Virtual Environments and Other Media for Being There Together: Towards a Convergence of Technologies, Uses, and Research Agendas

Ralph Schroeder, Oxford Internet Institute, Ralph.Schroeder@oii.ox.ac.uk

Abstract

This essay compares virtual environments (VEs) with three other technologies: videoconferencing, online spaces for socializing and gaming, and online awareness and social networking technologies. The aim is to see what we can learn about VEs from other communications media – and vice versa. The essay discusses both small group interaction and interaction in large online populations. The focus is on the characteristics of interaction and communication, but these are directly related to presence and copresence: some technologies are focused more on engagement with persons (copresence), whereas others are focused more on interactions with objects and the environment (presence). It is argued that with the increasing convergence between these technologies in their practical uses, research on VEs will need to take the different affordances and commonalities of these systems into account in its research agendas in order to develop useful and enjoyable systems.

Keywords: Presence, copresence, virtual environments, communication.

1. Background

Virtual environments (VE) technology is still evolving, but with the increasing use of computer-mediated communication (CMC), there are more and more overlaps between VEs and other forms of 'being there together' online. So far, however, there is hardly any exchange between research on VEs and other forms of CMC. This is because research on VEs has operated in a technology-specific way and been confined to certain research areas or disciplinary specialisms, with little interaction with other areas of research such as that on videoconferencing, online gaming, mobile phones, Instant Messaging (IM), or social networking technologies. This may be appropriate for studies of single-user VEs that are psychological or human-factors oriented, but single-user VEs are increasingly being eclipsed by uses of multi-user or collaborative VEs.

This essay will focus on the uses of multi-user VEs. In what follows there will be no separation between findings from social or media psychology, human-computer interaction and computer-supported cooperative work, and other computer and social sciences. This is based on the idea that these findings should ultimately converge in our understanding of online interaction and communication. Nevertheless, the essay will return, in the conclusion, to how different disciplines are contributing to the study of interaction in VEs and how the various research agendas are still separate or overlapping.

In this paper, VEs are defined as providing the sensory experience of being in a place other than the you are physically in, and being able to interact with that place [1, 2] A shorthand is to say that these are technologies for 'being there', and multi-user VEs for 'being there together' [3]. The other technologies that are discussed here also provide a sense of 'being there together', though they do so in a somewhat different way from VEs:

- Videoconferencing achieves the sensory perception of other(s) being there together by means of high-fidelity video capture of participants.
- Online spaces and games provide the experience of being there together by means of avatar embodiment which the user experiences with a first- or third-person point of view, a space is shared with other users, and communication is often via text.
- Online awareness or social networking technologies allow users to identify whether other users are there, available online, and/or how they are represented symbolically or in the form of images of themselves.

In the full version of this paper (available from the author on request), the paper summarizes some of the major findings about non-VE technology and then relate this in each case to social interaction in VEs, making a variety of comparisons, including key advantages and disadvantages of each technology. In this short version, the discussion and conclusion will synthesize and draw out the implications of these comparative findings for research.

2. Discussion

As we increasingly move online, there will be increasing engagement with video-captured persons and avatars. As the spaces which include these representations of others become more common, they will provide a familiar and conducive context for interacting with others. In this sense, a shift to being

there together which includes a variety of online representations is becoming a preferred 'media rich' mode of communication. At the same time, in certain circumstances users may prefer a mode of communication without online representations or that is less media rich.

It needs to be stressed here that fewer social cues or less media richness, however, does not necessarily mean that people do not have sense of the other person. As Walther [4] has argued, in CMC where social cues are minimal, such as text-only collaboration, people make more of an effort to represent themselves in words; they put more into constructing their identities and revealing more about themselves. This may take longer than in richer media, but it may mean that people get to know each better than in face-to-face interaction or in rich media interaction since they establish 'hyperpersonal' relationships.

The need for richer media arises in specific circumstances. One of the major requirements for fidelity of expression and awareness in small groups is to support turn-taking. In small groups, the expressiveness of avatar faces and bodies (non-verbal communication) has been a major research agenda within videoconferencing and VE research [5,6]. This applies to facial expression in virtual- and videoconferencing – plus object or task support in virtual conferencing. However, whether our appearance or how we present ourselves is crucial depends on if we already know people offline. And for object-focused tasks, it may not matter if the medium includes facial expressions or if the collaboration is between 'strangers' or 'friends' [7].

In small groups, a common focus of attention needs to be maintained, the flow of the interaction needs to be kept going, and the absence of attention of any one participant is perceived as such (if we consider that active participants are annoyed and distracted if other participants' attention is noticeably preoccupied elsewhere). In larger groups where participants face each other and form small clusters - in which case we are back to small group conferencing. As we move away from smaller groups to larger ones, however, the requirements shift from supporting turn-taking and a common focus of attention – to the rules and conventions governing social behaviour that apply to the populations of large spaces.

In larger online spaces, there are different requirement for online representations. In some cases, symbolic representations of availability and awareness of others may be sufficient. There may also be a requirement of consistent self-representation (where avatars need to recognize each other by appearance, or if one needs to find the same person via a profile on a social networking site). In other situations such as social networking sites a continual modification of one self-representation may be the appropriate norm to keep others interested or engaged [8].

In the larger online world, consistency may also have a broader significance, not in terms of a consistent representation of one's self, but in terms of being consistently represented as being available or aware of others in the same space. The online world – like virtual spaces for socializing and gaming – is very large – but unlike online spaces for gaming and socializing, it is not one space or world but many. Thus it may be important to know which online space the other person is available is available in, or whether they are available in several. The same applies to availability: where are others available? And again, are they available in multiple online spaces? The implication is that online populations must be distributed in such a way across different worlds or spaces that people and groups know where and when they can find each other (and people will, of course, want to be unavailable in some spaces and worlds). On the other hand, despite being vast, our engagement with others or with our networks only consists of a few people in terms of routine interaction and engagement.

Finally, the different functionalities of the four technologies discussed here, in small groups or large, and with a greater or lesser focus on objects and the environment, can be seen as distinct, or they can be seen as overlapping. They are distinct in that virtual and video conferencing are mainly a modality for meetings and small group tasks. Online spaces and gaming are a means for informal interaction and collaboration in the pursuit of game tasks. Social networking technologies provide a means to maintain a constant yet changing online identity vis-à-vis others, while IM and mobile phones provide a means to maintaining day-to-day availability and awareness. Here we can make a number of comparisons (see Figure 1 after references). These comparisons may be drawn too sharply – in reality many of them overlap. Still, the figure crystallizes some of the most typical characteristics.

On the other hand, these functionalities can also be regarded as a continuum varying on a single dimensions, whether they are more intensive or more extensive in terms of interpersonal engagement (though in two varieties, virtual- or video-mediated). Online environments for being there together have two peaks with a high level of interpersonal engagement (see Figure 2 after references) - immersive videoconferencing and immersive VEs – and they typically also afford the greatest copresence. In dyads and small groups, these two peaks allow the most powerful engagement with each other. (Although these peaks are currently separate, there is also a technology which merges video and 3D computer- generated environments, the blue-c system [9] which allows video-captured 3D images of users' faces and bodies to be represented in immersive 3D computer-generated spaces.)

From these peaks, there is a descending order of intensity as groups get larger and the extent of interaction and ability to communicate with each other are diminished – and as we move towards desktop systems with more limited capabilities. However, this is not a straightforward descent: online chat (as opposed to immersion) can have a high degree of engagement, and crowds can be highly engaged if there is common focus of attention.

On the other hand, we could start at the base of these peaks where most people are online with each other for a variety of purposes for large parts of the day via online awareness technologies – this is the most extensive interpersonal online engagement. Broadband, mobile phones and other devices are used as always-on technologies at work and in the home, with people drifting in and out of the awareness and engagement in smaller and larger groups online and being almost constantly connected with each other. Some of these technologies will have video-based representations of people, others will have avatars, and the two are merging to some extent as representations of people take mixed forms. Still, this is a much more common form of being there for most people, and consists of far greater spaces and worlds than immersive VEs and high-end videoconferencing – even if, as mentioned earlier, the way people interact with others on a routine basis is in smaller groups. Still, in comparison with these much larger groupings which are used on a day-to-day basis, immersive VEs and high-end videoconferencing are marginal to most people most of the time.

3. Conclusion

This paper has compared a number of technologies for being there together, mainly for communication but also for a variety of other purposes. A number of related technologies and practices have been left out (mainly for reasons of space) which are also used for being there together: for example, shared virtual or video spaces for interacting not primarily with people but with objects and environments. This includes shared visualizations, workspaces, and augmented and mixed reality systems which often enhance the real environment with a virtual space. Other forms of mobile computing and tools for online collaboration could also be included. There is a range of interaction here - with more collaboration at one end and more communication at the other.

Still, the technologies discussed here are the main ones for computer-mediated-communication and even if they will continue to exist side-by-side with each other, a number of characteristics of convergence can be anticipated:

1.The current problems of voice and video/graphics quality will be overcome.

2.Users will not be forced to use a particular modality of representing themselves or communicating because of technology constraints or tradeoffs. Instead, using text as opposed to voice, or a realistic or constructed avatar representation, or engaging with a small or large number of simultaneous users – will become a matter of choice, convenience, and suitability of different means of communicating for different circumstances.

3.Communication via videoconferencing, which has been a separate technology, will increasingly merge with digital telephony and with 3D virtual environments.

As these technologies converge, both in technological terms and in terms of uses, a number of questions arise which have already been discussed in passing which go beyond the traditional research agendas of presence research.

So far, presence and copresence have been studied either in terms of particular applications or as a measurable psychological state (and often both), and mostly for individual users. If, however, there is a shift towards a variety of technologies and uses which overlap and which are used for similar purposes, there will need to be a shift in research agendas to include how the self is presented and others are perceived in various online modalities; how (intensively or extensively) engaging the various representations of users are; and how various modalities support communication and interaction. The range of issues raised by ‘being there together’, apart from presence and copresence, could therefore be broadened to address a number of other questions:

1.What kind of appearance is conducive to interacting in situations of online copresence?

2.What kind of environment or space, small or large, is appropriate for different copresent encounters and for developing appropriate social norms to govern copresence?

3.When is a more realist, and when a more artificially constructed, identity conducive to copresence?

4.What kind of technological system, with what affordances [see 10 for this concept], is suitable for mutual availability and awareness in situations of online copresence?

5.How should online spaces and worlds be designed to support maintaining awareness of others and signaling availability, especially across a range of spaces and spaces?

6.What type of engagement, extensive or intensive, video- or virtual-mediated communication, is best suited for people to communicate throughout the day?

7.How to combine technologies and uses such they provide the most useful and enjoyable experience of being connected to others in online spaces and worlds?

People will increasingly traverse online networks with different online identities, making use of online

spaces and representations of themselves to engage with others. As our interaction with others moves online, we need to present and make ourselves available to others, just as we depend on other's representations of themselves and awareness of where they are online. As online connections become ever denser, the lessons that can be drawn across them will become ever more important. There are signs that videoconferencing of different kinds – high-end and low-end – may finally be turning into more than a rare occurence, even if it remains confined to certain niches. Yet the differences between high-end videoconferencing and immersive VEs and other technologies may become increasingly eroded. Research on VEs will benefit from engaging with this larger changing landscape of technologies for being there together.

References

[1] S.Ellis. Origins and Elements of Virtual Environments. In W. Barfield and T. Furness (eds.) *Virtual Environments and Advanced Interface Design*. New York: Oxford University Press, 14-57, 1995.

[2] M.J. Scheumie, P. van der Straaten, M. Krijn, C. van der Mast, C. Research on Presence in Virtual Reality: A Survey. *Cyberpsychology and Behaviour*, 4(2), 183-201, 2001.

[3] R. Schroeder. Being There Together and the Future of Connected Presence, *Presence: Journal of Teleoperators and Virtual Environments*, 15(4), 438-454, 2006.

[4] J.B.Walther. Computer-mediated communication: Impersonal, Interpersonal, and Hypersonal Interaction. *Communication Research*, 23, 3-43, 1996.

[5] M.Garau. *The Impact of Avatar Fidelity on Social Interaction in Virtual Environments*. Ph.D. thesis, Department of Computer Science, University College London, 2005.

[6] V. Vinayagamoorthy. *User Responses to Virtual Humans in Immersive Virtual Environments.* Ph.D. thesis, Department of Computer Science, University College London, 2006.

[7] A. Steed et. al. Strangers and Friends in Caves: An Exploratory Study of Collaboration in Networked IPT Systems. *ACM SIGGRAPH 2003 Proceedings on Interactive 3D Graphics*. New York: ACM Press, 51-54, 2003.

[8] A. Lenhart, M. Madden. Social Networking and Teens: An Overview. Pew Internet and American Life Project, http://www.pewinternet.org/PPF/r/198/report_display.asp, 2007.

[9] M. Gross et al. blue-C: A Spatially Immersive Display and 3D Video Portal for Telepresence. *Proceedings of ACM SIGGRAPH*, July, 819-827. 2003.

[10] I. Hutchby. *Conversation and Technology: From the Telephone to the Internet*. Cambridge: Polity, 2001.

	Virtual Environment	Videoconferencing	Online Spaces for Gaming and Socializing	Social Networking and Awareness Tools
Appearance	Face with limited expressiveness, body	Head and torso	Avatar	Iconic representation, photo
Environment	Room and larger spaces	Small space within room	World	Spaces consisting of pages and windows
Realism	High	High	Low	Low
Object and environment interaction	High	Very limited	High (but restricted by field of view)	High (but restricted)
Facial Expressiveness	Low	High	Low	Low
Group Size	Small	Small	Large	Large
Communication and Interaction	Synchronous, brief	Synchronous, brief	Synchronous, extensive sessions	Synchronous and asynchronous, constant
Communication modality	Voice	Voice	Text, sometimes voice	Text, sometimes voice
Key disadvantage	Expense, poor facial expressiveness	Expense, poor audio and bodily cues	Poor facial expressiveness	Lack of social cues
Key advantage	Object interaction	Facial expressiveness	Engaging setting for interaction	Awareness, availability, and self-presentation can be managed

Figure 1. Comparison of four technologies for 'Being There Together'

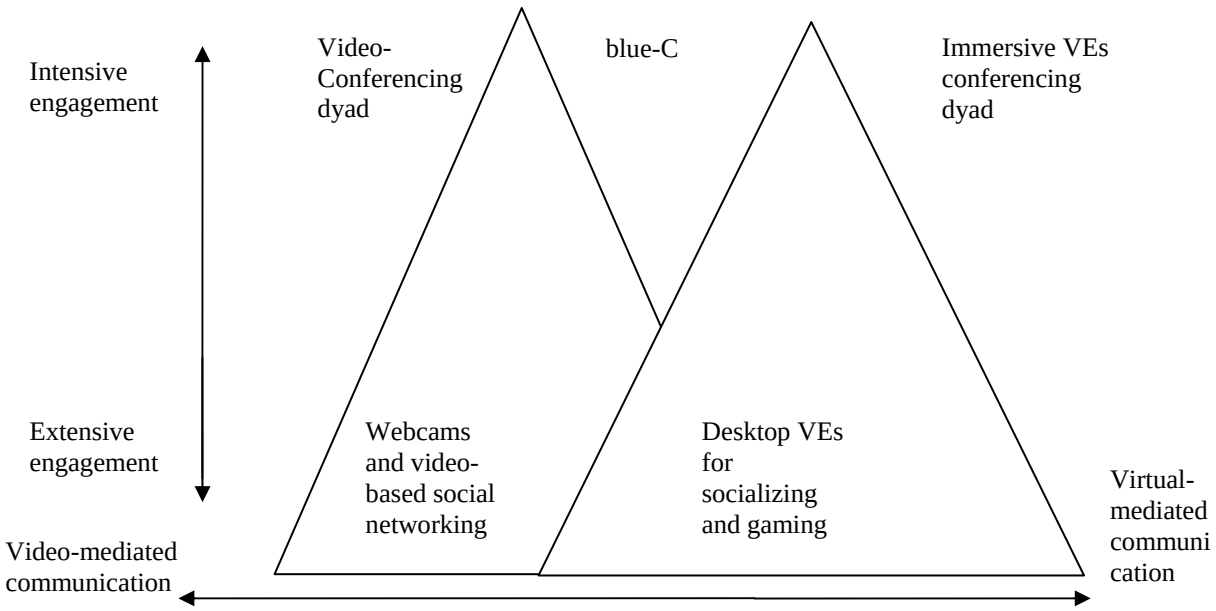

Figure 2. Level of engagement in four technologies for 'Being there Together'.