 1998, pp. 17-38.

The Virtual Immersion Center for Simulation Research: Interactive Simulation Technology for Communication Disorders

Stacy L. Williams
Case Western Reserve University
{Slw30@case.edu}

Abstract

The Communication Sciences Department at Case Western Reserve University is building an immersive virtual reality (IVR) cave simulation research laboratory to increase student clinical competency skills and carry over of therapeutic skills for children and adults with communication disorders. The use of this IVR simulation lab known as the Virtual Immersion Center for Simulation Research (VICSR) has the potential to provide computerized training in a safe, controlled, learner-centered environment wherein students and clients can repeatedly practice a range of skills in natural contexts. Previous research has shown that VR learning environments demand a high level matrix of knowledge, skill and judgment; qualities that contribute to successful, competent clinicians and rehabilitated clients [1]. The focus of VICSR is to design and implement a clinical VR simulation training program to target a variety of communication disorders to determine its effectiveness in increasing student clinical competency skills and generalization of therapeutic skills for clients with speech and language disorders.

Introduction

According to a recent survey conducted by the University of Cincinnati (2004), speech-language pathology (SLP) graduate students often feel apprehensive in their clinical abilities when diagnosing and treating childhood speech and language disorders. They report difficulty with understanding children's developmental levels, applying and interpreting appropriate assessment measures, and translating this into an appropriate therapy program [2]. These skills are critical for successful speech-language clinicians.

Furthermore, a large number of clients with speech and language disorders often report high levels of frustration with their therapy programs due to lack of generalization [3]. Owens (2005) defines generalization as the transfer of speech and language gains made in treatment to real world settings [4]. Generalization is the most challenging part of any client's speech and language rehabilitation program – learning to apply newly acquired therapeutic techniques and skills in a variety of contextual, natural based settings.

Given these data, there is an urgent need to examine learning from a practice-based approach using routine, everyday activities which promote successful transfer of clinical and therapeutic skills. This practice based approach is referred to as a situated learning paradigm [5]. According to this learning model, knowledge is situated; it is a product of the activity, context and culture in which it is developed and used [6].

There are a number of advantages to a situated learning paradigm such as (1) learning is always contextual, situated in a particular everyday practice (2) learning is culturally and socially diverse in nature (3) learning is dynamic with a community of learners [5]. Research shows that situated learning experiences are highly motivating and create autonomous learners – gaining insight and understanding into information presented [7].

Immersive virtual reality (IVR)-based learning environments support situated learning paradigms [7]. Immersive virtual reality can be defined as an artificial environment or setting that is created by a computer in which the user feels present [8]. Steuer (1992) refers to this phenomenon as telepresence – the extent to which the user feels present in an artificial environment rather than the immediate physical environment [9]. This sense of immersion or telepresence in a virtual setting is critical for creating a situated learning environment that promotes transfer of skills to real world situations.

The Virtual Immersion Center for Simulation Research (VICSR) has the potential to provide computerized training in a safe, controlled, learner-centered environment wherein students and clients can repeatedly practice a range of clinical skills without endangering patients. VICSR has the power to create an effective virtual learning environment for the field of speech language pathology that will promote transfer of skills into real world settings.

Initial Research Agenda for VICSR

The research agenda for the creation and implementation of this IVR simulation training center consist of the following:

1. To determine the effectiveness in increasing clinical competency skills for SLP graduate

students for identifying speech and language disorders using IVR simulation technology compared to traditional based learning methods

2. To determine the effectiveness in increasing speech and language skills for clients with communication disorders using IVR simulation technology compared to traditional intervention learning methods.
3. To measure transfer of speech and language skills to every day situations for clients with communication disorders using IVR simulation technology compared to traditional intervention learning methods.
4. To explore the relationship between learning environments and telepresence for SLP graduate students and clients with communication disorders

Significance of VICSR

While virtual reality training simulation technology has been applied to the military, airline industry and a variety of medical specialties, it has yet to be investigated in the field of speech-language pathology. A field that requires high levels of patient interaction and communication skills.

In the medical field, an increasing number of higher education institutions are beginning to seek more effective and efficient means in both teaching and assessing their medical students' implementation of clinical knowledge and critical reasoning skills. A survey of medical institutions found that 73 of 124 schools are using some form of simulation for evaluation or teaching purposes [10]. Students participating in computer simulations performed significantly better and faster when compared to students in a traditional classroom [11]. Furthermore, studies have found that VR training simulations can improve the acquisition and retention of knowledge in comparison with the traditional lecture [12].

Successful use of VR technology with patients requiring rehabilitation has been demonstrated in fields such as cognitive rehabilitation [13] spinal cord injury [14], children with visual impairments [15] and children with Asperger's Syndrome [16,17]. Preliminary data regarding VR technology with these patient populations has shown that applied VR training shows promise for learning specialized skills and generalization of those skills into natural contexts.

These research studies regarding the effective use of computer simulation demonstrate the need to not only expand simulation technology to other fields, but improve upon the technology to create a highly sophisticated simulated environment. This optimal learning environment would allow learners to interact and carry out "sustained deliberate practice with expert feedback which is an essential prerequisite for expert performance, yet within an authentic context relevant to learning needs" [1].

Steuer (1992) provides the conceptual framework necessary for developing new IVR simulation technology that incorporates visualization and interactivity - key telepresence variables that promote learning and bridge the transfer of knowledge to real world situations [9]. This sense of "being in a virtual environment" is critical for students and clients to become engaged and active in the overall learning process [18].

Creation of VICSR

Recently, the Department of Communication Sciences at Case Western Reserve University and Virtra Systems, Inc. has designed a revolutionary interactive 180 degree immersive virtual reality (IVR) cave simulator for communication disorders called the Virtual Immersion Center for Simulation Research (VICSR). This IVR-Cave research environment surrounds the user with three eight by ten foot screens and enables them to experience an interactive training simulation which utilizes state-of-the-art rear projection-based technology.

VICSR's features include:

- Customized two dimensional branched HD films targeting "real life" scenarios
- External displays and supplemental computer monitoring/recording systems to track users virtual experiences
- Specialized authoring software allowing for high levels of social interactions with "real life" scenarios
- Dynamic instructor station allowing SLP clinicians/instructors the ability to control social interactions within a given virtual setting
- Biometric feedback data for monitoring indices of telepresence throughout a virtual experience (heart-rate, skin conductance, respiration)
- Observation room for family members, teachers and other professionals to observe virtual experiences

Interactive IVR simulation films consist of multiple branched video segments to allow for a variety of user interactions and are created using high definition (HD), two-dimensional recording devices. Scripts for the HD films are created by faculty members from the Communication Science Department at Case in collaboration with the Cleveland Hearing and Speech Center.

MediaVision, a Department of Instructional Technology and Academic Computing (ITAC) at Case carries out the IVR filming production at low cost which promotes interdisciplinary research to be conducted into effective real learning experiences for students as well as clients with communication disorders. The interactive nature of the IVR environment is the unique element that makes this simulator one of the first of its kind.

Future Research Initiatives for VICSR

This simulation technology allows for the education, evaluation and preparation of students and patients with a variety of communication disorders targeting "real life" scenarios. This situated approach to learning speech-language pathology offers the potential to increase generalization of both clinical skills and therapeutic skills into everyday situations. This is the premiere objective for every student and client with a communication disorder.

Currently, IVR film production is targeting two research initiatives:

Simulation program for working with clients with communication disorders targeting fast food restaurant ordering

The first IVR film production has targeted aMcDonald's fast food restaurant setting to work with clients with stuttering disorders as well as non-verbal communication disorders requiring the use of augmentative communication devices for ordering food items at the counter as well as the drive-up window.

The McDonald's IVR simulation film consists of several branched segments allowing clients the ability to interact freely with the McDonald workers and restaurant patrons. Interactions with the McDonald's workers and patrons are determined by the questions, comments and social behaviors derived from the client in the simulator. The speech-language pathologist controls the overall interaction by selecting the appropriate response and behaviors for the virtual worker. Hence, there is a need for multiple branched film clips to allow for a variety of interactions and outcomes.

VICSR is equipped with digital video recording equipment to provide visual support and feedback to clients following each virtual experience. In addition, clients will be encouraged to wear biometric feedback instrumentation to track their heart rate, skin conductance and respiration to gage autonomic levels of presence during interactions in the virtual settings. Clients will be permitted to repeat the IVR- simulations as many times as they desire since there are many outcomes depending on their communication skills.

Preliminary data for determining the effectiveness of interactive IVR-CAVE simulation training will be available Fall 2006.

Simulation program for training graduate students diagnostic procedures for identifying speech and language disorders

The second IVR film production has targeted a third grade child with a suspected speech and language disability and her family for graduate students learning to become future SLPs. This patient scenario consists of several branched segments allowing students to interact freely with the virtual patients and their families while attempting to determine if a speech and language problem is present.

SLP graduate students will be provided with a basic case history of the virtual patient prior to entering VICSR. A designated instructor will be seated at the instructor station so that they can control the responses and behaviors of the VR patient. Once in the simulator, the student will introduce himself/herself to the VR patient and his/her family and begin conducting a thorough case history. Interactions with the virtual patients and their families are determined by the questions and behaviors derived from the student in the simulator. Students can request to observe the child in a particular setting if they feel they need additional information to make an informed decision, but this will only be made available upon the student's request.

Once the student feels that they have obtained all necessary background information, the student will complete a more in-depth analysis of the case by determining appropriate diagnostic assessment protocol and administering a thorough diagnostic battery. After the evaluation process, SLP graduate students will determine an appropriate diagnosis and consult with the VR patient and family regarding treatment options.

Digital tracking capabilities and biometric feedback data will be recorded throughout the virtual experience for student performance feedback and physiological measurement of presence.

Preliminary data for determining the effectiveness of interactive IVR-cave simulation training will be available Fall 2006.

Summary

IVR simulation technology for communication disorders has many uses beyond preparing students for real world challenges; it also has the power to offer new therapeutic training techniques for children and adults with a variety of communication disorders. The goal of VICSR is not only to develop IVR training modules for students to become expert clinicians in their fields of study, but to offer this technology as a means to promote generalization of therapy skills into real life settings. IVR technology has the ability to offer an unlimited supply of lifelike settings for students and clients with speech and language disorders to practice their clinical skills in a safe and controlled environment.

References

Reality and Associated Technologies, Alghero, Sardinia, Italy, (eds. P.Sharkey, A. Cesarani, L. Pugnetti & A Rizzo), pp.67073. University of Reading, Reading.

[1] Kneebone, R.L., Scott, W., Darzi, A., & Horrocks, M. (2004). Simulation and clinical practice: strengthening the relationship. Medical Education, 38, 1095-1102.

[2] Rudolph, C & Breen, P. (2004) [Supervisors feedback regarding SLP student clinician performance in practicum settings]. Unpublished raw data.

[3] Fey, M.E. (1988). Generalization issues facing language interventionists: An introduction. *Languae, speech and Hearing Services in Schools*, 19, 272-281.

[4] Owens, R.E. (2004) *Language disorders: A functional approach to assessment and intervention*. Boston: Pearson Education, Inc.

[5] Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

[6] Wilson, B. & Myers, K.M. (2000). Situated cognition in theoretical and practical context. In D. Jonassen & S. Land (Eds), Theoretical foundations of learning environments (pp. 57-88). Mahway, JF: Erlbaum.

[7] McLellan, H. (1991). Virtual environments and situated learning. *Multimedia Review*, 2(3), 25-37.

[8] Jacobson, L. (1993). Welcome to the virtual world. In: Richard Swadley (Ed.). *On the cutting edge of technology* (69-79). Carmel, IN: Sams.

[9] Steuer J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication* 4(2) 73-93.

[10] Gordon, J.A., Wilkerson, W.M., Shaffer, D.W., & Armstrong, E.G. (2001). Practicing medicine without risk: Students' and educators' responses to high-fidelity patient simulation. *Academic Medicine*, 76, 469-72.

[11] Holmboe, E.S. & Hawkins, R.E. (1999). Computers and evaluation of clinical competence. *Annals of Internal Medicine*, 130, 244-245.

[12] Grantcharov, T.P., Kristiansen, V.B., Bendix, J., Bardram, L., Rosenberg, J. & Funch-Jensen, P. (2004). Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery*, 91, 146-15

[13] Rothbaum, B., & Hodges, L. (1999). The use of virtual reality exposure in the treatment of anxiety disorders. *Behavior Management* 23, 507-25.

[14] Riva, G. (2000). Virtual reality in rehabilitation of spinal cord injuries: a case report. *Rehabilitation Psychology* 45,81-8.

[15] Sanchez, J. & Lumbreras, M. (2000). Usability and cognitive impact of the interaction with 3-D virtual interactive acoustic environments by blind children. In: The 3rd International Conference on Disability, Virtual

[16] Cobb S., Beardon L., Eastgate R., Glover T., Kerr S., Neale H., Parsons S., Benford S., Hopkins E., Mitchell P., Reynard G. & Wilson J. (2002) Applied virtual environments to support learning of social interaction skills in users with Asperger's Syndrome. *Digital Creativity* 13(1), 11-22.

[17] Parsons, S. & Mitchell (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research* 46(5), 430-443.

[18] Jelfs, A. & Whitelock, D. (2000). The notion of presence in virtual learning environments: what makes the environment “real”. British Journal of Educational Technology 31(2) 145-152.

The Effects of Fully Immersive Virtual Reality on the Learning of Physical Tasks

Kayur Patel
University of Washington
{kayur@cs.washington.edu}

Jeremy N. Bailenson
Stanford University
{baileson@stanford.edu}

Sang-Hack Jung, Rosen Diankov, Ruzena Bajcsy
University of California at Berkeley
[sangj,ffxvross,bajcsy]@eecs.berkeley.edu

Abstract

Fully immersive virtual settings are different from traditional virtual reality settings in that they are able to capture full body motion. This ability allows people to use their full range of physical motion to interact with other avatars, computer controlled agents, and objects in the virtual environment. As such, fully immersive virtual reality presents a novel mediated learning environment in which people can learn physical activities. Capturing human motion for virtual settings has traditionally been a model-based approach where a few degrees (on the order of tens) of freedom are mapped to virtual model. In contrast, we use an image-based solution that sacrifices visual fidelity for motion fidelity and increased degrees of freedom (on the order of hundreds). Due to the difficulties involved with building such an image-based immersive system, very little work has been done to assess the effectiveness of this form of mediated learning. In the current work, participants were taught several tai chi moves in either a 2D video system or a 3D immersive system equipped with features not possible to implement in traditional video systems. We demonstrated via blind coder ratings that people learned more in the immersive virtual reality system than in the 2D video system, and via self-report ratings the social presence was higher as well. We discuss these findings and the resulting implications for designing and testing fully immersive systems.

Keywords--- **virtual reality, human computer interaction, mediated learning, computer vision**.

1. Introduction

From entering a holodeck on a starship to plugging into the matrix, our culture has been fascinated with virtual settings, indistinguishable from reality, which we can augment at our whim. In the pursuit of this ideal, researchers have developed various virtual reality (VR) systems. Creating high fidelity virtual environments is a difficult endeavor; no current system has adequately solved this problem. Nevertheless, as systems have improved, VR has proven useful for various domains; in particular, it has been shown to be a successful environment for learning [1].

Research has shown that, for learning physical motion, repeating a task reinforces learning [2]. Virtual reality, as a platform for mediated learning, provides instruction on-demand, allowing students to repeat difficult motions whenever instruction is needed and as often as is needed away from the social pressures of a classroom setting. Learning adapts to the user's schedule, goals, and speed. Previous research has shown the affordance of on-demand learning provides a distinct advantage over face-to-face human interaction [3].

While some virtual systems have proven successful for learning, most VR systems lack the ability to capture the user's full range of motion, limiting their ability to fully immerse the user in the virtual setting. In contrast, fully immersive virtual reality allows for full mobility by capturing human motion and reproducing the same motion in the virtual representation. Full immersion is crucial to any type of learning activity that involves any type of body coordination such as medical training for surgery, learning physical therapy exercises, recreational activities (e.g., martial arts, dance, yoga), and manual skills (e.g., repair, combat training). Recent advances in computer graphics, computer vision, motion capture, and computer power have made it possible to build systems that allow us to assess the effectiveness of fully immersive virtual reality [4, 5].

Full immersion can be achieved though model-based techniques (e.g., capturing a few degrees of freedom and reproducing the captured human motion in a 3D avatar) or image-based techniques (e.g., creating a model of the human from large scale camera arrays and computer vision). Image based methods provide higher degrees of freedom and more accurate representations of the individual at the expense of model quality. Since empirical data supports the notion that motion fidelity is more important and visual fidelity [6], we believe an imaged based methods are most promising for learning fully body motions.

In the current study, we compare learning in an image-based fully immersive VR to instructional videos, the current ubiquitous method of learning full body motion without human interaction. Video-based learning [7, 8] is a good example of mediated learning since it affords instruction on-demand. Videos have a particular advantage over books in that they allow the user to view live, fluid motion of an expert performing a motion. By aggregating the results from 63 separate papers, McNeal and Nelson [9] show that across many different contexts video is a more effective form of instruction than books.

Immersive virtual reality extends the affordances of video, allowing the user to enter the same world as the teacher. First, immersive settings allow users to see in full three dimensions, greatly increasing detail, presence (i.e., learners feel psychologically as if they are in the digital learning environment, as opposed to the physical space [10, 11]), and social presence (i.e., they feel as if the digital reconstruction of the instructor is a real person [12]). Second, as opposed to stationary video, immersive virtual settings allow users to control how they view the environment by allowing them to change aspects such as camera position and orientation.

Third, video settings only allow users to watch the instructor; immersive virtual reality allows the user to interact with the instructor and the environment, as well as to perform novel functions such as sharing body space with the instructor.

In the current work, we compare image-based immersive technology and the established video training tools in their effectiveness in teaching tai chi. We choose tai chi as the learning context because it involves complicated full-body motion and provides clear guidelines for correct performance, and previous work on model-based fully immersive virtual reality has utilized similar learning content [4].

2. Related Work

Since the advent of personal computing and the World Wide Web, there has been a surge of interest in the field of technology mediated learning [13, 14, 15, 16]. Virtual reality has proven to be a promising field for mediated learning, leading to a variety of solutions that address the need for flexible, on-demand training [1]. Virtual training has been met with wide success in the fields of aviation (e.g., flight and space simulators [17]), military (e.g., mission training [18]), medicine (e.g., invasive surgery [19, 20]), emergency (e.g., fire fighting and paramedics [21, 22]), art (e.g., calligraphy [23]), and classroom education [48],

Most empirical research measuring presence between people and virtual humans (either human controlled avatars or computer controlled agents) has utilized systems that track and render the virtual humans along a finite set of degrees of freedom (see [24], for a recent review). In the current work, we are one of the first studies to examine social interaction and social presence using a system in which a user's entire head and body are tracked with extremely high resolution. Our results demonstrate that the high levels of immersion afforded by the fully tele-immersive system result in a higher degree of self-reported presence than video systems.

However, it must be noted the definitions of presence vary between author, and the current research was in no way designed to attempt to provide insights into the intricacies of the semantic definition of presence. Instead, the research took a simulation that was extremely immersive and tested to see if people learned better in it than in a traditional video.

For learning physical motions involving the entire body, it is essential that the user's actions in the physical space correspond to actions rendered on the digital representation of the user (i.e., the avatar) in the virtual space. Yang gives a high level overview of both the hardware and software used in many fully immersive systems, including the image-based system utilized in this study [5], and a typical method of extracting three dimensional models of humans in real-time using off the shelf components is described by Daniilidis and colleagues [25].

In another study, Chua has developed a model-based fully immersive virtual reality system that uses a head mounted display (HMD) along with motion capture to help a student learn tai chi [4]. Their study utilized algorithms that measured deviation between joint movements between the student and the teacher while the student was trying to mimic the teacher, and did not show any advantage in performance for any virtual reality conditions over control conditions. Though Chua and colleagues' experiment provided an important foundation to the study of immersive learning, the current work provides an approach that combines immersive technology with rigorous social scientific methodologies. Consequently, while the previous work did not demonstrate improvements in learning, the current work provides measurement tools sensitive enough to demonstrate the superiority of virtual reality systems over other forms of learning media. There are four notable advancements of the current study over previous work studying learning in virtual reality.

First, because our system is a projection-based system, there is no cumbersome HMD to inhibit naturalistic body motions. We provide extremely high presence without having to drastically limit the user. Second, the image-based reconstruction is not limited to a set number of measured degrees of freedom in the way that model-based motion capture systems are, so the user's physical motions are highly realistic. Third, given that learning such a high level system of movements is likely a gestalt phenomenon, we evaluated our system both via self-report and blind coder ratings. These methods give us the power to test how well participants learn the overall gestalt of the tai chi system of movements, as opposed to only analyzing the success of the learning on a micro, joint-by-joint basis.

Finally, in the current study, we not only measure performance while the virtual teacher is present, but also measure learning, that is, how well the student can perform the tai chi moves later on outside of the simulation when the teacher is no longer present. In sum, the current research is unique both technologically in terms of graphics and methodologically in terms of social scientific learning measurement.

3. Study

We ran a study in our immersive virtual reality system that reconstructed a three-dimensional model of participants as they moved. We placed the participant in a virtual setting where they could interact with their surroundings using natural movement in a tai chi learning environment. Our choice of tai chi was motivated by its slow, fluid motions, which allow us to capture fluid motion in real-time and allow the student to better follow the teacher. Moreover, since tai chi is well established, there is a correct way to perform each move. This allowed us to accurately judge the performance of each participant. Previous research examining virtual performance has utilized similar tasks [4].

3.1. Design

We compared the differences between immersive virtual reality and video as a way to isolate the effect of immersion. Since training videos for physical activities are ubiquitous, we find that this is an apt and crucial comparison for gauging the impact of immersive environments for training physical motion. In this paper the video condition refers to learning from two-dimensional images, and the virtual reality condition refers to interaction in three dimensions with feedback.

3.2. Apparatus

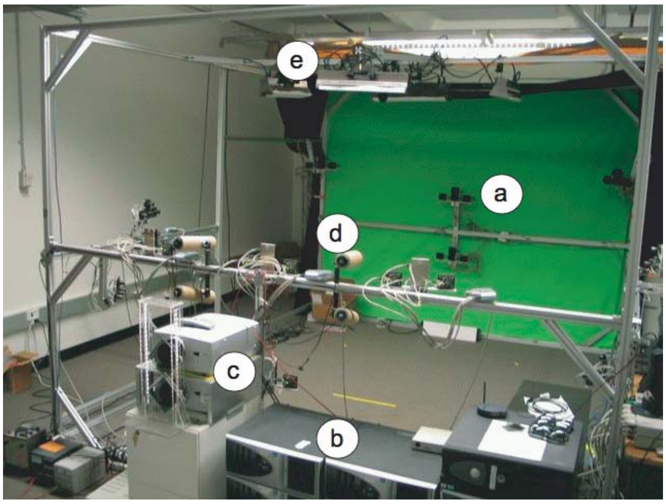


Figure 11: The physical workspace for our immersive reality setting. The camera clusters (a) capture images that are processed by the computers (b) and projected from the projectors (c). The infrared cameras (d) and ceiling lights (e) assist our vision algorithm in capturing a three-dimensional model of the participant.

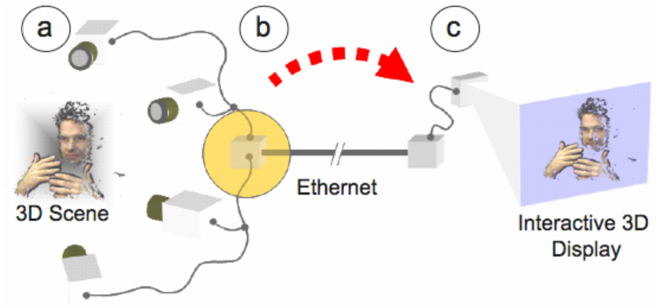


Figure 12: Our system consists of three components image processing (a), data transmission (b), and visualization (c).

Previous work describes our hardware in detail [5]. Figure 1 shows the physical environment in which the participant's image is captured. Our system consisted of three components: image processing, data transmission, and visualization. Figure 2 shows an overview of our system.

In the image-processing phase (Figure 2a), twelve camera clusters (Figure 1a), consisting of three black and white cameras and a color camera, sent images to dedicated computers. Each computer then processed its image (Figure 1b) and computed a 3D representation of the scene from different viewpoints. Partial 3D representation from each viewpoint was combined later in the rendering machine to produce a full representation of the scene. The multiple cameras provided full coverage within a physical workspace of 1.2 meters by 1.2 meters by 2.1 meters.

To enhance the performance of the 3D reconstruction algorithm, ceiling and ground lights illuminated the physical setting (Figure 1e). Infrared cameras (Figure 1d) projected patterns onto objects within the workspace. These patterns did not appear on the model; however the computer used them to increase reconstruction accuracy.

The second component of the system was the network transmission of 3D data (Figure 1b). To achieve immersive virtual environments that allow realistic interactions, 3D streams must be transmitted to a rendering system synchronously with minimal delay.

Figure 13: The projected image of the virtual teacher in our environment. The move name is displayed at the top center of the screen; the reconstructed image is in the middle of the scene.

A simple compression and encoding algorithm reduced bandwidth of data increasing the speed. In visualization phase (Figure 2c), the rendering computer combined the 3D streams from the camera clusters into a single 3D model and placed the model into a virtual environment (Figure 3). Dual projectors (Figure 1c) displayed the virtual environment on a screen in front of the physical workspace. A special pair of polarized glasses combined the overlapping images into one 3D image. Blocking a projector removed one image, which in turn removed depth cues that created the perception of three dimensions. By removing depth, we created the 2D setting for the video condition. Polarized glasses could be worn in either condition without affecting the clarity of image on the screen.

The rendering machine could also record interactions and replay them offline, allowing users to review their performance. A stereo monitor connected to a separate desktop replayed the recorded scene. The monitor simulated two overlapping images by rapidly switching between two images from different viewpoints. As with the projected scene, a special pair of glasses was worn to achieve 3D images of the replay. Image switching was turned off to achieve two dimensions for the video condition.

3.3. Experimental Materials

To create our training program we first captured a 3D model of a tai chi expert as she performed a series of moves. Moves were carefully chosen taking into consideration difficulty and the constraints of the physical workspace and the amount of training time available in a one-hour session. Moves that involved traveling more than a meter or 360-degree rotations were ruled out to accommodate the space restrictions of the workspace while the participant viewed the screen.

The recorded image of the virtual teacher contained three moves, each performed three times with a pause between the discrete moves. Moves were named as follows: Part the Wild Horses Mane, Brush Knee Twist, and Throwing the Loom. While the moves had similar lower body movements, they were quite distinguishable by the upper body movement, especially hand and arm motions.

In both conditions, the virtual world displayed the name of the current move performed by the teacher centered on the top of the screen. In the video condition, the student only saw 2D images of the environment and of the virtual teacher, and only could see the teacher from a single, front-on camera angle (Figure 4). In the virtual reality condition the student saw four stereoscopic human representations, an image of herself rendered in the third person and the teacher from behind, as well as a reflection of both those images in a virtual mirror (Figure 5). We chose a mirror as opposed to arbitrary figures duplicated in space to make the interaction as natural as possible for the participants. Furthermore, the ability to duplicate images of from different vantages is an inherent affordance of our system, and the inclusion of a mirror has been shown to be beneficial when learning physical motion [26].

In order to replay and compare the student's motions during training, the software recorded a model of the student while they were performing the moves. While rendering the scene for review, the virtual mirror was removed, and the student model was synchronized

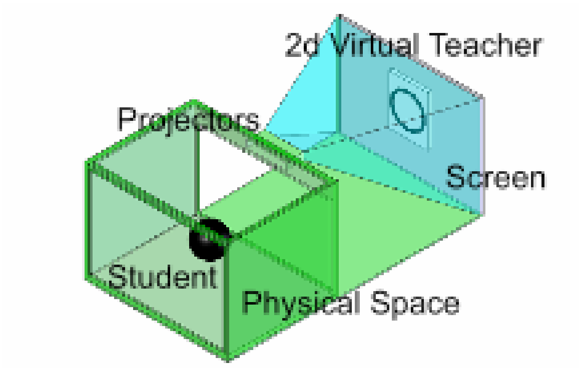


Figure 14: This shows the physical environment and virtual environment in the video condition from a three quarters perspective view. Participants in the physical environment can see a 2D image of the teacher in the screen in front of them.

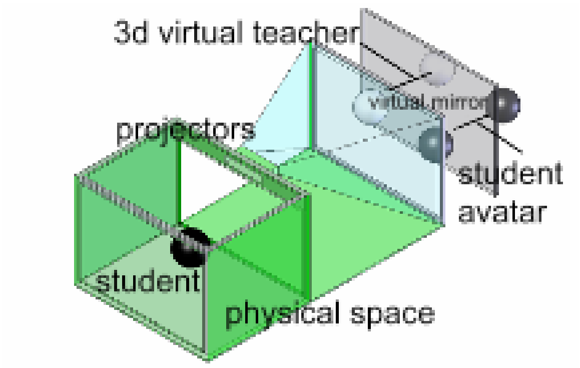


Figure 15: This shows the physical environment and virtual environment in the VR condition from a three quarters perspective view. Students in the physical environment can see the following 3D avatars: the student, the student in the mirror, the teacher, the teacher in the mirror

3.4. Participants

Twenty-six undergraduate students participated in the study and were compensated for their time. The participant pool was split evenly by condition (13 video and 13 virtual reality) and gender (13 male and 13 female). The virtual reality condition contained six male and seven female participants, while the video condition contained seven male and six female participants.

3.5. Procedures

Before entering the physical environment shown in Figure 1, participants were asked to fill out consent from and a demographic survey. After filling out these forms, participants entered the room and were given a briefing of all five phases of the task. Before they began the task, all participants were required to put on the same special clothing that assisted our vision algorithm.

During the first phase, the participants learned tai chi from the prerecorded teacher model. Participants watched and mimicked the teacher to the best of their ability. The environment varied by condition as described in the software section above. In addition to displaying the move names on the screen, the experimenter also called out the name of the move at the beginning of each trial. This phase was videotaped in order to be judged by blind coders during data analysis. In phase two the participant reviewed their performance in phase one on a separate desktop with a stereo monitor. This phase lasted for ten minutes. In the video condition, participants reviewed the actions of the video recording of the teacher in during phase one. The video recording played at the same rate as in phase one, and participants were not able to control any aspect of the recording or playback. In the virtual reality condition, participants received depth cues, control over the recording (e.g., angle, speed, distance), and the ability to examine a three-dimensional rendering of themselves as they interacted with the virtual teacher.

Phase three repeated the same learning task as in phase one. The participant learned from the teacher by watching and mimicking their motions. However, they had the benefit of added learning reinforcement from phase two [27]. Actions in this phase were also videotaped.

Phase four was the same for both conditions. The participants were tested on the individual moves and were asked to recreate the motions without the benefit of seeing the teacher. The experimenter verbally provided the participant with the name of all three moves and the participant performed the moves, one at a time, to the best of his or her ability. This phase was video recorded.

Phase five involved the participant filling out a questionnaire that contained questions about social presence and the usability of the environment. The questions from the questionnaire are listed in Appendix I.

4. Results

4.1. Blind Coder Ratings

Videos recorded during the experiment were separated by participant, phase, and move. For each of the 26 participants, there were three recorded phases with three moves each for a total of nine videos per participant or 234 possible videos in total. Of the 234 total videos, there were four missing or corrupted videos and 11 videos where the participant chose not to perform the task.

Two independent reviewers were trained to judge the specific Tai chi moves, and each evaluator inspected all remaining 219 videos and rated the participants performance. Each video was rated according to 13 separate categories that are depicted in Appendix II.

Of the 13 categories, eight described the steps of the Tai chi move and two described overall form; these were rated on a seven point Likert scale. Coders also rated the participant's knowledge of tai chi and overall performance on a five point Likert scale, and the participant's coordination as either somewhat coordinated or very coordinated. The coder inter-reliability (i.e., how similarly the two coders performed across the various videos and categories) was low but acceptable, Cronbach's α = .55. We then averaged the ratings from the two coders such that, for a given video we had only 13 scores that indicated the mean of the two coders. We then averaged the 13 scores into a single learning measure for each video (Cronbach's α = .98). Table 1 indicates the scores by condition and phase. Participants in the virtual condition consistently outperformed participants in the video condition, especially during the crucial trial of testing which was our strongest measure of actual learning. We also ran analyses to partial out the variance due to individual differences such as age, body size, gender, familiarity with technology, and previous tai chi experience. The difference between learning conditions persists when accounting for these other variables.

	Condition		Significance	
	Video	VR	*t*	*p*
Training I (Phase I)	2.10(0.69)	2.83(1.04)	2.01	0.03
Training II (Phase III)	2.23(0.71)	2.88(1.08)	1.81	0.04
Testing (Phase IV)	2.34(0.54)	3.04(1.18)	1.94	0.03

Table 1: The table above gives the mean and standard deviation values for participants ratings in each condition across all tests. The standard deviation is the value in parenthesis. The* p *value (one tailed due to our a-priori, directional prediction) and* t *scores show students learned better in the VR condition.

4.2. Self Report

We examined the subjective response questions depicted in Appendix I. We first examined all of the questions surrounding social presence of the instructor. There were ten total questions gauging social presence; we ran a factor analysis using a promax rotation to arrive at a more parsimonious representation of the items. One single factor accounted for 42% of the variance [28], and three questions loaded above .60 on this factor: To what extend did you feel in the same place as the instructor, Did you experience this task as something you did together/jointly with the instructor, and how much the instructor was rated as an expert. We took a mean score of these three questions and ran t-tests to assess the differences between video and VR on social presence ratings. As Table 2 demonstrates, participants reported significantly higher social presence in the virtual reality condition than in the video condition.

We next examined the questions that related to participants analysis of the task itself. There were seven total questions gauging subjective task performance; we ran a factor analysis using a promax rotation to isolate the responses that fell together in a single scale. One single factor accounted for 44% of the variance, and four questions loaded above .60 on this factor: Did you find this task pleasant, How easy was this task, How easy was it to move around in the environment, and ratings of how pleasant the task was. We took a mean score of these four questions, and as Table 2 demonstrates, this difference did not approach significance, and there was no significant effect of gender and no interaction between gender and learning condition. We attributed the lack of significance to the inherent bias that is shown when people evaluate their own behavior [6, 29].

	Condition		Significance	
	Video	VR	*t*	*p*
Social Presence	0.81(0.76)	1.47(1.04)	1.87	0.04
Task Presence	1.81(0.710	1.67(0.58)	0.54	0.30

Table 2: The table above shows the results of our self report questionnaire. The mean and standard deviation for each condition are given with the standard deviation in parenthesis. Ratings of task performance from participants within the virtual setting did not differ significantly from participants within the video setting. However, participants in the virtual setting felt significantly higher social presence.

Conclusions and Future Work

We presented a study assessing the affordances given by immersive virtual environments for technology mediated learning of full body physical motion. Immersive virtual environments were compared to video, the current, ubiquitous method for mediated learning. Through our comparisons, we showed that immersive virtual environments were rated better by blind coders and reported to be better by participants than their video counterparts. We showed that participants within the virtual settings performed better during every phase of the experiment. Self-reports showed participants felt higher social presence within the virtual setting.

In the current study, we designed our control condition of video only to provide a benchmark to compare our immersive system. In doing so, we may have stacked the deck in favor of demonstrating benefits of our immersive virtual reality system. There were learning features in the virtual reality condition (such as adding the mirror in the learning phases and being able to change the angle and distance in the reviewing phases) that were designed to give learners as much advantage as possible. However, these extra features are directly related to the differences between video based systems that we discussed in the introduction. Moreover, as previous work did not demonstrate reliable differences with different types of virtual tai chi instruction systems [4], we wanted to first demonstrate significant differences and then, in future work, scale the differences between virtual reality and video back in order to isolate the theoretical variables that contributed to the improved learning (e.g., stereoscopic viewing versus viewpoint control versus rendering the self in third person).

Our system captures three-dimensional models of motion in real-time, however the quality and speed of reconstruction is far from perfect. Increasing the number of cameras and infrared projectors increases coverage and redundancy leading to denser, more accurate images. Cameras that identify, focus on, and follow certain body parts could increase the accuracy in key regions (e.g., face, hands) [30, 31].

The speed of our current system is constrained by two main factors. First, extracting and combining partial models from cameras clusters is computationally expensive. Porting vision algorithms to existing graphics hardware could improve speed in this regard [32, 33]. Second, the sheer volume of data saturates the network and storage bandwidths. Using techniques that have low bandwidth consumption help ease the data management load [34, 35].

The current system reconstructs both the teacher and student models using the same real-time algorithm, however only the student model needs to be captured in real-time. Real-time performance comes at the expense of reconstruction quality; by using offline algorithms to overlay meshes on the teacher model, we can fill in holes and increase the accuracy making it easier for the students to distinguish subtle differences in the teacher's motion [36, 37, 38].

During the learning phases, participants watched and mimicked the virtual teacher. This level of interaction is acceptable for tai chi; however, other full body physical activities require feedback and the ability to manipulate the environment. Interactivity can be achieved by detecting collisions between manipulators (e.g., hands, tools) and virtual objects [39, 40, 41]. Small vibration units can give haptic feedback, indicating successful interaction with a virtual object [42, 43]. For example, a participant could dance with a virtual partner and know that their avatars are making contact; or a participant could swing a golf club to hit a virtual ball. Increased interactivity can also be achieved by increasing the intelligence of the virtual agent. The current system uses a pre-recorded model as the teacher. While human level artificial intelligence is far beyond our capabilities, it can be approximated by creating simple reactive avatars or simulated by remotely controlling the teacher model. It has been shown that the simple actions such as mimicking non-verbal gestures are enough to give the perception of intelligence [44]. As such, encoding these actions within the teacher model can increase the feeling of immersion. A simpler approach involves connecting a remote teacher to student by networking two immersive reality systems [5]. The remote teacher can interact though natural motion without having to be collocated with the student. Comparative studies can assess the benefits of each approach.

Exciting work has been done regarding collaboration within virtual environments [45, 46], however space constraints of the current physical environment limit both the number of participants in the virtual environment and the amount any given participant can move. Since each participant retains their full mobility in the virtual world, enabling our physical environment to support multiple people would expand our learning domain to include collaborative tasks (e.g., disaster relief, construction, and surgery). Retaining natural movement bolsters the validity of statements about virtual environments and the effect on social interaction and learning.

In the current system, the participant must look at a screen in front of them to see the virtual world. Exploring other immersive technologies such as HMDs or full CAVE [47] environments may increase mobility and lead to a higher degree of immersion and better learning. Different technologies are better suited to different tasks [48], studying the interaction between technology and task can help build better systems.

Changes in technology can be assessed to discover their effects on learning physical motion. As interactivity and realism of our environment increases, we hope to bridge the gap between learning from a virtual teacher, in a virtual environment, to learning from face-to-face interaction with a real teacher. In the current work, we have demonstrated persuasive evidence that immersive virtual reality provides better learning of physical movements than a two-dimensional video. As technology, and our understanding of how to use that technology, improves we should see larger gains in learning from virtual reality.

Acknowledgements

We would like to thank the Stanford LIFE lab and Nick Yee for helpful suggestions on the design of this research. We thank Nancy Ru for providing us with our teacher model, Alice Kim, Zi Han, and Claire Carlson for encoding and inputting data, and Jaireh Teccaro for creating figures.

References

[1] F Mantovani, *Virtual Reality Learning: Potential and Challenges for the Use of 3D Environments in Education and Training*, chapter 12, Towards Cyber-Psychology: Mind, Cognitions and Society in the Internet Age. IOS Press, Amsterdam, 2001.

[2] J Bruner, *Towards a theory of instruction*, WW Norton, New York, 1966.

[3] E Trondsen and K Vickery, "Learning on demand," *Journal of Knowledge Management*, vol. 1, no. 3, pp. 169–180, March 1997.

[4] P T Chua, R Crivella, B Daly, N Hu, R Schaaf, D Ventura, T Camill, J Hodgins, and R Pausch, "Training for physical tasks in virtual reality environments: Tai chi," in *Proc of IEEE Virtual Reality*. IEEE, 2003, pp. 87–94.

[5] Z Yang, K Nahrstedt, Y Cui, B Yu, J Liang, Sang Hack Jung, and Ruzena Bajscy, "Teeve: The next generation architecture for tele-immersive environments," in *Proc IEEE International Symposium on Multimedia*, Irvine, CA, USA, 2005, IEEE.

[6] J N Bailenson, K R Swinth, C L Hoyt, S Persky, A Dimov, and J Blascovich, "The independent and interactive effects of embodied agent appearance and behaviour on self-report, cognitive, and behavioural markers of copresence in immersive virtual environments.," *PRESENCE: Teleoperators and Virtual Environments*, vol. 14, pp. 379–393, 2005.

[7] R D Pea, Video-as-data and digital video manipulation techniques for transforming learning sciences research, education and other cultural practices, *International Handbook of Virtual Learning Environments*. Kluwer Academic Publishing, 2005.

[8] Daniel L Schwartz and K Hartman*, It is not television anymore: Designing digital video for learning and assessment*, Video research in the learning sciences. Erlbaum, Mahwah, NJ, 2006.

[9] B J McNeal and K R Nelson, "Meta-analysis of interactive video instruction: A 10 year review of achievement effects," *J. COMP. BASED INSTR*, vol. 18, no. 1, pp. 1–6, 1991.

[10] J M Loomis, "Distal attribution and presence," *PRESENCE: Teleoperators and Virtual Environments*, no. 1, pp. 113–119, 1992.

[11] M Slater, "Presence and the sixth sense," *PRESENCE: Teleoperators and Virtual Environments*, vol. 11, no. 4, pp. 435–439, 2002.

[12] J Blascovich, J Loomis, A Beall, K Swinth, C Hoyt, and J Bailenson, "Immersive virtual environment technology as a methodological tool for social psychology," *Psychological Inquiry*, vol. 13, pp. 103–124, 2002.

[13] M Alavi, B C Wheeler J S, and Valacich, "Using it to reengineer business education: An exploratory investigation of collaborative telelearning," *MIS Quarterly*, vol. 19, no. 3, pp. 293–313, 1995.

[14] G Piccoli, R Ahmad, and B Ives, "Web-based virtual learning environments: A research framework and a preliminary assessment of effectiveness in basic it skills training," MIS *Quarterly*, vol. 25, no. 4, pp. 401–426, 2001.

[15] M-N Lamy and R Goodfellow, "Reflective conversation," *Virtual Language Classroom. Language Learning and Technology*, vol. 2, no. 2, pp. 43–61, 1999.

[16] M Warschauer, *Interaction, Negotiation and Computer-mediated Learning*, Educational Technology in Language Learning: Theoretical Considerations and Practical Applications. 1998.

[17] R Lofin and P Kenny, "Training the hubble space telescope flight team," *IEEE Computer Graphics and Applications*, vol. 15, pp. 31–37, 1995.

[18] Albert Rizzo, Jacquelyn Ford Morie, Josh Williams, Jarrell Pair, and J Galen Buckwalter, "Human emotional state and its relevance to military VR training," in *International Conference on Human Computer Interaction*, 2004.

[19] J Berkley, G Turkiyyah, D Berg, M Ganter, and S Weghorst, "Real-time finite element modelling for surgery simulation: An application to virtual suturing," *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, no. 3, pp. 1–12, 2004.

[20] S Delp, P Loan, C Basdogan, and J Rosen, "Surgical simulation: An emerging technology for training in emergency medicine," *PRESENCE: Teleoperators and Virtual Environments*, vol. 6, pp. 147–159, 1997.

[21] S Stansfield, D Shawver, and A Sobel, "Medisim: A prototype virtual reality system for training medical first responders," in *Proc of IEEE Virtual Reality*. IEEE, 1998.

[22] D Tate, L Silbert, and T King, "Virtual environments for shipboard firefighting training," in *Proc IEEE Virtual Reality Conference*. IEEE, 1997.

[23] U Yang, "Just follow me: An immersive virtual reality based motion training system," in *International Conference on Virtual Systems and Multimedia*, 1999.

[24] K M LEE, "Presence, explicated," *Communication Theory*, vol. 14, pp. 27–50, 2004.

[25] K Daniilidis, J Mulligan, R McKendall, G Kamberova, D Schmid, and R Bajcsy, *Real-Time 3D Tele-immersion*, pp. 253–266, The Confluence of Vision and Graphics. Kluwer Academic, 2000.

[26] L P Sewall, T G Reeve, and R A Day, "Effect of concurrent visual feedback on acquisition of a weightlifting skill," *Percept Mot Skills*, vol. 67, no. 3, pp. 715–718, December 1988.

[27] D L Weeks and L P Anderson, "The interaction of observational learning with overt practice: effects on motor skill learning," *Acta Psychol*, vol. 104, no. 2, pp. 259–271, May 2000.

[28] A K Romney, S Weller, and W Batchelder, "Culture as consensus: a theory of culture and informant accuracy," *American Anthropologist*, vol. 88, pp. 313–338, 1986.

[29] [29] R E Nisbett and T D Wilson, "Telling more than we can know: Verbal reports on mental processes," *Psychological Review,* vol. 84, pp. 231–259, 1977.

[30] Henry Schneiderman, "Learning a restricted bayesian network for object detection," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, June 2004.

[31] Ying Wu, Gang Hua, and Ting Yu, "Tracking articulated body by dynamic markov network," in *IEEE International Conference on Computer Vision*, Nice, France, October 2003, IEEE, pp. 1094–1101, IEEE Computer Society.

[32] James Fung and Steve Mann, "Openvidia: parallel gpu computer vision," in *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia,* New York, NY, USA, 2005, pp. 849–852, ACM Press.

[33] William R. Mark, R. Steven Glanville, Kurt Akeley, and Mark J. Kilgard, "Cg: a system for programming graphics hardware in a c-like language," *ACM Transactions on Graphics (SIGGRAPH 2003)*, vol. 22, pp. 896–907, 2003.

[34] Marc Levoy and Pat Hanrahan, "Light field rendering," in *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1996, pp. 31–42, ACM Press.

[35] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen, "The lumigraph," in *SIGGRAPH*

'96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, New York, NY, USA, 1996, pp. 43–54, ACM Press.

[36] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis, "SCAPE: Shape completion and animation of people," in *Proceedings of the 32nd International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH),* Los Angeles, California, August 2005.

[37] Diego Nehab, Szymon Rusinkiewicz, James Davis, and Ravi Ramamoorthi, "Efficiently combining positions and normals for precise 3d geometry," *ACM Transactions on Graphics (SIGGRAPH 2005)*, vol. 24, no. 3, Aug 2005.

[38] Greg Turk and Marc Levoy, "Zippered polygon meshes from range images," in *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1994, pp. 311–318, ACM Press.

[39] Ming C. Lin, "Fast and accurate collision detection for virtual environments," in *Scientific Visualization*, Washington, DC, USA, 1999, IEEE, pp. 171–180.

[40] Naga K. Govindaraju, David Knott, Nitin Jain, Ilknur Kabul, Rasmus Tamstorf, Russell Gayle, Ming C. Lin, and Dinesh Manocha, "Interactive collision detection between deformable models using chromatic decomposition," *ACM Transactions on Graphics (SIGGRAPH 2005)*, vol. 24, no. 3, pp. 991–999, 2005.

[41] Jan Klein and Gabriel Zachmann, "Point cloud collision detection," in *Computer Graphics forum (Proc. EUROGRAPHICS)*, M.P. Cani and M. Slater, Eds., Grenoble, France, Aug 30 – Sep 3 2004, number 3 in 23, pp. 567–576.

[42] Karon E. MacLean, "Designing with haptic feedback," in *Proc of International Conference on Robotics and Automation*, San Francisco, CA, USA, April 2000, IEEE.

[43] Allison M. Okamura, Stanford Dennerlein, and Robert D. Howe, "Vibration feedback models for virtual environments," in *Proc of International Conference on Robotics and Automation*. IEEE, 1998, pp. 674–679.

[44] Jeremy N Bailenson and Nick Yee, "Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments," *Psychological Science*, vol. 16, pp. 814–819, 2005.

[45] Veronique Normand, Christian Babski, Steve Benford, Adrian Bullock, Stephane Carion, Yiorgos Chrysanthou, Nicolas Farcet, John Harvey, Nico Kuijpers, Nadia Magnenat-Thalmann, Soraia Raupp Musse, Tom Rodden, Mel Slater, and Gareth Smith, "The coven project: Exploring applicative, technical, and usage dimensions of collaborative virtual environments," *PRESENCE: Teleoperators and Virtual Environments*, vol. 8, no. 2, pp. 218–236, 1999.

[46] Jaron Lanier, "Virtually there," *Scientific American*, pp. 66–75, April 2001.

[47] C Cruz-Neira, D J Sandin, and T A DeFanti, "Surround-screen projection-based virtual reality: The design and implementation of the cave," in *Proc of SIGGRAPH-93: Computer Graphics*, Anaheim, CA, 1993, pp. 135–142.

[48] Pantelidis, Veronica S. Virtual Reality in the Classroom. *Educational Technology* vol. 33 no. 4 (April 1993), p. 23-27. (General.) Reprinted in John Hirschbuhl, ed., *Computer Studies: Computers in Education*, p.175-179. Guilford, Connecticut: The Dushkin Publishing Group, Inc., 1993.

Appendix

I. Self Report Questions

The following questions were rated on a five point Likhert scale. The extremes of each scale are in brackets.

- To what extent did you have the sense that you were in the same place as the instructor? [from a very small extent to a very large extent]
- How would you rate your awareness of the instructor's intentions/wishes in this task? [from low awareness to high awareness]
- Did you find this task pleasant or unpleasant? [from very pleasant to very unpleasant]
- Did you experience this task as something that you did together/jointly with the instructor, or as something you did on your own/separately)? [from a very large extent on my own to every large extent together]
- How easy or difficult was this task? [from very difficult to very easy]
- How easy or difficult was it to move around in the environment? [from difficult to very easy]The following questions were rated on a five point Likhert scale. The extremes of each scale are in brackets.
- To what extent did you have the sense that you were in the same place as the instructor? [from a very small extent to a very large extent]
- How would you rate your awareness of the instructor's intentions/wishes in this task? [from low awareness to high awareness]
- Did you find this task pleasant or unpleasant? [from very pleasant to very unpleasant]
- Did you experience this task as something that you did together/jointly with the instructor, or as something you did on your own/separately)? [from a very large extent on my own to every large extent together]
- How easy or difficult was this task? [from very difficult to very easy]
- How easy or difficult was it to move around in the environment? [from difficult to very easy]
- For each of the word pairs below, please circle the number that best suits your experience of the learning environment. [from personal to impersonal, from social to antisocial, from lively to lifeless, and from pleasant to unpleasant]
- For each of the word pairs below, please circle the number that best suits your experience of the instructor during the task that you have just done. [from close to distant, from responsive to unresponsive, from active to passive, from warm to cold, from helpful to unhelpful, from realistic to fake, and from an expert to a novice]

II. Blind Coder Rating Criteria

The following questions were rated on a seven point Likhert scale from very poor to very well. Coders were given guidelines on how to evaluate each step for each move and how to evaluate fluidity and posture.

- You are provided with descriptions of how to perform each move broken down into distinct steps. On the scale below rate the participant's performance on each step of the move.
- To correctly perform tai chi each move part must flow smoothly into the next. In addition, the participant must have good posture. Good technique is described in the attached handout. On the scale below rate the fluidity of the moves and the participant's posture throughout the entire move.

The following questions were rated on a five point Likhert scale from poor to excellent.

- Please rate the participant's knowledge of tai chi
- Please rate the participant's overall tai chi performance

For the final question the coders rated the participants overall coordination as either somewhat coordinated or very coordinated.

Examining the relationship between violent video games, presence, and aggression.

Kristine L. Nowak, Marina Krcmar, Kirstie M. Farrar
University of Connecticut
{kristine.nowak, marina.krcmar, kirstie.farrar}@uconn.edu

Abstract

The level of presence, or immersion, a person feels with media influences the effect media has on them. This project examines both the causes and consequences of presence in the context of violent video game play. In a between subjects design, 227 participants were randomly assigned to play either a violent or a non violent video game. Causal modeling techniques revealed two separate paths to presence. First, individual differences predicted levels of presence: men felt more presence while playing the video game, as did those who play video games more frequently. Secondly, those who perceived the game to be more violent felt more presence. Those who felt more presence, felt more resentment, were more verbally aggressive, and that led to increased physically aggressive intentions.

Keywords--- Presence as immersion, video games, aggressive affect, violence, aggression, and social learning theory.

1. Introduction

The sense of presence, or the sense of involvement with and engagement in media, influences how people respond to mediated stimuli, and how it affects them [1-3]. As with social learning from media and media effects generally, the sense of presence varies across contexts, content, and media, as well as from person to person and from one exposure to the next. Both the level of presence with the media [4-6] and aggressive responses to mediated stimuli [7, 8] have been shown to be influenced by individual differences (e.g., gender, media use), and features of the medium (e.g., screen size, interactivity, vividness, agency). Therefore, understanding the factors that influence the sense of presence and it's potential influence on aggression may provide insight into the process of media effects and when and why some people respond differently to the same mediated stimuli, whether violent or not, than others.

Being able too predict when, and to what extent, media will influence people is of critical importance particularly when considering media violence. In this study, we examine how individual difference variables (e.g., gender, video game use) influence people's responses to an experimental manipulation (i.e. violent game play) in terms of their sense of presence and subsequent level of aggression.

After 4 decades of research, scholars have concluded that exposure to media depictions of violence can cause aggressive behavior including the imitation of violent acts [9-11]. There is also evidence of desensitization, with those who see more violence having a greater acceptance of, and tolerance for, violent behavior [12, 13]. As Anderson [14] argued, the "scientific debate over whether media violence has an effect is over" (p.114). Research using experimental, longitudinal and cross sectional methods has concluded that exposure to violent television, film, and video games causes increased aggressive behaviors and attitudes and decreased prosocial behaviors among adults, children, men and women [14, 15]. In most, although not all cases, researchers have found stronger effects following exposure to video game violence as compared to exposure to television violence [16].

More importantly, this effect seems to be getting stronger. Gentle, et al [15] argued that video game research that is 10-15 years old likely underestimates the effect of current video games on players. Video games have become more realistic, engaging, and increasingly violent. They use more vivid and sophisticated graphics, including vividly depicted violence against human characters. They are also more engaging--requiring active participation in games made possible by interactivity and increasingly involving input devices such as head mounted displays and data gloves. These changes to video games have led to increased concern over the effects of exposure to, and interaction with, violence video games. In addition, research has shown that these more current, technologically advanced violent games have led to increased levels of aggression from video game players when compared to the much tamer games of a decade or more ago [17].

Despite what is known about the effects of violent television, and to a lesser extent, the effects of violent video games, there are still unanswered questions regarding how different kinds of people respond to games and how those responses might influences people's perceptions of acceptable behavior, or social learning, and how this might in turn affect outcomes. Exploring the processes that predict when, how, and to what extent, exposure to violent video games will influence both people's perceptions of acceptable behavior, their sense of presence, and their overall levels of aggression is of critical importance not only for researchers, but also for policy makers and society in general [18].

This article reports the results of an experiment that examines both the causes and consequences of presence following people's exposure to violent video games, their perception of the violence and how presence is related to aggressive affect. We first measure individual difference variables (gender, previous video game use), then manipulate the presence or absence of violence experienced while playing video games. Finally, we use causal modeling techniques to show the extent to which variations in degree of presence causes variations in levels of aggression. We specifically examine whether a person's level of presence can predict levels of aggressive outcomes including resentment, hostility, and verbal aggression and physically aggressive intentions.

2. Learning and modeling from the media: Social Cognitive Theory

Social cognitive theory (SCT) [19] grew out of Bandura's earliest work on social learning theory [20]. Whereas social learning theory proposed that learning can occur through modeling and imitation, more recent theoretical work on social cognitive theory [21] "accords a central role to cognitive, vicarious, self-regulatory and self-reflective processes….Most external influences affect behavior through cognitive processes rather than directly. Cognitive factors partly determine which environmental events will be observed, what meaning will be conferred on them, whether they leave a lasting effect, what emotional impact and motivating power they will have, and how the information they convey will be organized for future use, " (p. 122). In other words, learning can occur through imitation of direct and mediated modeled behavior [10]; however, the cognitive experience of the observer has an effect on any outcomes that will result.

In the case of video game violence, various aspects of the theory are more specifically relevant. For example, early social learning theory argued that behaviors that were rewarded were more likely to be imitated by an observer than behaviors that were punished or went unrewarded [22]. More recent advances in the theory would argue that it is the cognitive interpretation and experience of that witnessed event that would lead people to act in ways that create desirable outcomes, avoid negative outcomes, and utilize information about the consequences of others' actions in making their decisions about how to act, and what behaviors to imitate [23]. Therefore, both the rewarded event being witnessed, and the cognitive experience of that event by the observer might impact likelihood of imitation.

In support of SCT, early research on television violence has found that when media depict a character being rewarded, or even not punished, for aggressive behaviors, viewers are more likely to imitate the behaviors [24]. This is particularly relevant in the case of video games, where aggressive behaviors are often required, and players are rewarded for aggressive behavior. The interactive nature of video games means that these aggressive actions influence the outcome of the game, as well as what the player sees. By acting aggressively, players earn points, move up to the next level and may ultimately win the game. By rewarding the aggressive and violent actions of video game players, it may promote the perception that violence is useful, appropriate and even a good way of dealing with conflict [18]. Furthermore, the interactivity of video games requires players to engage in the game to attain the skills required to read the output devices and to quickly react and manipulate the input devices provided with the interface [25]. In accord with social cognitive theory, this level of interactivity and engagement is the cognitive experience of the game that may explain why exposure to video game violence results in higher levels of aggression than exposure to television violence [16].

Essentially, both social cognitive theory, and the extant literature on mediated violence, suggests that observing rewarded behavior makes that behavior seem more attractive and more likely to be imitated. The interactive nature of video games requires active participation and a higher level of presence, or involvement [6]. This active participation and increased presence may make it even more likely that people will imitate and repeat the behaviors they learned and practiced when playing the game. In the case of violent video games, this could increase aggression.

3. Hostility

The majority of research to date has examined the effect of media violence on verbal and physical aggression. However, outcomes such as aggressive affect or hostility have also been explored. For example, violent video game play can influence aggressive cognitions or thoughts and has been found to increase hostility [6, 26, 27].. Meta-analyses conducted on the research on violent video games have also supported an effect of game play on levels of aggressive cognitions and hostility [16, 17]. Therefore, aggressive cognitions and hostility are affected by exposure to media violence.

Furthermore, affective hostility mediates the relationship between exposure to violent video games and aggressiveness [15]. Therefore, it is important to include measures of hostility in a complete model that explores the relationship between video game play, aggression, and, as we explore here, presence.

4. Presence and the Suspension of Disbelief

The sense of presence may be thought of as the level of involvement with the medium. Although the concept of presence is multidimensional, we use the term here as associated with immersion, or a sense of being 'in' the medium. This dimension of presence is also sometimes called telepresence [1-3, 28]. The International Society for Presence Research [29] defines presence, which it calls a shortened version of the term 'telepresence,' as "a psychological state or subjective perception in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience." If, as some researchers have argued [2], presence can only occur when it is brought about by technology, than 'tele' as in telepresence becomes unnecessary [30]. Therefore, we use the term presence.

To experience true presence then, one must suspend awareness that the source of stimuli is mediated, which is often called a willing suspension of disbelief. This suspension of disbelief can be conscious or not. This implies that a person is engaging their senses (hearing, sight) to receive the stimuli coming from a mediated environment, (Virtual reality, television, radio, video game, etc). This focus on the mediated stimuli comes at the expense of information, or stimuli, coming from natural, or unmediated sources (honking cars, people in the room, etc).

Although presence is traditionally associated with research in virtual reality [3, 31]), a theory of presence is applicable to all media [1-3]. For example, even television related variables such as screen size [2, 4, 5] and a number of other factors including level of interactivity, and avatar anthropomorphism [1, 32, 33] have been shown to influence people's sense of presence.

Presence or a sense of being "in" the mediated environment is also likely to increase when one is familiar with the medium. In the case of video games, it is clear that presence is enhanced once the initial frustration of learning the game is passed and a player can fully engage in the challenge of the game itself [6]. Therefore, increased play will likely lead to decreased frustration which in turn will lead to increased presence. Finally, some individuals are more prone to feeling present with media than others [1]. For example, interest in the message or environment has been shown to increase a sense of presence and people's willing suspension of disbelief [34]. Also, males and females have been shown to experience presence differently [35]. Therefore, these variables are examined in the present study.

5. Considering the relationship between presence and social learning.

Just as individuals differ in their hostile or aggressive reactions to media, there are individual differences in players' feelings of presence. The two factors known to be related to individual differences in presence are gender, and experience with the medium, though these two factors are highly correlated.

On average, boys play video games for greater duration and with greater frequency than their female counterparts [36]. This increased frequency represents a fairly large investment of time in the game, which should increase skill with the game over time. As with any learned activity, practice improves skill, which increases enjoyment, and acts as its own reinforcement.

Perhaps in part due to this greater experience and enjoyment, men play more games and also experience a greater sense of presence following exposure to violent games. Wilfred's [37] examination of the effectiveness of VR based simulation training revealed that males felt more presence than women following a game based interaction in a virtual environment, although he did not control for previous experience with VR technology.

Previous research has shown that in addition to playing games less frequently, and feeling less presence, women tended to perceive depictions of violence as more severe than men who watched the same depictions [38]. Further, men were more aggressive after playing a violent video game than women [39], though men are more aggressive than women, in general [40]. Overall, then, it would seem that men are more likely than women to enjoy the game, and to experience a greater sense of presence when playing a violent game, as well as afterwards. Social cognitive theory would argue that the cognitive experience of the game by the user would influence any emotional and behavioral outcomes of game play. For example, because males play more often and have more experience in game play, not only is play less frustrating, but their greater experience allows them an opportunity for greater immersion or presence in the game.

However, gaming experience and presence are not the only predictions of, or explanations for, the level of a person's aggressive responses. Players' interpretations of violent stimuli may be equally important. For example, individual interpretations of violent stimulus have been found to affect aggressive outcomes [41] and men are likely to perceive the depicted violence as less severe than women [38]. In addition, because game play is largely a cognitive experience where players' experiences occur between the screen and the mind, it is important to explore how the cognitive experience of this event affects outcomes. Here, social cognitive theory would argue that *perceptions* of violence and *perceptions* of presence in the game are likely to result in greater hostility and greater aggression.

It is possible the perception that those who perceive violence as less severe are more likely to accept, adapt, or just believe that behaviors that others would classify as aggressive were more normative. If this is true, then social learning would predict that they would behave more aggressively after exposure to violence.

Essentially, as discussed above, various factors are likely to affect game play outcomes. Gender is likely to affect frequency of game play in general, but also to affect feelings of presence during game play. In part due to this greater frequency of game play, males are also less likely to feel frustration during play.

Here we examine the extent to which presence may moderate the relationship between perceptions of violence and aggressive affect. If we extend the predictions of SCT to this context, then we could predict that combining the vicarious experience of playing games with the cognitive perception that the game is violent would result in a more intense and involving, game experience, or increased presence. Therefore, we predict that participants' perception that the video game is violent will lead to increased presence, and that the sense of presence will increase aggression. That is, we predict that presence, and the *experience* of game play cause differences in aggression.

Overall, there has been little empirical testing of the effect of increased presence on learning or outcome variables, such as aggression. As Wilfred, [37] pointed out, "There is much more focus on the technology of virtual reality than there is evaluation of its impact on learning. The efficacy of the VR systems developed, and what factors mediate this effectiveness are seldom studied." The same is true of presence research, with a few exceptions. In this project, we examine some of the predictions outlined above with a causal model of the relationship between gender, game use, presence and aggression to shed further light on the relationship between these variables.

6. Methodology

This study uses an experimental design and causal modeling techniques to examine the causes and consequences of presence in the context of violent video games. Participants were randomly assigned to play either a violent or non-violent game before responding to a number of questionnaire items.

6.1. Participants

Participants in this study were 227 undergraduate students (109 males, 117 females, one subject did not report gender) enrolled in lower division Communication courses at a large East Coast University. Participants received extra credit for taking part in this research.

6.2. Stimulus Materials

For the purposes of this research, violence is defined using the operational definition from the National Television Violence Study (NTVS): "any overt depiction of a credible threat of physical force or the actual use of such force intended to physically harm an animate being or group of beings" [42], p. 30). So, in order to qualify as a violent game for this research, the main character in the game had to demonstrate actual intent to physically harm others. The violent game used for this study is Hitman II, Silent Assassin. This game received an Entertainment Software Rating Board (ESRB) rating of "M" for Mature. The nonviolent control game is Tony Hawk, Pro Skater3. Tony Hawk received an ESRB rating of "T" for teen.

6.3. Measurement Instruments

All scales were tested for internal consistency and parallelism through confirmatory factor analysis using the software CFA (Hamilton & Hunter, 1997).

Demographic variables. Subjects indicated their gender, age, year in school and race.

Game Use was measured with a 5 item likert-type scale on a 7 point metric (Standardized Alpha = .84). Participants were asked to indicate how frequently they play different types of video games.

Perceived Violence was measured with a 3 item likert type scale with a 7 point metric (Standardized Alpha = .85). Participants were asked to think about the game they played and rate it in terms of violent content.

Frustration with the game was measured with a 2 item scale (easy/difficult, or frustrating/not frustrating) with a 7 point metric (Standardized Alpha = .77).

Presence as immersion, or the extent to which participants felt that they were "inside" the video game [43, 44], was measured with a 5 item likert type scale with a 7 point metric (Standardized Alpha = .89).

Hostility was measured using 8 items [45]. These were likert type items on a 4 point metric (Standardized alpha = .92).

Aggression. Finally, aggressiveness was measured by using a modified version of the Buss-Perry version of the aggression questionnaire [46]. The items in this study were slightly reworded to reflect state rather than trait aggression. Before responding to these items, participants were asked to: "Imagine that you leave this building when you're done completing this survey. Someone bumps into you, spilling your drink and the contents of your backpack." They were then asked to rate whether each potential reaction was "0" (extremely uncharacteristic of me) to "6" (extremely characteristic of me). This reworded version of the Buss Perry aggression scale has been found to be reliable, and to more accurately tap participants' responses to an aggressive prime [47]. Confirmatory factor analysis tests revealed three separate dimensions on this scale, including resentment, verbal aggression and physically aggressive intentions. Items and the construct they measured are detailed below.

Resentment was measured with 5 aggression items (Standardized alpha = .88). These items included 'this person always seems to get the breaks,' 'I think this person talks about me behind my back,' 'I would be suspicious of this person being overly friendly.'

Verbal Aggression was measured with 5 aggression items (standard alpha = .89). They included 'I would tell this person openly that I disagree with him or her,' and 'this person would say that I'm somewhat argumentative.'

Physically Aggressive Intentions was measured with 5 aggression items (standard alpha = .88). They included 'given enough provocation, I would hit this person,' and 'if this person hit me, I would hit back.'

6.4. Procedure

Male and female undergraduates were randomly assigned to play either the nonviolent control game or the violent game for 12 minutes. All games were played on a Sony PlayStation II gaming console hooked up to a 13-inch color television monitor. Participants filled out a post-test immediately after playing.

7.0. Results

To test the relationship between these variables, a causal modeling technique was employed using Path Version 5.0. The model shown in Figure 1 contains only significant structural coefficients at $p < .01$, with no significant missing paths indicating that all variables with direct relationships have paths in this model. The overall goodness-of-fit of this model, as measured by the Root Mean Square Error of Approximation (RMSEA), is .05, matching the desirable value of .05 or less and far from the unacceptable value of .10 or greater. The overall model chi-square is 11.83 with 24 d.f., which does not differ significantly from the original data, $p = .98$. This is well above the desirable significance level of .05 or greater. Finally, no reproduced errors were above .10. On balance, the structure of the final model appears to be a good fit with the observed data. This section will first discuss the separate paths to presence and then discuss the impact of presence on aggression.

This model shows two separate paths to presence: one through individual differences (biological sex and previous game use), and the other through condition (violent game or non-violent game). Both of these variables have paths directly to presence and also to level of frustration. There is a large path showing that perceived violence increases presence, but also a moderate negative path from game type to presence, indicating that those who played a violent game felt less presence. Perceived violence increased frustration, and men felt less frustration than women. Those who were more frustrated felt less presence. The model also shows that presence increases hostility and aggressive affect.

The first path to presence is from individual differences including gender, and previous game use, both of which also predicted the level of frustration. Also, biological sex is a very strong predictor of previous game use. Males play significantly more video games than females, and those who play more games felt more present. However, there was no link from gender to perceived violence and no difference between men and women in terms of how violent they perceived the game, though men felt less presence than women overall.

However, two somewhat contradictory factors are influencing presence. First, the direct path from biological sex to presence shows that, all things being equal, females felt more presence than males. However, females, and people who do not use video games a lot, felt the game was more frustrating, and frustration reduced presence. The path showing that frustration decreased presence was larger than the path from biological sex to presence, revealing that frustration (and game use) has a stronger effect on presence than biological sex, though again, both of these factors are predicted by biological sex.

There was also a direct path from biological sex to both verbal and physically aggressive intentions, with males demonstrating more verbal and physically aggressive intentions than women. This was true regardless of previous game use, level of frustration, or which game they played (violent or nonviolent). People who frequently played video games may have been more aggressive from the beginning, but there is not a direct path from game use to aggression; only from game use to frustration and presence.

The second path to presence comes from both condition (violent or nonviolent game) and the level of perceived violence, and there appears to be contradictory influences on presence from this direction as well. Those randomly assigned to play a violent game perceived that game to be more violent than those assigned to the control game. This perception of violence strongly increased presence, with moderate to small influences on resentment, and frustration with the game, though it decreased people's hostility, regardless of gender or previous game experience. However, there is a direct path from condition to presence, which indicates that those who played a violent game felt less presence when condition was not moderated by perceived violence.

Those who play more games in general were less frustrated and felt more presence, and this in turn resulted in greater verbal and physically aggressive intentions and hostility. Thus, previous game use did not predict verbal aggression directly; however, the relationship between these variables is mediated by presence.

There were only two direct links from presence: hostility and verbal aggression. Presence directly predicts hostility, which predicts resentment, which predicts verbal aggression and physically aggressive intentions. Also, presence directly increases verbal aggression, which increases physical aggression. There were direct links to physical aggression from violent game, biological sex, and verbal aggression.

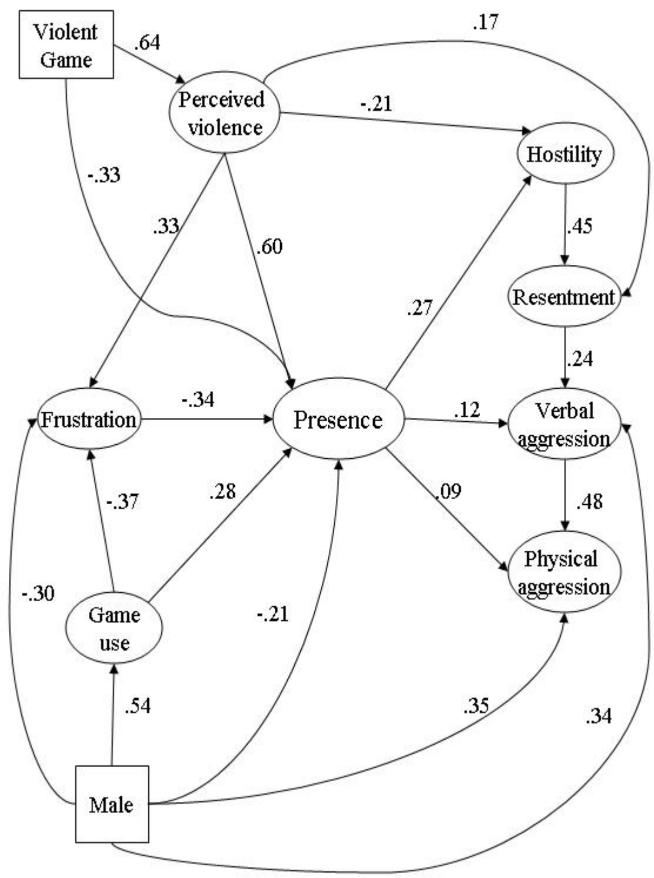


Figure 16 Path Model

This sense of presence and reduced frustration among males seems to occur both because of more frequent game use and independent of it. In the former case, it is reasonable that experience allows one to experience more presence with the medium, become more involved in the story line and less frustrated by the experience of game play. People who were not frequent game players may have had to allocate more cognitive resources to figuring out the rules of the game, or to learning how to work the interface. This may well have increased frustration, and consistently reminded the participant that the experience was mediated, which would reduce presence. In the latter case, it may be that males, socialized to enjoy more aggressive story lines, can become more involved even without the benefit of previous exposure to games. In this case, the different reactions to media stimuli would depend upon interest in, or comfort with, the content. However, we note that the difference in level of frustration held was predicted by the perception of violence and not by whether participants played a violent or non-violent game.

It is important to recognize that the previous game use measure in this study did not specifically measure violent video game usage. This may explain why there is not a separate path directly from game usage to aggression. Future research should measure violent video game play separately to test whether those who play more violent video games are more aggressive than those who play other video games following a short exposure to violent stimuli. This would allow for the distinction between familiarity with the medium, and aggressive priming, as the cause of increased presence and aggression.

Also, this limited test of video games and aggression seems to suggest that both overall game use and the experimental manipulation of game play have an influence on the aggression measures. However, as with the research by Gentile, et al [15], in this study there was no direct link between game use and hostility. Instead, presence mediated the relationship between game use, and hostility and aggressive outcome variables. In addition, frustration mediated the relationship between games use and aggression, suggesting that both the cognitive experience of presence and the affective experience of frustration and presence are the processes by which aggression is increased.

In the case of the game manipulation, there was a direct impact on presence, with those playing a violent game feeling less presence. However, the very strong effect of perceived violence on presence reveals that those who perceived the game as more violent felt a greater sense of presence, and they felt more resentment directly associated with play. However, those that perceived more violence felt less hostility. In line with our earlier argument, those socialized to perceive violence might also become more immersed in a violent story line.

Finally, presence, both that associated with game use and that associated with the experimental manipulation, resulted in greater verbal aggression and more physically aggressive intentions. These findings are consistent with social cognitive theory [21] in that the symbolic experience of game play, in part as indicated by presence, affects outcomes. Note that there are no direct paths from game use, or the experimental manipulation, to verbal or physical aggression (though there are small paths from perceived violence to resentment and hostility). As Bandura would argue: "An extraordinary capacity for symbolization provides humans with a powerful tool for comprehending their environment and creating and regulating environmental events that touch virtually every aspect of their lives ([21], p. 122)."

Despite support for social cognitive theory, there are some anomalous findings in the model. Specifically, those who perceived greater violence in the game experienced less hostility. Perhaps when players perceived the game as violent, but did not experience greater presence, they did not engage as much with the content and played the game as a way to pass time. So, it is not that exposure to violence reduced hostility but that without presence, hostility was not increased as much. Essentially, the lower level of presence may have resulted in lower levels of affective hostility. This argument is supported in part because when presence did not mediate the link between game play and hostility, the path was negative. In other words, aggression results from true involvement in the violent game (i.e., presence) and not from game play as a means to pass time. This argument is consistent with a uses and gratifications approach to media [48] that argues that the way media are used influences resulting outcomes.

Understanding the causes and consequences of presence may be central to our understanding of who will be most affected by media violence and under what conditions. Since we know that increased presence increases aggressive affect, we can look to things that increase presence (both individual differences and features of the media) for cues about how to predict the influence of violent media on people.

The results from this study, as well as previous studies, show significant effects on hostility and aggression as a result of playing a game for very small amounts of time- particularly when they are present, or engaged. Men and women did not differ in terms of how violent they perceived the game, and those who played the violent game were more aggressive than those who played the control game. This raises the question of the effects on adolescents of playing hours of video games per week, especially given that more frequent use of video games can raise increase both presence and aggressive affect in short exposures. It suggests that the over time use of video games may prime users to quickly experience aggressive affect and perhaps engage in aggressive behaviors when they encounter violence or unpleasant stimuli, whether mediated or not. This is particularly true when presence is increased. Also, as Anderson [14] suggested, there is a need for more longitudinal studies, as well as a need to educate the public about the potential effects of exposure to violent media and particularly the effects of playing violent video games. As suggested by Sherry's [17] meta-analysis, the effect of video games on people's aggression is likely to continue to increase. Video games are becoming more vivid, are being presented on larger screens, including surround sound, engaging characters, and compelling storylines, all of which have been shown to increase presence, which these data show increases aggression.

Acknowledgements

We would like to thank Mark Hamilton for his assistance with this project.

References

[1] T. Kim and F. Biocca, "Telepresence via Television: Two dimensions of Telepresence may have different connections to memory and persuasion," *Journal of Computer Mediated Communication*, vol. 3, pp. Retrieved from http://www.ascusc.org/jcmc/vol3/issue2/kim.html, 1997.

[2] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," *Journal of Computer-Mediated Communication*, vol. 3, pp. Retrieved March 10, 2006, from (http://www.ascusc.org/jcmc/vol3/issue2/), 1997.

[3] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *Journal of Communication*, vol. 42, pp. 73-93, 1992.

[4] M. Lombard, "Direct responses to people on the screen: Television and personal space," *Communication Research*, vol. 22, pp. 288-324, 1995.

[5] B. Reeves, B. Detenber, and J. Steuer, "New televisions: The effects of big pictures and big sound on viewer responses to the screen," presented at Information Systems Division of the International Communication Association, Washington, D.C., 1993, May.

[6] R. Tamborini, M. Eastin, K. Lachlan, P. Skalski, T. Fediuk, and R. Brady, "Violent virtual video games and hostile thoughts," *Journal of Broadcasting and electronic Media*, vol. 48, pp. 338-154, 2004.

[7] B. J. Bushman, "Moderating role of trait aggressiveness in the effects of violent media on aggression," *Journal of Personality and Social Psychology*, vol. 69, pp. 950-960, 1995.

[8] B. J. Bushman, "Individual differences in the extent and development in aggressive cognitive associative networks," *Personality and social psychology bulletin*, vol. 22, pp. 811-819, 1996.

[9] E. Donnerstein, Slaby, R., & Eron, L., " The mass media and youth violence," in *Violence and youth: Psychology's response*, vol. 2, E. R. J. Murray, & G. Comstock, Ed. Washington, DC: American Psychological Association, 1994, pp. 219-250.

[10] H. Paik and G. Comstock, "The effects of television violence on anti-social behavior: A meta-analysis," *Communication Research*, vol. 21, pp. 516-546, 1994.

[11] B. Wilson, D. Kunkel, D.Linz, W. J. Potter, E. Donnerstein, and S. Smith, "Violence in television programming overall: University of California Santa Barbara study," in *National Television Violence Study*, vol. 1. Thousand Oaks, CA: Sage, 1997.

[12] [M. H. Thomas, R. W. Horton, E. C. Lippencott, and R. S. Drabman, "Desensitization to portrayals of real-life aggression as a function of exposure to television violence," *Journal of Personality and Social Psychology*, vol. 35, pp. 450-458, 1977.

[13] J. B. Funk, H. B. Baldacci, T. Pasold, and J. Baumgardner, "Violence exposure in real-life, video games, television, movies, and the Internet: is there desensitization?" *Journal of Adolescence*, vol. 27, pp. 23-39, 2004.

[14] C. A. Anderson, "An update on the effects of playing violent video games," *Journal of Adolescence*, vol. 27, pp. 113-122, 2004.

[15] D. A. Gentile, P. J. Lynch, J. R. Linder, and D. A. Walsh, "The effects of violent video game habits on adolescent

hostility, aggressive behaviors, and school performance," *Journal of Adolescence*, vol. 27, pp. 5-22, 2004.
[16] C. A. Anderson and B. J. Bushman, "Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature," *Psychological Science*, vol. 12, pp. 353-359, 2001.
[17] J. Sherry, "The effects of violent video games on aggression: A meta-analysis," *Human Communication Research*, vol. 27, pp. 409-431, 2001.
[18] B. Krahe and I. Moller, "Playing violent electronic games, hostile attributional style and aggression-related norms in German adolescents," *Journal of Adolescence*, vol. 27, pp. 53-69, 2004.
[19] A. Bandura, "Social cognitive theory and social referencing. In (Ed.), (pp.)." in *Social referencing and the social construction of reality*, S. Feinman, Ed. New York: Plenum Press, 1992, pp. 175-208.
[20] A. Bandura, *Aggression: A social learning analysis*. Upper Saddle-River, NJ: Prentice-Hall., 1973.
[21] A. Bandura, "Social Cognitive theory of mass communication," in *Media Effects: Advances in theory and Research*, J. B. a. D. Zillmann, Ed. Mahwah, NJ: Lawrence Earlbaum Associates, 2002, pp. 121-153.
[22] L. Berkowitz and E. Rawlings, "Effects of film violence on inhibitions against subsequent aggression," *Journal of Abnormal and Social Psychology*, vol. 66, pp. 405-412, 1963.
[23] A. Bandura, "Social Cognitive Theory: An Agentic Perspective," *Annual Review of Psychology*, vol. Annual 1, pp. 1, 2001.
[24] R. G. Geen and D. Stonner, "Context effects in observed violence," *Journal of Personality and Social Psychology*, vol. 25, pp. 145-50, 1973.
[25] J. Jansz, "The Emotional Appeal of Violent video games for adolescent males," *Communication Theory*, vol. 15, pp. 219-241, 2005.
[26] C. A. Anderson and K. E. Dill., "Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life," *Journal of Personality and Social Psychology*, vol. 78, pp. 772-790, 2000.
[27] S. J. Kirsh, "Seeing the world through mortal kombat-colored glasses. Violent video games and the development of a short-term hostile attribution bias," *Childhood*, vol. 5, pp. 177-184, 1998.
[28] C. Heeter, "Being there: The subjective experience of presence," *Presence: Teleoperators and Virtual Environments*, vol. 1, pp. 262-271, 1992.
[29] ISPR, "Resources for the study of presence: About Presence. Retrieved April 14, 2004, from <http://www.ispr.info/>." 2000.
[30] [K. M. Lee, "Presence, Explicated," *Communication Theory*, vol. 14, pp. 27-50, 2004.
[31] M. Slater, M. Usoh, and A. Steed, "Depth of presence in virtual environment," *Presence: Teleoperators and Virtual Environments*, vol. 3, pp. 130-144, 1994.
[32] F. Biocca and K. L. Nowak, "Plugging your body into the telecommunication system: Mediated Embodiment, Media Interfaces, and Social Virtual Environments," in *Communication Technology and Society: Audience adoption and Uses*, D. Atkin and C. Lin, Eds. Cresskill, NJ: Hampton Press, 2002, pp. 407-447.
[33] K. L. Nowak, "The Influence of Anthropomorphism and Agency on Social Judgment in Virtual Environments," *Journal of Computer Mediated Communication*, vol. 9, pp. available http://jcmc.indiana.edu/vol9/issue2/nowak.html, 2004.
[34] M. Slater and M. Usoh., "Representations systems, perceptual position, and presence in immersive virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 2, pp. 221-233, 1993.
[35] M. Lombard, R. D. Reich, B. E. Grabe, C. C. Bracken, and T. B. Ditton, "Presence and Television; The role of screen size," *Human Communication Research*, vol. 26, pp. 75-95, 2000.
[36] D. F. Roberts, U. G. Foehr, and V. Rideout, "Generation M: Media in the lives of 8-18 year olds," Kaiser Family Foundation, Menlo Park, CA 2005, March.
[37] L. M. Wilfred., "Learning in affectively intense virtual environments," in. Rolla. Retrieved December 10, 2005 from: http://campus.umr.edu/lite/theses/2004_1_Wilfred_Lawrence.pdf: University of Missouri, 2004.
[38] G. Gunter, *Dimensions of television violence*. Aldershots, England: Gower, 1985.
[39] B. D. Bartholow and C. A. Anderson, "Effects of violent video games on aggressive behavior: Potential sex differences.," *Journal of Experimental Social Psychology*, vol. 38, pp. 283-290, 2002.
[40] E. Huesmann, "Psychological processes promoting the relation exposure to media violence and aggressive behavior by the viewer," *Journal of Social Issues*, vol. 42, pp. 125-139, 1986.
[41] M. Krcmar and M. C. Cooke, "Moral development and children's interpretations of justified and unjustified television violence," *Journal of Communication*, vol. 51, pp. 300-316, 2001.
[42] S. L. Smith, B. J. Wilson, D. Kunkel, D. Linz, W. J. Potter, C. M. Colvin, and E. Donnerstein, "Violence in television programming: University of California, Santa Barbara study," in *National Television Violence Study*, vol. 3. Newbury Park, CA: Sage, 1998, pp. 5-220.
[43] M. Lombard and T. Ditton, "Presence Measures," Temple University, Unpublished Manuscript February, 1999.
[44] K. Nowak and F. Biocca, "The Effect of Agency and Anthropomorphism on Users' sense of Telepresence, Social Presence and Copresence in virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 12, pp. 481-494. MIT Press, 2003.
[45] C. A. Anderson, W. E. Deuser, and K. DeNeve, "Hot tempers, hostile affect, hostile cognition and arousal: Tests of a general model of affective aggression," *Personality and Social Psychology Bulletin*, vol. 21, pp. 434-448, 1995.
[46] A. H. Buss and M. Perry, "The Aggression Questionnaire," *Journal of Personality and Social Psychology*, vol. 63, 1992.
[47] K. M. Farrar and M. Krcmar, "Measuring state and trait aggression: A short, cautionary tale," *Media Psychology*, In press.
[48] A. M. Rubin, "A multivariate analysis of "60 Minutes" viewing motivations," *Journalism Quarterly*, vol. 58, pp. 529 - 534, 1981.

Walk A Mile in Digital Shoes: The Impact of Embodied Perspective-Taking on The Reduction of Negative Stereotyping in Immersive Virtual Environments

Nick Yee, Jeremy Bailenson
Stanford University
{nyee@stanford.edu, bailenson@stanford.edu}

Abstract

In social psychology, perspective-taking has been shown to be a reliable method in reducing negative social stereotyping. These exercises have until now only relied on asking a person to imagine themselves in the mindset of another person. We argue that immersive virtual environments provide the unique opportunity to allow individuals to directly take the perspective of another person and thus may lead to a greater reduction in negative stereotypes. In the current work, we report on an initial experimental investigation into the benefits of embodied perspective-taking in immersive virtual environments. It was found that negative stereotyping of the elderly was significantly reduced when participants were placed in avatars of old people compared with those participants placed in avatars of young people. We discuss the implications of these results on theories of social interaction and on copresence.

1. Introduction

A great deal of research in social psychology has focused on the nature of stereotypes and prejudice. For example, since Allport's classic work developing a theoretical framework of social categories [1], it has become apparent that strong beliefs about inter-group differences can be created with minimal and oftentimes arbitrary decisions [2, 3]. For example, individuals randomly assigned to one of two different groups will have more positive expectations of members in their own group and more negative expectations of members of the other group despite the fact that there is no rational reason to differentiate members of one group over another.

Researchers have also demonstrated that stereotype activation oftentimes occurs with an automaticity that is beyond conscious control [4] and that the presence of these stereotypes leads to prejudicial interactions unless conscious intervention is applied [5]. These stereotypes not only impact minority groups in social interactions due to the prejudicial treatment they receive from others, but also create cognitive burdens for these minorities themselves as well. For example, negative stereotypes can lead to systematic underperformance via a mechanism known as stereotype threat [6-8]. In a landmark study of stereotype threat, when black students were given a verbal test that they were told was a direct measure of their intellectual ability, they performed worse than another group of black students who were told the test was about understanding different problem-solving strategies [8]. The authors of that study argued that the extra pressure caused by the fear of reinforcing a negative stereotype causes systematic underperforma
nce in ability tests.

These findings all point to the explicit and implicit hold that stereotypes and prejudice have on our society, and the difficulty in preventing tensions and conflicts due to the existence of stereotypes. In the current work, we explore potential interventions for decreasing the application of stereotypes, and in particular, we present empirical data from an experimental design that implemented an intervention method in immersive virtual reality (VR) with beneficial results.

1.1 Decreasing impact of stereotypes

The line of research into the nature of prejudice has led other researchers to explore ways to decrease the accessibility and application of stereotypes. The earliest method of intervention suggested was by Allport himself in his early work [1]. The *Contact Hypothesis* was the suggestion that social interaction between two groups of individuals would decrease the existing conflicts and tensions between them as mutual understanding occurred. The success of this form of intervention, however, depended on several factors, which Allport noted. For example, 1) the two groups must have equal status, 2) and share resources or power in such a way as to create a mutual interdependence, and 3) the context for interaction must be conducive to positive and friendly interactions. While these factors may be introduced in some artificially-created groups [2] to decrease inter-group tension, these factors may be hard to introduce in groups where power imbalances are deeply entrenched (such as with racial or gender stereotypes). Thus, the Contact Hypothesis offers one potential solution that is unfortunately highly situational.

Another intuitive intervention method is *thought suppression*. For example, individuals who try to avoid treating minorities prejudicially may attempt to suppress stereotypical references before the interaction. It has been shown, however, that deliberate thought suppression often backfires [9, 10]. Because a representation of the target stereotype must be articulated in order to suppress it, deliberate thought suppression has the effect of making stereotypical concepts and traits hyper-accessible. This search also functions as a form of repetitive priming. In other words, suppressed stereotypes often become more, rather than less, accessible and salient after the intervention.

1.2 Perspective Taking

One intervention method that has yielded positive results derives from the concept of *perspective-taking*. In social interactions, the fundamental attribution error affects how we think about and evaluate ourselves and others [11]. When we judge ourselves, we tend to rely on situational factors (i.e., "I did poorly on the test because I didn't sleep well the night before."). On the other hand, when we judge others, we tend to rely on dispositional factors (i.e., "He did poorly on the test because he's not that bright."). Thus, when people are forced to observe their own actions (via a video tape), they tend to make more dispositional rather than situational attributions [12]. The reverse is also true. When participants are asked to take the perspective of the person they are observing, participants tend to make situational rather than dispositional attributions [13].

More importantly, it has been found that perspective-taking leads to an increased overlap between the self and other. In one study, participants rated themselves and another person more similarly on a set of trait words in the perspective-taking condition in which they asked to take the point of view of another person than in a control condition [14]. It was also found that participants felt the target was more similar to themselves than control participants and liked the other person more after the perspective-taking exercise. Thus, on an individual level, perspective-taking had been shown to generate positive interpersonal effects.

Galinsky and Moskowitz [15] extended this work on perspective-taking into the domain of stereotypes and prejudice. They hypothesized that the benefits of perspective-taking may extend to inter-group evaluations and interactions. In other words, by encouraging people to focus on situational rather than dispositional factors via perspective-taking, they may rely less on stereotypes in evaluating and interacting with members of minority groups. In their work, they worked with stereotypes about the elderly because it has been shown that college-age students automatically associate negative traits with the elderly [16]. Participants were shown a photograph of an elderly man and were asked to write a short narrative essay about a typical day in the life of this individual. Participants also performed an implicit association task involving recognition of words related to old age. In a lexical decision task [17], words are flashed briefly on a computer screen and participants are asked to categorize the flashes as words or non-words. It has been shown that concepts and associations that are more accessible in a person's mind will be recognized faster than concepts that are less accessible. Thus, if an individual has a strong negative association with the elderly, then they are more likely to recognize words with negative connotations quicker (i.e., frail, wrinkled, sick) than neutral words or worlds with a positive connotation. It was found that perspective-taking decreased the amount of implicit and explicit stereotyping while increasing the amount of self-other overlap. The significant contribution of their work was in showing that the positive effects of perspective-taking can extend to the group level (i.e., improving evaluations of all elderly people after taking the perspective of one elderly man) rather than simply on the individual level (i.e., improving the evaluation of Tom after taking the perspective of Tom).

1.3 Collaborative Virtual Environments

If it were possible to convincingly place an individual into the body of an elderly person, rather than simply asking them to imagine this, we may expect this perspective-taking exercise to have an even stronger effect. Collaborative Virtual Environments [CVEs, see 18, 19, 20] make this manipulation possible. CVEs are communication systems in which multiple interactants share the same three-dimensional digital space despite occupying remote physical locations. In a CVE, immersive virtual environment technology monitors the movements and behaviors of individual interactants and renders those behaviors within the CVE via avatars (digital representations of people). These digital representations are tracked naturalistically by optical sensors, mechanical devices, and cameras. Thus, CVEs offer unique opportunities for social science research [21, 22].

1.4. Transformed Social Interactions

Unlike telephone conversations and videoconferences, the physical appearance and behavioral actions of avatars can be systematically filtered in immersive CVEs idiosyncratically for other interactants, amplifying or suppressing features and nonverbal signals in real-time for strategic purposes. Theoretically, these transformations should impact interactants' persuasive and instructional abilities. Previously, we outlined a theoretical framework for such strategic filtering of communicative behaviors called Transformed Social Interaction [23]. In a CVE, every user perceives their own digital rendering of the world and each other and these renderings need not be congruent. In other words, the target may perceive his or her own avatar as being attractive while the perceiver sees the target as being unattractive.

Previous work on transformed social interaction has demonstrated quite resoundingly that changing one's representation has large implications on other's in terms of social influence. In other words, transforming Avatar A strategically causes Avatar B to behave consistently with the representation of Avatar A (as opposed to the actual representation of Avatar A). A recent review [24] summarizes a number of studies that show social influence resulting from transformations in facial similarity, mimicry, and eye gaze.

Other research has also shown that alterations in digital self-representation can have a large impact on how a person behaves in virtual environments - a phenomenon termed The Proteus Effect [25]. For example, participants given attractive avatars are willing to walk closer and share more personal information to a stranger in a virtual environment than participants given unattractive avatars. Also, participants in taller avatars are significantly more willing to make unfair offers to their own advantage than participants in shorter avatars in a negotiation task. In other words, previous work has shown that alterations in self-representation can lead to significant changes in behavior.

In the current study, we were interested in altering self-representation for a different goal. Specifically, we wanted to explore whether changing a person's self-representation may help in reducing the negative stereotypes against particular social groups. For example, we could place college-age users into an elderly avatar to test whether embodied perspective-taking increases the positive evaluations of the elderly.

Presence is the concept of "being there", a measure of how immersive an environment is [26-31]. We argue that the beneficial effects of perspective-taking would be evident in a high presence situation such as the one proposed. Instead of asking people to imagine the world from the perspective of another person, immersive virtual reality allows us to place a person directly into the body of another person. Thus, in our study, we immersed participants into a virtual reality environment and presented them with an avatar via a virtual mirror. Following the work by Levy [32] and Galinsky and Moskowitz [15], we worked with stereotypes of the elderly in the study. Thus, participants were either given an avatar of a young or an elderly person and were forced to interact with another person while wearing the old or young avatar. We predicted that participants given avatars of elderly people would come to have fewer negative stereotypes of the elderly.

2. Method

2.1. Design

In a between-subjects design, participants were randomly assigned to have an avatar of an elderly person or an avatar of a young adult of the same gender. Participants then interacted with a confederate of the same gender who was blind to condition. Finally, participants completed a survey measure of their attitudes towards the elderly.

2.2. Participants

Forty-eight undergraduate students (24 men and 24 women) participated in the study for course credit.

2.3. Materials

2.3.1 Face Pretest. We selected faces of old and young people based on the results of a pretest. First, we found 12 digital photographs of individuals in each of the gender and age conditions needed in the study (thus 48 images altogether). These digital photographs were frontal photographs of individuals in well-lit conditions consisting of at least 400 by 400 pixels. To reduce variation, we selected only photographs of Caucasians who had no facial hair and were not wearing glasses.

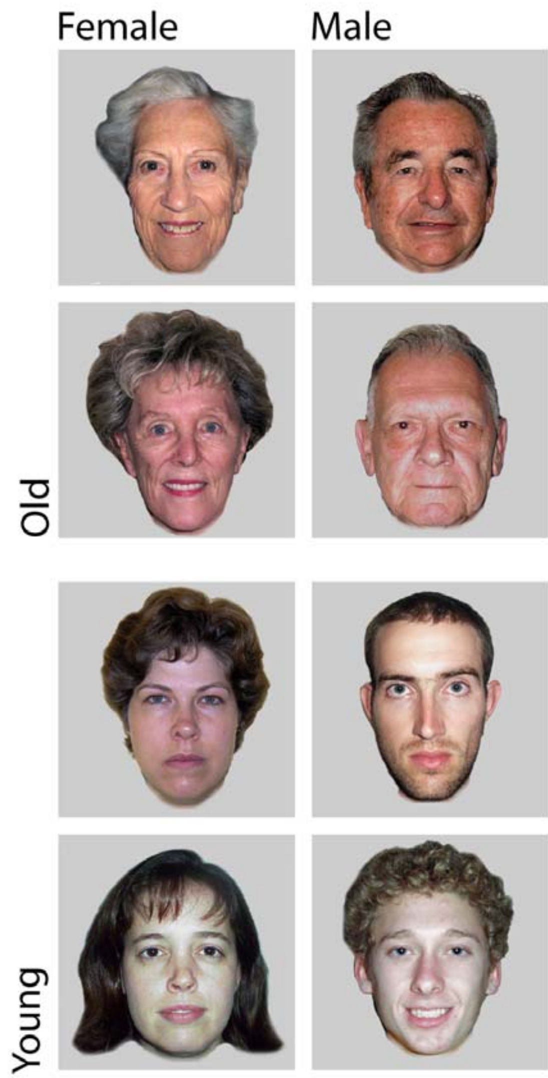


Figure 17 - Avatar faces selected from the pretest for the old and young conditions

Thirty-three undergraduate students who did not participate in the study were presented with each photograph sequentially in a randomized order. They were asked to estimate the age of the individual in the photo on a 6-point

Figure 18 - Equipment setup: A) head-mounted display, B) tracking camera, C) rendering machine.

scale (labeled as 16-25, 26-35, 36-45, 46-55, 56-65, 66-75) as well as rate the attractiveness of the individual on a 7-point fully-labeled scale (from "Extremely Attractive" to "Extremely Unattractive").

Our goal was to select photographs of individuals for both genders that had significantly different age ratings but non-significantly different attractiveness ratings near the mid-point of the attractiveness scale (labeled as "Average"). We selected the two faces for each gender and age condition (thus eight faces altogether) that were closest to the mid-point of the attractiveness scale. A repeated-measures analysis of these eight faces using face trial as the independent variable and attractiveness as the dependent variable was not significant ($F[7, 224] = 1.32$, $p = .24$). A repeated-measures analysis of these eight faces using age group and face trial as the independent variables and age estimate as the dependent variable showed that the effect of age group was significant ($F[1, 32] = 2693.60$, $p < .001$). In particular, the photographed individuals selected for the old condition were rated as significantly older ($M = 5.14$, $SE = .06$, thus between the labels of 55-65 and 66-75) than those in the young condition ($M = 1.67$, $SE = .04$, thus between the labels of 16-25 and 26-35). See Figure 17.

2.3.2. The Physical Lab Setting. The lab consisted of two rooms with an open doorway. In the room where the study took place, a black curtain divided the room. To ensure that confederates and participants were not biased by the attractiveness each other's real faces, confederates stayed behind this black curtain until the VR interaction began and thus never saw the participant's real face and vice versa.

2.3.3. The Virtual Setting. The virtual setting was a white room that had the same exact dimensions as the physical room participants were in. Two meters behind the participant was a virtual mirror that reflected the z-rotation (roll) of the head and body translation (translation on X,Y, and Z) of the participant with the designated face (See Figure 19). Thus, the mirror image tracked and reflected four degrees of freedom such that when the participant moved in physical space, his or her avatar moved in synchrony in the mirror. The confederate's avatar was located 5 meters in front of the participant, facing the participant, and remained invisible until the conversational portion of the experiment began. The confederate's avatar always had a young face of average attractiveness and also had an automated blink animation based on human blinking behavior.

2.4. Apparatus

Perspectively-correct stereoscopic images were rendered by a 1700 MHz Pentium IV computer with an NVIDIA 5950 graphics card, and were updated at an average frame rate of 60 Hz. The simulated viewpoint was continually updated as a function of the participants' head movements, which were tracked by a three-axis orientation sensing system (Intersense IS250, update rate of 150 Hz). The position of the participant along the x, y, and z planes were tracked via an optical tracking system (WorldViz PPT, update rate of 60 Hz). The system latency, or delay between a participant's movement and the resulting concomitant update in the head-mounted display (HMD) was 45 ms maximum. The software used to assimilate the rendering and tracking was Vizard 2.17. Participants wore an nVisor SX HMD that featured dual 1280 horizontal by 1024 vertical pixel resolution panels that refreshed at 60 Hz. The display optics presented a visual field subtending approximately 50 degrees horizontally by 38 degrees vertically. See Figure 18.

2.5. Procedure

Two research assistants - one male and one female - were present for each trial. Because the confederate was always the same gender as the participant, one research assistant would greet the participant and guide the study while the other would be the confederate based on the gender of the participant.

After informed consent, participants were told that the goal of the experiment was to study social interaction in virtual environments and that they would be having a conversation with another person in a virtual environment. Participants were then led into the room with the black curtain and shown how to wear and adjust the HMD. Once the virtual world was loaded, participants saw themselves in a room that was exactly the same dimensions as the physical lab room, as depicted in Figure 19. The research assistant drew open the curtains at this point.

Figure 19 - Screenshot of participant's point of

view in the virtual room with the mirror when he is embodied in an elderly avatar.

Participants were then asked by the lead research assistant to turn around 180 degrees and asked to verify that they saw a mirror in front of them. After verbal affirmation, participants were then told that this is how they appeared to others in the virtual room. Several procedures were used to make sure participants had enough time to observe their avatars' faces. First, they were asked to tilt or nod their head and verbally affirm whether the reflection was following them. Then they were asked to walk up closer to the mirror to get a good look at the face and verbally affirm that the mirror image responded correctly. They were then asked to tilt or nod their head again and verbally affirm that their mirror image followed them. And finally, they were asked to bend down at the knee and come back up and verbally affirm that the mirror image was following them. Every participant was thus exposed to the designated face for between 60 to 75 seconds.

Participants were then asked to turn back around to face the front (i.e., their original orientation). Slightly ahead of time, the research assistant had triggered the program to make the confederate's avatar visible to the participant in the virtual world. Participants were then reminded that others in the virtual room saw them as they had just seen themselves in the mirror. The lead research assistant then told the participants that the other person in the room would initiate the next portion of the experiment.

Figure 20 - Screenshot of participant's point of view with confederate across the room and beam of light for the movement task.

The confederate followed a strict script that was displayed in their HMD so they could follow the specific verbal and nonverbal procedures while interacting with the participant inside the CVE. In order to provide the subjects with the time and maximal opportunity to feel self-presence [26] from the designated avatar, subjects were asked to perform a variety of social interactions. First, the confederate asked the participant to walk to several locations in the room under the guise of getting used to moving in a virtual environment. These locations were shown to the participant as cones of light within the virtual environment. There were three such locations that the participants walked to sequentially (see Figure 20). Then participants - now back at the starting location, facing the confederate again - were asked to answer several questions: 1) "Tell me a little about yourself.", 2) "What makes you happy in life?", and 3) "What do you think are the most important things to have in life?". Participants also performed a brief memory exercise of trying to repeat a list of fifteen words. The purpose of the memory exercise was to reinforce the fact they were wearing an elderly avatar. Finally, the confederate asked participants to approach him or her to make the interaction more social. Participants were then taken out of the virtual environment and completed a questionnaire.

2.6. Measures

The questionnaire included three attitudinal measures towards the elderly. This collection of attitudinal measures were taken from Levy's [32] study on the effect of self-stereotyping on attitudinal differences towards the elderly.

2.6.1. Word Association. In the word association task, participants were asked the open-ended question "When you think of somebody old, what are the first five words that come to mind?" Two individuals who were blind to condition rated the responses according to the scoring criteria first developed in an earlier study by Levy and Langer [6]. In brief, each of the five words was scored on three dimensions: 1) whether it described an internal or external trait (i.e., wise vs. wrinkled), 2) whether it had a positive or negative connotation (i.e., kind vs. agitated), and 3) whether the word was associated with activity or inactivity (i.e., gardening vs. slow). Finally, the scores for each of the three ratings (low was always negative; high was always positive) were averaged across the five words for each participant. The inter-coder reliability was .66.

2.6.2. Indirect Attitudes. The second measure of participant's views on aging was derived from Palmore's [33] Fact on Aging Quiz - a true or false quiz that could be used to measure underlying biases towards the elderly. Again, the scoring was based on the earlier Levy and Langer [6] study. We used both positive and negative bias items in this measure. In a positive bias question, a true negative statement about the elderly is made (i.e., “Old people usually take longer to learn something new.”). If participants choose “false” as the response, then it means they have a positive bias towards the elderly. Conversely, in a negative bias item, a false negative statement about the elderly is made (i.e., “The majority of people over the age of 65 reside in nursing homes.”). If participants choose “true” as the response, then it means they have a negative bias towards the elderly. In this measure, we had a total of 6 positive bias items and 6 negative bias items (see Appendix A). We calculated a bias score for each individual by subtracting their negative bias score from their positive bias score.

2.6.3. Ambiguous Story. Finally, we included an ambiguous story - the Margaret story - drawn from Levy’s [32] study (see Appendix B). The story described a 73-year-old woman named Margaret who moves in with her adult daughter and attends a college reunion. The story consisted of 21 sentences and presented information that could be judged in both positive, neutral, or negative ways. For example, Margaret’s inability to concentrate during the long reunion speech could be attributed to her age (negative dispositional attribution), her lack of sleep the night before (neutral situational attribution), or the tedium everyone experiences in listening to a long speech (a neutral dispositional attribution).

Participants were asked to summarize the story in their own words as well as to describe Margaret. Two coders blind to condition read through and scored each participant’s responses on four dimensions: 1) was Margaret dependent or independent in her relationship with her daughter?, 2) is Margaret described as being imaginative or losing touch with reality (becoming senile)?, 3) is Margaret forgetful because she is getting old or because of situational factors?, and 4) is Margaret described overall positively or negatively? The inter-coder reliability was .68. These four scores were then combined to form an aggregate rating.

3. Results

3.1. Word Association

We ran a t-test on the word association scores from the two Age Conditions (old and young). We found a significant difference (t[44] = 2.60, p = .01) where participants in the old condition (M = 7.67, SD = 1.72) associated traits that were significantly more positive to the elderly than participants in the young condition (M = 6.42, SD = 1.54).

3.2. Indirect Attitudes

We ran a t-test on the overall bias score from the two Age Conditions. The effect was not significant (t[45] = .30, p = .77).

3.3. Ambiguous Story

We ran a t-test on the aggregate rating from the two Age Conditions. The effect was not significant (t[31] = .77, p = .45). See Table 1 for the detailed results of all our significance tests.

Table 1. Means and standard deviations for all our dependent measures.

	Young *M* (*SD*)	Old *M* (*SD*)	*t*	*p*
Word Association	6.42 (1.54)	7.67 (1.72)	2.60	.01
Indirect Attitudes	-1.13 (1.61)	-1.17 (2.06)	.30	.77
Ambiguous Story	10.10 (2.52)	9.47 (2.12)	.77	.45

4. Discussion

While our results did not show a consistent effect across all three dependent measures, it was extremely encouraging to find that such a short virtual interaction can change a person’s negative stereotypes at all. We found a significant effect in the word association task, but not in the indirect attitudes or ambiguous story tasks. On the other hand, similar inconsistencies were found in Levy’s [32] original study, which only demonstrated significant differences on one of several dependent measures. Thus, perhaps these other dependent measures were not particularly sensitive to the manipulations we chose. Nevertheless, it is encouraging that a brief immersion into the avatar of an elderly person has a significant effect on attitudes towards the elderly in general.

There were several limitations to the current study. First of all, the dependence on explicit attitudinal measures made it impossible to assess whether the manipulation had an effect on implicit stereotype activation. For example, the use of a lexical decision task (with words associated with the elderly) may be used in future studies to explore the effect of embodied perspective-taking on implicit stereotype activation. And secondly, it is difficult to minimize demand characteristics in the current experimental paradigm. During debriefing, we demonstrated that many participants in the old condition guessed the goal of the study. On the other hand, intervention methods (such as the contact hypothesis or thought suppression) are typically explicit as well. So while there are demand characteristics in both the narrative and embodied perspective-taking tasks, it may be the case that the increased amount of presence from the latter type of simulation is a more effective tool in reducing stereotypes. Furthermore, it is not clear how one might conceal the nature of the intervention in practice. After all, diversity training in school or business settings are never covert interventions.

Future work can expand on the current findings in several ways. First of all, as mentioned before, implicit measures should be included. The inclusion of implicit measures may also yield more consistent results, as lexical decision tasks are less susceptible to social desirability bias among other demand characteristics. Thus, these implicit measures may improve upon the inconsistent results apparent in both Levy's study and the current study. Secondly, it would be interesting to compare the effects of embodied perspective-taking with the effects of narrative perspective-taking. In other words, an estimation of the added benefit of using this intervention in immersive virtual reality. And finally, future work might explore how embodied perspective-taking impacts other stereotypes, such as race or gender. Previous work has demonstrated that VR is an effective tool for treating many types of psychological disorders (such as phobias or post-traumatic stress syndrome) [34-37]. In future work, we hope to add prejudice to this list.

Social inequality caused by stereotypes and prejudice is a problem that has no simple solutions. Perspective-taking has been shown to be a practical intervention method that decreases both implicit and explicit stereotyping. While the current study found inconsistent results for the use of embodied perspective-taking, the findings do suggest that this intervention in immersive virtual reality can have a positive effect on reducing negative stereotypes.

5. Appendix A

The following items were drawn from Palmore's original work [33] on the stereotypes people have of the elderly.

5.1. Positive bias items

1. Physical strength tends to decline in old age.
2. The five senses (sight, hearing, taste, touch, and smell) all tend to weaken in old age.
3. Old people usually take longer to learn something new.
4. Older people tend to react slower than younger people.
5. Lung vital capacity tends to decline in old age.
6. A person's height tends to decline in old age.

5.2. Negative bias items

1. About half of the people over the age of 65 in the US have Alzheimer's Disease.
2. The majority of people over the age of 65 reside in nursing homes.
3. The majority of old people lose interest in and capacity for sex.
4. About half of old people live below the poverty line.
5. Drivers over the age of 65 are more likely to get into car accidents than drivers under the age of 65.
6. Old people are more likely to victims of theft, murder, and burglary than people under the age of 65.

6. Appendix B

Because the "Margaret story" was not reproduced in Levy's earlier paper [32], we requested a copy from her and she kindly shared her original stimulus. Since the story has not been reproduced in full before, we include it here for reference.

6.1. The Margaret story

Margaret had just moved in with her daughter, son-in-law and grandchildren (aged 7, 5, and 1). Margaret's daughter, Anne had been worried about her mother. Margaret had been living alone since her husband died soon after her 70 th birthday, which was three years ago. Anne convinced her mother to move into her nearby home where she could ask her mother to baby-sit for her grandchildren and where she could better help her mother with meals and household chores.

The week Margaret moved in with her daughter happened to be the week of Margaret's 50 th college reunion. During Margaret's first night in her daughter's home, she woke up many times because her one year old grandson had an earache and cried throughout the night. The next day, Margaret was supposed to attend the dinner to honor the reunion of her college class. Margaret's college roommate Essie had convinced Margaret to attend the reunion which was to be held at the nearby Hyatt Hotel. Margaret tried to walk to the hotel but could not recognize the local streets and got lost. She finally wandered back to her daughter's house. When she walked in, her grandchildren started to giggle at their disheveled grandmother. Margaret's daughter decided to give her mother a ride to the reunion dinner and go as her escort.

When they arrived, Margaret introduced her daughter to some of her college acquaintances but could not recall many of their names. During the salmon dinner, Margaret found it difficult to concentrate on the speeches by the long-winded members of her class. She noticed that one of the speakers resembled a poodle. Margaret's attention shifted downward. She saw a napkin that had fallen under the table and thought how it looked very much like a squirrel. She then looked over at her daughter and started to think about how she used to take her to a playground with lots of swings and a big green sandbox. She smiled at her daughter and commented half out loud, "That sandbox - how you loved it." Anne looked at her mother with a concerned look and whispered, "Shhh! Be quiet mom!" After dinner, Anne drove her mother home.

Acknowledgements

The authors would like to thank Alexia Nielsen and Keith Avila for their assistance in running the experimental study, Andy Orin and Anna Jayaram for their assistance in coding the open-ended responses, as well as Becca Levy for sharing the "Margaret story" stimulus from her earlier study. The current work was partially supported by NSF grant 0527377 and funding from Stanford's Media X.

References

1. Allport, G., *The nature of prejudice*. 1954, Reading, MA: Addison-Wesley.
2. Sherif, M., et al., *Intergroup cooperation and competition: The Robbers Cave experiment*. 1961, Norman, OK: University Book Exchange.
3. Tajfel, H., et al., *Social categorization and intergroup behavior.* European Journal of Social Psychology, 1971. **1**: p. 149-178.
4. Bargh, J.A., *Automaticity in social psychology*, in *Social psychology: Handbook of basic principles*, E.T. Higgins and A.W. Kruglanski, Editors. 1996, Guilford: New York. p. 169-183.
5. Devine, P.G., *Stereotypes and prejudice: Their automatic and controlled components.* Journal of Personality and Social Psychology, 1989. **56**: p. 5-18.
6. Levy, B. and E. Langer, *Aging free from negative stereotypes: Successful memory in China and among the American deaf.* Journal of Personality and Social Psychology, 1994. **66**: p. 989-997.
7. Steele, C.M., *A threat in the air: How stereotypes shape the intellectual identities and performance of women and African Americans.* American Psychologist, 1997. **52**: p. 613-629.
8. Steele, C.M. and J. Aronson, *Stereotype threat and the intellectual test performance of African Americans.* Journal of Personality and Social Psychology, 1995. **69**: p. 797-811.
9. Macrae, C.N., et al., *Out of mind but back in sight: Stereotypes on the rebound.* Journal of Personality and Social Psychology, 1994. **67**: p. 808-817.
10. Wegner, D.M., et al., *Paradoxical effects of thought suppression.* Journal of Personality and Social Psychology, 1987. **58**: p. 409-418.
11. Jones, E.E. and R.E. Nisbett, *The actor and the observer: Divergent perceptions of the causes of behavior.* 1971, Morristown, NJ: General Learning Press.
12. Storms, M.D., *Videotape in the attribution process: Reversing actors' and observers' points of view.* Journal of Personality and Social Psychology, 1973. **27**: p. 165-175.
13. Regan, D.T. and J. Totten, *Empathy and attribution: Turning observers into actors.* Journal of Personality and Social Psychology, 1975. **32**: p. 850-856.
14. Davis, M.H., et al., *Effect of perspective taking on the cognitive representation of persons: A merging of self and other.* Journal of Personality and Social Psychology, 1996. **70**: p. 713-726.
15. Galinsky, A.D. and G.B. Moskowitz, *Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism.* Journal of Personality and Social Psychology, 2000. **78**: p. 708-724.
16. Perdue, C.W. and M.B. Gurtman, *Evidence for the automaticity of ageism. Journal of Experimental Social Psychology.* Journal of Personality and Social Psychology, 1990. **26**: p. 199-216.
17. Greenwald, A.G., D.E. McGhee, and J.K.L. Schwartz, *Measuring individual differences in implicit cognition: The implicit association test.* Journal of Personality and Social Psychology, 1998. **74**: p. 1464-1480.
18. Normand, V., et al., *The COVEN Project: Exploring Applicative, Technical, and Usage Dimensions of Collaborative Virtual Environment.* Presence: Teleoperators and Virtual Environments, 1999. **8**: p. 1999.
19. Slater, M., et al., *Small Group Behavior in a Virtual and Real Environment: A Comparative Study.* Presence: Teleoperators and Virtual Environments, 2000. **9**: p. 37-51.
20. Schroeder, R., ed. *The social life of avatars*. 2002, Springer-Verlag: London.
21. Blascovich, J., et al., *Immersive virtual environment technology as a methodological tool for social psychology.* Psychological Inquiry, 2002. **13**(2): p. 103-124.
22. Loomis, J., J. Blascovich, and A. Beall, *Immersive virtual environments as a basic research tool in psychology.* Behavior Research Methods, Instruments, and Computers, 1999. **31**: p. 557-564.
23. Bailenson, J., et al., *Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual Environments.* Presence, 2004. **13**(4): p. 428-441.
24. Bailenson, J., *Transformed Social Interaction in Collaborative Virtual Environments.*, in *Digital Media: Transformations in Human Communication*, P. Messaris and L. Humphreys, Editors. in press, Peter Lang: New York.

25. Yee, N. and J. Bailenson, *The Proteus Effect: The effect of transformed self-representation on behavior*. under review.
26. Lee, K.M., *Presence, explicated.* Communication Theory, 2004. **14**: p. 27-50.
27. Lombard, M. and T. Ditton, *At the heart of it all: The concept of presence.* Journal of Computer-Mediated Communication, 1997. **3**.
28. Biocca, F., C. Harms, and J. Burgoon, *Towards a more robust theory and measure of social presence: Review and suggested criteria.* Presence: Teleoperators and virtual environments., 2003. **12**: p. 156-480.
29. Slater, M., Usoh, M., & Steed, A., *Depth of presence in virtual environments.* Presence: Teleoperators and Virtual Environments, 1994. **3**: p. 130-144.
30. Bente, G., et al. *Measuring behavioral correlates of social presence in virtual encounters*. in *International Communication Association Conference, May 27–31*. 2004.
31. Ijsselsteijn, W., J. Freeman, and H. de Ridder, *Presence: Where are we?* Cyberpsychology and Behavior, 2001. **4**: p. 179-182.
32. Levy, B., *Improving memory in old age through implicit self-stereotyping.* Journal of Personality and Social Psychology, 1996. **71**: p. 1092-1107.
33. Palmore, E., *Facts on Aging Quiz: A handbook of uses and results.* 1988, New York: Springer.
34. Rizzo, A., et al., *Design and Development of a VR Therapy Application for Iraq War Veterans with PTSD.*, in *Technology and Informatics*, J. Westwood, Editor. 2005, OIS Press: Amsterdam.
35. Riva, G., B. Wiederhold, and E. Molinari, *Virtual Environments in Clinical Psychology and Neuroscience, Methods and Techniques in Advanced Patient-Therapist Interaction*. 1998, Amsterdam: IOS Press.
36. North, M., S. North, and J. Coble, *Virtual Reality Therapy: An Innovative Paradigm*. 1997, Colorado Springs: IPI Press.
37. Wiederhold, B. and M. Wiederhold, *Clinical observations during virtual reality therapy for specific phobias.* CyberPsychology & Behavior, 1999. **2**: p. 161-168.