COMPUTER SUPPORTED COOPERATIVE WORK

Avatars at Work and Play:

Collaboration and Interaction in Shared Virtual Environments

Edited by

Ralph Schroeder and
Ann-Sofie Axelsson

Springer

AVATARS AT WORK AND PLAY

Computer Supported Cooperative Work

Volume 34

Avatars at Work and Play

Collaboration and Interaction in Shared Virtual Environments

Edited by

Ralph Schroeder

Oxford University, Oxford, U.K.

and

Ann-Sofie Axelsson

Chalmers University, Gothenburg, Sweden

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-3883-6 (HB)
ISBN-13 978-1-4020-3883-9 (HB)
ISBN-10 1-4020-3898-4 (e-book)
ISBN-13 978-1-4020-3898-3 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Printed in the Netherlands.

List of Contributors

Ann-Sofie Axelsson, Department of Technology Management and Economics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden annaxe@ chalmers.se

Jeremy N. Bailenson, Department of Communication, Stanford University, Stanford CA 94305-2050, USA bailenso@stanford.edu

Andrew C. Beall, Department of Psychology, University of California Santa Barbara, Santa Barbara CA 93106-9660, USA beall@psych.ucsb.edu

Marek Bell, Department of Computer Science, University of Glasgow, Glasgow G12 8QQ, UK marek@dcs.gla.ac.uk

Jim Blascovich, Department of Psychology, University of California Santa Barbara, Santa Barbara, CA 93106-9660, USA blascovi@psy.ucsb.edu

Barry Brown, Department of Computer Science, University of Glasgow, Glasgow G12 8QQ, UK Barry@dcs.gla.ac.uk

Lars Bråthe, Volvo Powertrain, SE-405 05 Gothenburg, Sweden Lars.Brathe@volvo.com

Katy Börner, School of Library and Information Science, Indiana University, Bloomington, IN 47405, USA katy@indiana.edu

Mari Siân Davies, Childrens Media Center and Department of Psychology, UCLA, Los Angeles, CA 90095, USA marisian@ucla.edu

Maia Garau, Department of Computer Science, University College London, London WC1E 6BT, UK m.garau@cs.ucl.ac.uk

Patricia M. Greenfield, Childrens Media Center and Department of Psychology, UCLA, Los Angeles, CA 90095, USA greenfield@psych.ucla.edu

Ilona Heldal, Department of Technology Management and Economics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden ilohel@chalmers.se

Mikael Jakobsson, Arts and Communication, Malmö University, SE-205 06 Malmö, Sweden mikael.jakobsson@k3.mah.se

Oliver Otto, The Centre for Virtual Environments, University of Salford, Manchester M5 4WT, UK o.otto@salford.ac.uk

Susan Persky, Department of Psychology, University of California Santa Barbara, Santa Barbara, CA 93106-9660, USA persky@verizon.net

Shashikant Penumarthy, School of Library and Information Science, Indiana University, Bloomington, IN 47405, USA sprao@indiana.edu

David Roberts, The Centre for Virtual Environments, University of Salford, Manchester M5 4WT, UK D.J.Roberts@salford.ac.uk

Ralph Schroeder, Oxford Internet Institute, University of Oxford, Oxford OX1 3JS, UK ralph.schroeder@oii.ox.ac.uk

Diane H. Sonnenwald, The Swedish School of Information and Library Science, Gothenburg University & University College of Borås, SE-501 90 Borås, Sweden diane.sonnenwald@hb.se

Maria Spante, Department of Technology Management and Economics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden marspa@chalmers.se

Anthony Steed, Computer Science, University College London, London WC1E 6BT, UK A.Steed@cs.ucl.ac.uk

Francis F. Steen, Childrens Media Center and Department of Communication Studies, UCLA, Los Angeles, CA 90095, USA steen@commstds.ucla.edu

Brendesha M. Tynes[*]**,** Childrens Media Center and Department of Psychology, UCLA, Los Angeles, CA 90095, USA btynesb@ucla.edu

Nick Yee, Department of Communication, Stanford University, Stanford, CA 94305-2050, USA nyee@stanford.edu

Robin Wolff, The Centre for Virtual Environments, University of Salford, Manchester M5 4WT, UK r.wolff@salford.ac.uk

[*] Currently African American Studies and Educational Psychology, University of Illinois, Urbana-Champaign, IL 61820, USA

Contents

WORK AND PLAY IN SHARED VIRTUAL ENVIRONMENTS: OVERLAPPING THEMES AND INTERSECTING RESEARCH AGENDAS

Ralph Schroeder and Ann-Sofie Axelsson

This volume, like its predecessor *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments* [1], aims to provide a state-of-the-art overview of research about how people interact in shared virtual environments (SVEs). Unlike the first volume, which covered a wide variety of topics, the essays collected here focus on two applications of SVEs; collaborative work and online gaming. These two areas are rapidly emerging as key drivers of SVE development. (Sometimes work applications are discussed under the label of collaborative virtual environments—or CVEs—but SVE is a broader term since it includes online gaming and socializing, so SVE is more suitable here.)

One reason for examining the two areas or work and play jointly is that although they are often treated in different academic arenas, in fact many issues overlap. As argued in the introduction to *The Social Life of Avatars*, certain issues—presence and copresence, communication between people in the environment, the appearance of the avatar and the environment, differences in the size of groups interacting, and how technology and the offline world shape the interaction—apply to *all* SVEs. Yet despite common themes, several academic disciplines are represented in this volume to tackle them—including psychology, sociology, computer science, and information sciences. Clearly, the study of SVEs requires that a number of disciplines work together.

This volume begins with two essays that investigate the important topic of avatar appearance, the appearance of the person inside the SVE. The essays by Bailenson and Beall and by Garau come at this from quite different perspectives. While Bailenson and Beall explore the plasticity of avatars, or the way in which the manipulation of appearance and behavior of avatars can be exploited for different purposes, Garau investigates the fidelity of avatar appearance with special reference to behavioral realism and eye gaze.

Bailenson and Beall demonstrate that it is easy to manipulate people's appearance. Changing facial appearance, allowing people to appear to be looking at several other people at the same time (non-zero sum gaze), and giving avatars virtual trainers that others cannot see—these and many other possibilities exist in SVEs that are not possible in face-to-face interaction. Their research, which they call "transformative social interaction", opens the way for investigating a host of social science questions in settings that can be controlled and manipulated. Their chapter makes a start in this direction (though there is some earlier related work by Blascovich [2] and by Slater and Steed [3]) by investigating, for example, how people respond when their own face is blended into that of the group they interact with, or when people are able to direct their gaze at two conversational partners simultaneously.

Eye gaze may seem like a very specialized topic, but as anyone who has studied interaction between people will know, in many instances eye gaze is the single most important form of non-verbal communication (and non-verbal communication may, of course, be more important than verbal communication). It is also very difficult to reproduce accurately in SVEs, though as Garau's chapter shows, it will be more important to focus on behavioral realism than on representational realism (or photorealism), which will have major implications for the design of SVE systems. Further, her findings suggest that, as there will always be trade-offs in implementing eye gaze and avatar fidelity, it may be that there are easier ways to provide more effective means for believable social interaction than is often thought.

One advantage of SVEs is that the interaction between people in the environments can easily be captured and analyzed. The next chapter by Penumarthy and Börner gives an excellent demonstration of this. Their essay is also a good example of investigating larger groups of people interacting in SVEs rather than the small groups of two or three that are typically studied. Put differently, their chapter addresses the area beyond the micro of small group encounters. This level is often difficult to capture and analyze in social science about the *real* world. In virtual worlds, however, the analysis is easily scalable (for some other examples, see [4, 5])—although, as the authors point out, patterns of interactions in virtual worlds will be different from real world ones.

We can also see in Penumarthy and Börner's essay, as in the one that follows by Sonnenwald, the beginnings of the systematic investigation into some basic building blocks of social interaction in SVEs; such as cooperation and competition, leadership (see also [6]) and status. As Sonnenwald shows, collaboration over the course of time with larger groups across a number of sites requires not only smoothly functioning technology, but even more importantly the social coordination of people and their adaptation to new roles in SVE settings. A key issue that emerges in this and several other papers in this volume—and one that has not been studied sufficiently since many SVE trials and experiences have been for shorter periods—is that a different dynamic sets in with longer-term routine collaboration (see also [7, 8]).

Sonnenwald also reports, in relation to another study of collaboration in which two participants used a haptic system for a science lab exercise and which compared pairs working side by side and pairs working across a network—that the latter is in many ways superior to the former. This is an important result since it is often claimed that distributed collaboration can never be as good as face-to-face collaboration. The only previous result (to our knowledge) which shows that collaboration in a SVE is practically as good as working face-to-face is our own study of pairs solving a spatial task with a Rubik's cube-type puzzle using networked immersive projection technology systems [9].

The study of SVEs has to a large extent focused on presence, copresence and on doing different tasks with different systems. Much less is known so far about the patterns of how the bodies of avatars interact with each other and with the environment. The chapter by Heldal, Bråthe, Steed and Schroeder analyzes this interaction in detail, focusing on pairs of users using networked immersive projection technology systems doing a number of tasks together. By analyzing their movements and conversation in great depth, the authors are able to highlight certain common successful and less successful forms of interaction. It is clear from this analysis that some elements that one might expect to be problematic are not; for example, going through each other's avatar bodies and through objects during certain phases in the collaboration (and despite the fact that these are "unnatural" forms of interaction). Conversely, some forms of interaction that one might expect to find unproblematic in fact present obstacles to smooth interaction; such as moving a non-tracked arm to point to objects, or navigating together and orienting oneself in a large space. These findings can only be obtained by means of closely examining such small sequences of interaction. The problem for future research, as they point out, will be to find out how general lessons can be drawn from these very brief and specific sequences.

For open-ended and less true-to-life tasks (such as those in the chapter just described) these issues may not be so pressing since participants can develop workarounds for many of the problems. Roberts, Otto and Wolff's essay addresses a different type of collaboration; working together with objects on a closely coupled task which requires close coordination in building a small structure together. One of their aims is to show, as some others have done, the advantages of handling objects in an immersive SVE as opposed to a desktop one. Another is to highlight that for this type of—again, closely coupled task—a lot of decisions need to be made about how, in the virtual world, objects can be passed from one person to another (who "owns" them?) and how objects and tools are used (how is "gravity" implemented? How to indicate when a screw has been successfully screwed in?).

These are some problems that do not exist for physical world collaboration. Roberts, Otto and Wolff also describe how implementing the technical aspects of simultaneously handling objects and using tools is by no means a trivial task in terms of handling network traffic and software design—since time and coordination are critical. Still, the main point of their essay is that they demonstrate that even for a scenario in which people need to work closely and accurately together, which is perhaps the most demanding scenario to implement in immersive SVEs, solutions can be found for very difficult problems, such as delays, consistency of objects, and the like.

As we saw earlier, it is important how "truthful"—in behavioral terms—avatars are. The chapter by Spante, Axelsson and Schroeder deals with a related issue for people collaborating with others via different systems; namely, that it is important to let users know what the capabilities of each others' systems are. Unless this information is made explicit, users will often make assumptions about the other person's avatar or system that are incorrect, and this can lead to misunderstandings. Spante, Axelsson and Schroeder argue that greater transparency by means of more information will improve interaction and learning about the other person's system—or, that "putting yourself into the other person's shoes" can lead to an enhanced experience of collaboration. It should be noted, however, that there are also drawbacks to this: for example, the user will need to bear this information about the other person's system in mind throughout the interaction, and this means that another piece of information is added to concentrating on the task and other aspects of interaction.

Here, it can be recalled that the whole point of Virtual Reality (VR) technology is supposed to be that this is a "natural" interface, or that SVEs do away with the interface; that is, that the interface is so realistic that the user does not need to worry about commands or other pieces of information. So keeping in mind what kind of system the other person is using will put information between the user and the interface. These issues will also apply to the kinds of artificially enhanced or altered scenarios in Bailenson and Beall's paper: knowing that the encounter has "artificial" features could either detract from "realism", or it could be made transparent—but in this case detract from the naturalness of the interaction or add to the "cognitive load" of the participants.

The essay by Persky and Blascovich about immersive gaming provides an interesting transition between the two parts of the book—since immersive SVEs have to date been almost exclusively used for work or research purposes. Online gaming, on the other hand, is almost invariably associated with desktop computers. Nevertheless, it can be envisaged that online games will become increasingly immersive. Persky and Blascovich's experiments supply a number of findings which anticipate this development: one is that playing a violent game in an immersive SVE—as one might expect—has a more powerful effect on aggressive feelings than playing a non-violent one, and that these feelings are stronger in an immersive than in a non-immersive (desktop) SVE. The same does not apply to an art-themed game; in this case creative feelings are not heightened by playing on an immersive VR system. (Again, one of the limitations of these findings is that they apply to short-term experiences of VEs.) Nevertheless, although violence and addiction have been obvious topics for online gaming on desktop computers, they will take on a new dimension with immersive SVE systems.

Yee's chapter about the massively multiplayer online role playing game (MMORPG) Everquest is intended to go beyond the study of violence and addiction in long-term online gameplay. With his extensive questionnaire responses from 30000 MMORPG players, we begin to have a better understanding of what attracts people to interacting online. Apart from steering us away from the stereotype of the a-social male teenager, his findings are also relevant to why people are drawn to immerse themselves in virtual worlds—which is closely related to the question of "presence" and "copresence" analyzed in the other contributions in the volume. Yee shows, to give just a small example, that women are more motivated by the "relationship", "immersion" and "escapism" factors than men. Another interesting finding is the possibility raised by his research that partners or parents and their children can learn about aspects of each others' personalities that they may not been able to discover in face-to-face relations with each other. These findings could be relevant not only to the design of online games, but also to collaborative work and other applications of SVEs.

Everquest is one of the online games in Yee's study, and this popular game is also the focus of Jakobsson's chapter. Like Yee, Jakobsson is interested in why people are attracted to virtual worlds, but his approach is quite different: He charts, in the manner of an ethnographic participant observer, how the relationship to the game and to others changes over the course of time. He points out that few people, and certainly not game designers, have thought about questions to do with longer-term engagement with virtual worlds, such as how to maintain relations with friends when leaving a particular game and the continuity between different worlds ("continuity" is a problem for the economies of virtual worlds, see Castronova's essays [10]). Jakobsson also describes how gameplay increasingly entails more "managerial" functions at the more advanced levels, such as coordinating team play with others. In the end, however, even this more complex level faces the problem of where to take player progression—ultimately, towards being able to leave the game in a suitably rewarding way.

The last two chapters overlap in that they both focus on the social glue that makes online social interaction pleasurable—mostly successfully in the case of *There*, and mostly unsuccessfully, it seems, for *The Sims Online.* Brown and Bell's chapter about the online virtual world *There* argues for example that the design of the text bubbles for conversational turn-taking and how objects can be handled together provide a shared focus that enhances sociability. They also argue, like the first two chapters in this volume, that embodiment in online gaming plays an important role in facilitating social interaction (see also [11]). Their chapter is a good counterpoint to Steen, Davies, Tynes, and Greenfield's account of *The Sims Online.* Steen et al. argue that *The Sims Online* incorporated precisely the wrong elements—that is, the elaborate social structure—from the (highly successful) offline *Sims* game, and that the designers did not build enough features facilitating more immediate sociability around conversation and interaction with objects into the online version.

The essay by Steen et al. does not deal with SVEs in the strict sense that is used in the other contributions (for definitions of SVEs and Virtual Reality, see the introductory chapter in [1]), since control over one's first-person visual perspective and direct manipulation of the environment is lacking. Still, this environment is interesting because it is a large-scale and much discussed environment which hoped to replicate many of the complex features and depth of the real-world social interaction more thoroughly than other online social spaces.

As we have seen, this question—of the artificiality of the environment and the "structuredness" of interpersonal interaction—is one that is addressed in different ways in earlier chapters. Brown and Bell are thus surely correct to say that designers of collaborative work environments will benefit from studying online games. A further reason for this is that online gaming needs to engage the user over a long period of time. The interaction that is described in several of the work related chapters would, if it were to take place over longer periods, not only need smooth interaction with devices, but also promote a sense of sociability and of the participants enjoying each other's company.

Many other connections between these essays could be made. In the end, they are all linked by a common goal—of better understanding the uses of SVEs for practical work purposes and for leisure or socializing purposes. The first volume of essays *The Social Life of Avatars* was mainly exploratory and mapped out different research directions. With this volume, our hope is that research on SVEs is well on its way towards better insights into what makes them more effective and enjoyable—and to improved SVE design.

References

1. Schroeder, R. (Ed.) (2002). *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer.
2. Blascovich, J. (2002). Social influence within immersive virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 127–145.
3. Slater, M. & Steed, A. (2002). Meeting people virtually: Experiments in shared virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 146–171.
4. Craven, M., Benford, S., Greenhalgh, C., Wyver, J., Brazier, C.J., Oldroyd, A., & Regan, T. (2001). Ages of Avatar: Community building for inhabited television. In E. Churchill & M.

Reddy (Eds.), *CVE2000: Proceedings of the Third International Conference on Collaborative Virtual Environments.* New York: ACM Press, pp. 189–194.
5. Schroeder, R., Huxor, A., & Smith, A. (2001). Activeworlds: Geography and social interaction in virtual reality. *Futures: A Journal of Forecasting, Planning and Policy 33*: 569–587.
6. Slater, M., Sadagic, A., Usoh, M., & Schroeder, R. (2000). Small group behaviour in a virtual and real environment: A comparative study. *Presence: Journal of Teleoperators and Virtual Environments 9*(1): 37–51.
7. Hudson-Smith, A. (2002). 30 Days in Activeworlds—Community, design and terrorism in a virtual world. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 77–89.
8. Steed, A., Spante, M., Schroeder, R., Heldal, I., & Axelsson, A.S. (2003). Strangers and friends in caves: An exploratory study of collaboration in networked IPT Systems for extended periods of time. In *ACM SIGGRAPH 2003 Symposium on Interactive 3D Graphics.* New York: ACM Press, pp. 51–54.
9. Schroeder, R., Steed, A., Axelsson, A.S., Heldal, I., Abelin, Å., Wideström, J., Nilsson, A., & Slater, M. (2001). Collaborating in networked immersive spaces: As good as being there together? *Computers & Graphics 25*: 781–788.
10. Castronova, E. (2005). Available at http://mypage.iu.edu/∼castro/
11. Taylor, T.L. (2002). Living digitally: Embodiment in virtual worlds. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 40–62.

Chapter 1

TRANSFORMED SOCIAL INTERACTION: EXPLORING THE DIGITAL PLASTICITY OF AVATARS

Jeremy N. Bailenson and Andrew C. Beall

1. Introduction

What does it mean to be you? How drastically can a person change and still remain, in the eyes of either themselves or their peers, the same person? Until recently, these questions were typically asked in the context of philosophy, psychoanalysis, or science fiction. However, the increasingly common use of avatars during computer-mediated communication, collaborative virtual environments (CVEs) in particular, are quickly changing these once abstract questions into practical quandaries that are fascinating, thought-provoking, potentially paradigm shifting for those who study social interaction, and potentially devastating to the traditional concept of human communication.

Historically, even before the advent of computers, people have demonstrated a consistent practice of extending their identities. As Turkle [1, p. 31] points out:

> The computer of course, is not unique as an extension of self. At each point in our lives, we seek to project ourselves into the world. The youngest child will eagerly pick up crayons and modeling clay. We paint, we work, we keep journals, we start companies, we build things that express the diversity of our personal and intellectual sensibilities. Yet the computer offers us new opportunities as a medium that embodies our ideas and expresses our diversity.

Extending one's sense of self in the form of abstract representation is one of our most fundamental expressions of humanity. But abstract extension is not the only manner in which we manipulate the conception of the self. In addition to using abstract means to extend one's identity, humans also engage in the practice of using tangible means to transform the self. Figure 1-1 demonstrates some of these self transformations that occur currently, without the use of digital

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 1–16.

	Appearance	Nonverbal Behavior	Verbal Behavior
Short term	Haircuts Makeup	Mimicking Ingratiating Gestures	Lying Word Choice
Long term	Plastic Surgery Dieting	Habit Suppression Table Manners	Oratory Training Language-Acquisition

Figure 1-1. Non-digital transformations of self currently used.

technology. Before the dawn of avatars and computer-mediated communication, this process of self transformation was minor, incremental, and required vast amounts of resources.

However, given the advent of collaborative virtual reality technology [2–5], as well as the surging popularity of interacting with digital representations via collaborative desktop technology [6], researchers have begun to systematically explore this phenomenon of *Transformed Social Interaction* [7]. TSI involves novel techniques that permit changing the nature of social interaction by providing interactants with methods to enhance or degrade interpersonal communication. TSI allows interactants themselves, or alternatively a moderator of the CVE, to selectively filter and augment the appearance, verbal behavior, and nonverbal behavior of their avatars. Furthermore, TSI also allows the interactants to filter the context in which an interaction occurs. In our previous work outlining the theoretical framework of TSI, we provided three dimensions for transformations during interaction.

The first dimension of TSI is transforming *sensory abilities*. These transformations augment human perceptual abilities. For example, one can have "invisible consultants" present in a collaborative virtual environment, ranging from other avatars of assistants rendered only to you who scrutinize other interactants, to algorithms that give you real-time summary statistics about the movements and attentions of others (which are automatically collected in a CVE in order to render behaviors). As a potential application, teachers using distance learning applications can have "attention monitors" that automatically use eye gaze, facial expressions and other gestures as a mechanism to localize students who may not understand a given lesson. That teacher can then tailor his or her attention more towards the students higher in need. As another example, teachers can render virtual nametags (displayed to the teacher only) inserted over their students' avatars. Consequently, even in a distance learning classroom of hundreds, the students' names will always be at an instructor's disposal without having to consult a seating chart or a list.

The second dimension is *situational context*. These transformations involve changes to the temporal or spatial structure of an interaction. For example, each interactant can optimally adjust the geographical configuration of the room—in a distance learning paradigm, every single student in a class of twenty can sit right up front, next to the teacher, and perceive his or her peers as sitting behind. Furthermore, real-time use of "pause" and "rewind" during an interaction (while one's avatar exhibits stock behaviors produced by an "auto-pilot" algorithm) may be quite an effective tool to increase comprehension and productivity during interaction. Another example of transforming the situational contexts is to utilize *multilateral perspectives*. In a normal conversation, interactants can only take on a single perspective—their own. However, in a CVE, one can adopt the visual point of view of any avatar in the entire room. Either by bouncing her entire field of view to the spatial location of other avatars in the interaction, or by keeping "windows" in the corners of the virtual display that show in real time the fields of views of other interactants, it is possible for an interactant to see the behavior of her own avatar, as they occur, from the eyes of other interactants. Previous research has used either role playing scenarios [8] or observational seating arrangements [9] to cause experimental subjects to take on the perspectives of others in an interaction, and has demonstrated that this process is an extremely useful tool for fostering more efficient and effective interactions. Equipping an interactant with the real-time ability to see one's avatar from another point of view should only enhance these previous findings concerning the benefits of taking other perspectives.

The third dimension of TSI is *self-representation*. These transformations involve decoupling the rendered appearance or behaviors of avatars from the human driving the avatar. In other words, interactants choose the way in which their avatars are rendered to others in the CVE, and that rendering can follow as closely or as disparately to the actual state of the humans driving the avatars as they so desire. The focus of this chapter will be to discuss this third dimension in greater detail. While transforming situational contexts and sensory abilities are fascinating constructs, thoroughly discussing all three dimensions is beyond the scope of the current work.

This idea of decoupling representation from actual behavior has received some attention from researchers previously exploring CVEs. For example, [10] as well as [11] discussed *truthfulness* in representation, Biocca [12] introduced a concept known as *hyperpresence*, using novel visual dimensions to express otherwise abstract emotions or behaviors, and, moreover, numerous scholars debate the pros and cons of abstract digital identities [1, 13]. Furthermore, Jaron Lanier, considered by many to be one of the central figures in the history of immersive virtual reality, often makes an analogy between the human using immersive virtual reality and the "aplysia", a sea-slug that can quickly change its surface features such as body shape and skin color. Before virtual reality, humans had to resort to makeup, plastic surgery, or elaborate costumes to achieve these goals. William Gibson [14, p. 117] may have put it best when he declared that, once the technology supports such transformations, it is inevitable that people take advantage of "the infinite plasticity of the digital".

In sum, the idea of changing the appearance and behaviors of one's representation in immersive virtual reality has been a consistent theme in the development of the technology. The goals of the Transformed Social Interaction paradigm are threefold: (1) to explore and actually implement these strategies in collaborative virtual environments, (2) to put human avatars in CVEs and to measure which types of TSI tools they actually use during interaction, and (3) to examine the impact that TSI has on the effectiveness of interaction in general, as well as the impact on the specific goals of particular interactants. In the current chapter, we provide an overview of the empirical research conducted to date using avatars to examine TSI, and then discuss some of the broader implications of these digital transformations.

2. Transforming Avatar Appearance

This section reviews a series of TSI applications concerning the static appearance of one's avatar, some of which have been already tested using behavioral science studies in CVEs, others that have yet to receive empirical examination.

2.1. Identity Capture

The nature of a three-dimensional model used to render an avatar lends itself quite easily to applying known algorithms that transform facial structure according to known landmark points on the head and face. Once a face is digitized, there are an infinite number of simple morphing techniques that alter the three-dimensional structure and surface features of that face. This practice can be a powerful tool during interaction.

For example, persuaders can absorb aspects of an audience member's identity to create implicit feelings of similarity. Imagine the hypothetical case in which Gray Davis (the past governor of California, depicted in the leftmost panel of figure 1-2) is attempting to woo the constituents of a locale in which the voters are primarily fans of Arnold Schwarzenegger (the governor of California that ousted Davis) depicted in the rightmost panel of figure 1-2.

Research in social psychology has demonstrated large effects of similarity on social influence, in that a potential influencer who is more similar to a given person (compared to a less similar influencer) is considered more attractive

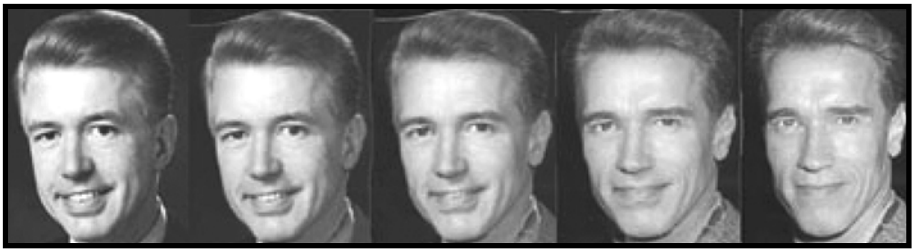


Figure 1-2. A digital morph of the two-dimensional avatars of Gray Davis (left) to Arnold Schwarzenegger (right).

[15] and persuasive [16], is more likely to make a sale [17], and is more likely to receive altruistic help in a dire situation [18]. Consequently, using digital technology to "absorb" physical aspects of other interactants in a CVE may provide distinct advantages for individuals who seek to influence others, either in a positive manner (e.g., a teacher during distance learning), or in a manner not so wholesome (e.g., a politician trying to underhandedly co-opt votes). Moreover, this type of a transformation may be particularly effective in situations in which the transformation remains implicit [19]. In other words, the effect of the transformation may be strongest when CVE interactants do not consciously detect their own face morphed into the face of the potential influencer.

To test this hypothesis, we brought Stanford University undergraduate students into the lab and used a simple morphing procedure with MagicMorph software [20, 21] to blend their faces in with an unfamiliar politician, Jim Hahn, a mayor of Los Angeles. Figure 1-3 depicts images of two undergraduate students as well as two blends that are each compromised of 60% of Jim Hahn and 40% of their own features.

The main hypothesis in this study [22] was that participants would be more likely to vote for a candidate that is morphed with their own face than a candidate that is morphed with someone else's face. In other words, by capturing a substantial portion of a voter's facial structure, a candidate breeds a feeling of familiarity, which is an extremely effective strategy for swaying preference [23].

Our findings in this study demonstrated two important patterns. First, out of 36 participants, only two detected that their own face was morphed into the candidate, even when we explicitly asked them to name one person like whom the candidate looked. Interestingly, their responses often demonstrated an implicit similarity (e.g., "He looks like my grandfather," or "He looks really familiar but I am not sure who he is"), but very rarely indicated a detection of the self. Second, overall there was a preference for candidates that were morphed with the self over candidates that were morphed with others, though the effect was strongest for white male participants (who were similar enough to the picture of Jim Hahn to create a successful morph) and for people interested in politics (who ostensibly were more motivated to pay attention to the photograph of the

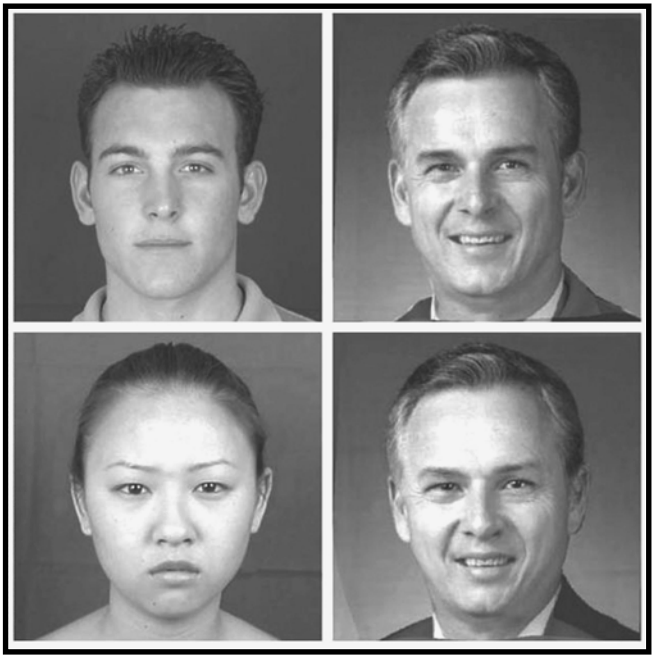


Figure 1-3. Pictures of the participants are on the left; the blend of 60% of an unfamiliar politician and 40% of the given participant is on the right.

candidate). In sum, very few participants noticed that their face was morphed into the political candidate, but implicitly the presence of themselves in the candidate gave the candidate a greater ability to influence those participants.

2.2. Team Face

A related study [24] examined the use of TSI for collaborative teams by creating a "Team Face". Given the underlying notion that teams function more cooperatively when they embrace commonalties (e.g., dress codes, uniforms) it is logical to consider that organizations would consider extending these team features to the rendering of avatars. Consider the faces in figure 1-4.

Figure 1-4. Four participants (left four panels) and their team face (far right), a morph that includes 25% of each of them.

The face on the far right is a morphed avatar that includes the faces from all four of the participants at equal contributions. In our study, participants (32 in total: four sets of four participants of each gender) received two persuasive messages: one delivered by their own team face, and one delivered by a team face that did not include their own face.

In this study, only three participants noticed their own face present inside the team face when explicitly asked to name one person like whom the face looked. In regards to persuasion, our results indicated that when participants received a persuasive message from an avatar wearing the team face, they were more likely to scrutinize the arguments. Specifically, arguments that were strong (determined by pre-testing) were seen as stronger when received by one's own team face than when received by a different team face, and the opposite pattern occurred for weak arguments.

This pattern is quite consistent with what would be predicted by the elaboration-likelihood model of Petty and Cacioppo [25]. According to that model, people processing a persuasive message utilize either the central route (i.e., dedicate cognitive resources towards actually working through the logical strengths and weaknesses of an argument) or the peripheral route (i.e., analyze the message only in terms of quick heuristics and surface features). In the study using team faces, participants were more likely to process a message centrally when the message was presented by their own team face than when presented by another team face—they were more likely to accept a strong argument and less likely to accept a weak argument. In sum, these preliminary data indicate that interacting with an agent wearing one's own team face causes that person to dedicate more energy towards the task at hand.

These two studies [22, 24] have been utilized solely with two-dimensional avatars in non-immersive displays. Current projects are extending this work to three-dimensional avatars in immersive virtual reality simulations that feature not only the texture being morphed between one or more faces but the underlying shape of the three-dimensional model as well. Previous research has demonstrated that three-dimensional models of a person's head and face built with photogrammetric software is sufficient to capture a majority of the visual features of one's physical self, both in terms of how people treat their own virtual selves [26] and in terms of how others treat familiar virtual representations of others [27].

2.3. *Acoustic Image*

While the majority of research and development in virtual environment technology has focused on stimulating the visual senses, the technology to richly stimulate the auditory senses is not far behind and possibly holds as much promise in its ability to transform social interactions amongst individuals as does its visual counterpart. Just a few years ago the process to render accurate spatialized (three-dimensional) sound required specialized and expensive digital signal processing hardware. Today, all this processing can be done on consumer-class PCs while easily leaving enough system resources left-over for the user's primary applications. In day-to-day living, we all take spatialized sound for granted just as we take binocular vision for granted. Only when you stop and reflect on the acoustical richness of our natural environments do you realize how much information is derived from the sensed locations of objects: without looking you know from where behind you your colleague is calling your name or that you better quickly step to one side and not the other to avoid being hit by a speeding bicyclist. Spatialization is partly what enables the "cocktail party phenomena" to occur—namely the ability to selectively filter out an unwanted conversation from an attended conversation. As such, our ability to synthetically render these cues in correspondence to three-dimensional visual images enables accurate reconstruction of physical spaces.

More interesting, however, are the possibilities arising from purposely altering the correspondence between the visual and acoustic images. By "warping" relational context, one can hand pick targets that are made maximally available along different channels. Research in cognitive psychology shows that human information processing is capacity limited and that these bottlenecks are largely independent for the visual and auditory channels. This means that by decoupling the visual and auditory contexts one could potentially empower a CVE user with the ability to maximize her sensory bandwidth and information processing abilities. For instance, in a meeting scenario one might place two different persons centered in one's field of attention, person A centered visually and person B centered acoustically. This way both A and B could be monitored quite carefully for their reactions to a presentation, albeit along different dimensions.

Just as it is possible to spatialize sound in real time, it is also possible to alter the characteristics of human speech in real time. Various software and hardware solutions are available on the consumer market today that can be used to alter one's voice in order to disguise one's identity. While it is not typically easy to transform a male voice into a female voice or vice versa, it is easy to alter a voice with a partial pitch and timbre shift that markedly changes the characteristics so that even someone familiar with the individual would unlikely recognize his identity. The implications of this regarding transforming social interaction are considerable. First, this technology enables the use of duplex voice as a communication channel while still maintaining the anonymity that digital representation allows. Already users in the online gaming community are using this technology to alter their digital personas.

But changing voice to disguise is just one possibility; voice can be transformed in a way that captures the acoustic identity just as the photographs can be morphed to do the same. One form of voice cloning is to sample a small amount of another's voice (e.g., 30 seconds or so) and analyze the frequency components to determine the mean tendencies and then use those statistics to modestly alter the pitch and timbre of your own voice using tools available today. In this way, you could partially transform your voice. While we know of no research that has done so, we believe the end result would be similar to the studies we have discussed in the visual domain. Perhaps a closer analogy to visual morphing is a voice cloning technology recently commercialized by AT&T Labs known as "concatenative speech synthesis." From a sample of 10–40 hours of recorded speech by a particular individual, it is possible to train a text-to-speech engine that captures the nuances of a particular individual's voice and then synthesize novel speech as if it came from that individual [28]. While the technology is impressive, it certainly still has a "robotic" ring to it—but its potential in CVE use is considerable.

As the next section demonstrates, extending TSI into immersive virtual reality simulations in which interactants' gestures and expressions are tracked bring in a host of new avenues to explore, and allow for extremely powerful demonstrations of strategies that change the way people interact with one another.

3. Transformations of Avatar Behavior

One of the most powerful aspects of immersive virtual reality, and in particular naturalistic nonverbal behavior tracking, is one that receives very little attention. In order to render behaviors onto an avatar as they are performed by the human, one must record in fine detail the actual behaviors of the human. Typically, the recordings of these physical movements are instantly discarded after they occur, or perhaps archived, similar to security video footage. However, one of the most powerful mechanisms behind TSI involves analyzing, filtering, enhancing, or blocking this behavior tracking data in real time during the interaction. In the current section, we review some previous research in which interactants have transformed their own nonverbal behavior as it occurs, and discuss some of the vast number of future directions for work within this paradigm.

3.1. Non-Zero-Sum Gaze

One example of these TSI "nonverbal superpowers" is *non-zero-sum gaze (NSZG)*: providing direct mutual gaze at more than a single interactant at once. Previous research has demonstrated that eye gaze is an extremely important cue: directing gaze at someone (compared to looking away from him or her) causes presenters to be more persuasive [29] and more effective as teachers [30–32]; it increases physiological arousal in terms of heartbeat [33], and generally acts as a signal for interest [34]. In sum, people who use mutual gaze increase their ability to engage a large audience as well as to accomplish a number of conversational goals.

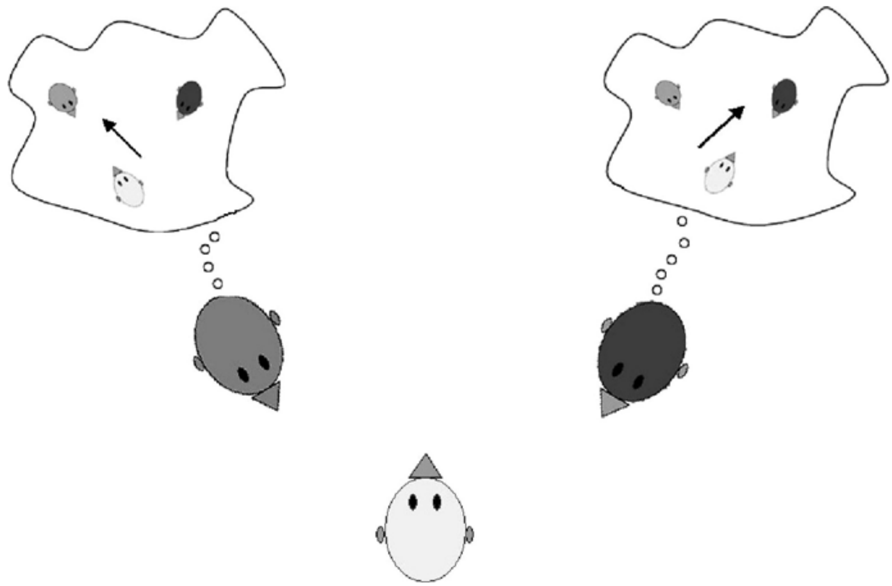


Figure 1-5. Non-zero-sum Gaze: Both the interactant on the top left and on the top right perceive the sole mutual gaze of the interactant on the bottom.

In face-to-face interaction, gaze is zero sum. In other words, if interactant X looks directly at interactant Y for 80% of the time, it is not possible for X to look directly at interactant Z for more than 20% of the time. However, interaction among avatars using TSI is not bound by this constraint. In a CVE, the virtual environment is individually rendered for each interactant locally at extremely high frame-rates. Consequently, with digital avatars, an interactant can have his avatar rendered differently for each other interactant, and appear to maintain mutual gaze with both Y and Z for a majority of the conversation, as figure 1-5 demonstrates.

NZSG allows a conversationalist to maintain the illusion that he or she is looking at an entire roomful of interactants. Previous research has implemented avatars that use "non veridical" algorithms to drive eye movements. For example, [35] implemented eye animations that were inferred from the verbal flow of the interaction. In other words, while head movements of interactants were tracked veridically, animation of the eyes themselves were driven not by the people's actual movements, but instead based on an algorithm based on speaking turns. These authors found that the conversation functioned quite well given this decoupling of rendered eye movements from actual eye movements, outperforming a number of other experimental conditions including an audio-only interaction.

Moreover, there has been research directly examining the phenomenon of NZSG. Two studies [36, 37] have utilized a paradigm in which a single presenter read a passage to two listeners inside an immersive CVE. All three interactants were of the same gender, wore stereoscopic, head-mounted displays, and had their head movements and mouth movements tracked and rendered. The presenter's avatar either looked directly at each of the other two speakers simultaneously for 100% of the time (augmented gaze) or utilized normal, zero-sum gaze. Moreover, the presenter was always blind to the experimental condition; in the augmented condition an algorithm automatically scaled down the magnitude of the presenter's head orientation movements (pitch, yaw, and roll) by a factor of 20 and redirected it at the eyes of both listeners.

Results across those two studies demonstrated three important findings: (1) participants never detected that the augmented gaze was not in fact backed by real gaze, despite being stared at for 100% of the time, (2) participants returned gaze to the presenter more often in the augmented condition than in the normal condition, and (3) participants (females to a greater extant than males) were more persuaded by a presenter implementing augmented gaze than a presenter implementing normal gaze.

The potential to use this tool should be extremely tempting across a number of conversational contexts ranging from distance education to sales pitch meetings to online dating chatrooms. Given the preliminary evidence described above, it is clear that avatar-gaze powered by algorithms, as opposed actual human behavior, can be at the very least innocuous, and most likely quite effective, during conversation.

3.2. Digital Chameleons

Chartrand and Bargh [38, p. 893] describe and provide empirical evidence for the Chameleon effect: when a person mimics our nonverbal behavior, that person has a greater chance of influencing us:

> Such a Chameleon effect may manifest itself different ways. One may notice using the idiosyncratic verbal expressions or speech inflexions of a friend. Or one may notice crossing one's arms while talking to someone else who has his or her arm's crossed. Common to all such cases is that one typically does not notice doing these things—if at all—until after the fact.

Data from Chartrand and Bargh's studies demonstrate that when people copy our gestures we like them better, interact more smoothly with them, and are more likely to provide them favor.

Given that typical rendering methods require capturing extremely detailed data concerning their gestures and actions, CVEs lend themselves towards utilizing mimic algorithms at very little added cost. Either from a "nonverbal profile" built from user historical archive data, or from slight adjustments to real-time gestures, it is quite easy for interactants to morph (or even fully replace) their own nonverbal behaviors with those of their conversational partners. There are many motives for interactants to implement the digital chameleon in CVEs, ranging from subtle attempts to achieve influence to powering their avatar with some type of “autopilot” while the user temporarily abdicates his or her seat in the CVE.

Previous research [37] demonstrated that participants often do not detect their own head movements when those movements are rendered at a delay onto other interactants in a CVE. Consequently, to test the digital chameleon hypothesis, Bailenson and Yee [24] ran an experiment in which undergraduate students sat in an immersive virtual environment, at a virtual table, across from an embodied agent. The agent proceeded to read a persuasive passage approximately four minutes long to the participants, whose head orientation movements were tracked while the scene was rendered to them stereoscopically through a head-mounted display. For participants in the *mimic condition*, the agent’s head movements were the exact same movements (on pitch, yaw, and roll) as the participants with a lag of 4 seconds. In other words, however the participant moved his or her head, the agent mimicked that movement 4 seconds later. For a separate group of participants in the *recorded condition*, the agent’s head movements were simply a playback of one of the other participants from the mimic condition.

Results of this study demonstrated a huge difference between groups. Agents that mimicked the participants were far more successful at persuading the participants and were seen as more likable than recorded agents. This effect occurred despite the fact that hardly any of the participants detected their own gestures in the behavior of the agents when given a variety of post-experiment questionnaires. These findings are extremely powerful. In order to render the behaviors of an avatar effectively, one must record in high detail all of the actions of the interactants. However, by doing so, the door is opened for other interactants (as well as embodied agents) to employ many types of nonverbal chameleon strategies. In this way, all interactants, some with less than altruistic motives, may achieve a new level of advantage in interaction.

Mimicry is also possible in the auditory channel. Recently, a team at ATR Media Information Science Laboratories in Japan succeeded in doing so [39]. Their idea was to avoid the obstacles of speech recognition and semantics and instead to mimic the overall rhythm and intonation of a speaker. To see if this idea would work, participants were asked to work with an animated agent whom they were told in advance would possess the speech skills of a 1-year-old child. The participants’ task was to make toy animals out of building blocks on the computer screen and to teach the agent the names of the toys being built. The agent child would then produce humming like sounds that responded in ways that mimicked the participants’ speech rhythms, intonations, and loudness. In a formal study, the levels of mimicry were varied and the effect on the participants’ subjective ratings of the agent were then assessed. Ratings were taken that measured cooperation, learning ability, task achievement, comfort, friendliness, and sympathy. The avatar that mimicked 80% of the time scored highest in user ratings. Just as with the studies reported above on head motions, these findings show that by isolating low-bandwidth dimensions of an interaction it is possible to create a sense of mimicry that does not require a top–down understanding of the interaction.

3.3. Other Behavioral Transformations

There are countless other ways to envision using TSI with the behavior of an avatar. For example, during interaction in CVEs, the automatic maintenance of a “poker face” is possible; any emotion or gesture that one believes to be particularly telling can just be filtered out, assuming one can track and categorize that gesture. Similarly, troubling habitual behaviors such as nervous tics or inappropriate giggles can be wholly eliminated from the behaviors of one’s avatars. On the other hand, behaviors that are often hard to generate in certain situations, such as a “genuine smile”, can be easily rendered on one’s avatar with the push of a button.

4. Implications and Outlook

The Orwellian themes behind this communication paradigm and research program are quite apparent. Even the preliminary findings discussed in this chapter concerning identity capture, face-morphing, augmented gaze, and digital mimicry are cause for concern, given the huge potential for misuse of TSI by advertisers, politicians, and anyone else who may seek to influence people via computer-mediated communication. On a more basic level, not being able to trust the very pillars of the communication process—what a person looks like and how they behave—presents interactants with a difficult position. One may ask whether or not it is ethical to keep the behaviors and appearance of your avatar close enough to veridicality in order to prove your identity to other interactants, but to then pick and choose strategic venues to decouple what is virtual from what is real. Is TSI fundamentally different from nose jobs, teeth-whitening, self-help books and white lies?

The answer is unclear. Currently, digital audio streams are “sanitized” over cell phone lines such that the digital information is transformed to present an optimal voice stream using simple algorithms. While this is an extremely mild form of TSI, it is important to point out that very few users of cell phones mind or even notice this transformation. Moreover, the potential ethical concerns of TSI largely vanish if one assumes that all interactants in a CVE are aware of the potential for everyone to rampantly use these transformations.

On a more practical note, an important question to consider is whether or not interactants will bother to pay attention to each other's behavior if there is no reason to suspect those behaviors are genuine. These strategic transformations utilized in CVEs may become so rampant that the original intent of a CVE—fostering multiple communication channels between physically remote individuals—is rendered completely obsolete. People may completely ignore the nonverbal cues of avatars, given that there is no reason to suspect the cue is genuine. On the other hand, as certain cues become non-diagnostic (e.g., it becomes impossible to infer one's mental state from one's facial expression), one can make the argument that interactants will always find the subtle conversational cues that are in fact indicative of actual behavior, appearance or mental state. For example, anecdotal evidence suggests that interactants speaking on the telephone (who do not have any visual cues available) are much more sensitive to slight pauses in the conversation than face-to-face interactants.

CVE programmers may be able to create an extremely persuasive illusion using an avatar empowered with TSI, but will it be possible to mask all truth from an interaction? If there is a lesson to be learned by various forms of mediated communication, it is that people adapt quite well to new technologies. Kendon [40] describes a concept known as *interactional synchrony*, the complex dance that occurs between (1) the multiple channels (i.e., verbal and nonverbal) of a single person during an interaction, and (2) those multiple channels as two interactants respond to one another. Kendon's studies indicated that there are extremely rigid and predictable patterns that occur among these channels during interaction. However, despite this consistent complexity of behavior during conversation, humans are quite adept at maintaining an effective interaction if a channel is removed, for example speaking on the telephone.

Taking away a channel of communication is one thing, but scrambling and transforming the natural correlation among multiple channels is another level of disruption entirely. Transformed social interaction does exactly that, decoupling the normal pairing of behaviors during interaction and, at the whim of interactants, changing the rules of the conversational dance completely. One would expect conversations to completely break down given such an extreme disruption to the traditional order of conversational pragmatics. However, given the results from the empirical investigations of TSI to date, which admittedly are quite limited and preliminary, this has not been the case. Interactants do not seem particularly disturbed by any of the TSI strategies discussed in this paper, and for the most part remain completely unaware of the breakdown among conversational channels.

As future research proceeds, and researchers and systems developers tamper more and more with the structure of interaction, we will provide a true test of the endurance of this conversational structure. One can imagine an equilibrium point in which sufficient amounts of conversational synchrony is preserved, but each interactant is utilizing TSI to the fullest advantage. As systems employing avatars that use these algorithms become widespread, it is essential that this balance point between truth and transformation is achieved. Otherwise, if actions by conversational partners are ships passing in the night, the demise of CVEs and computer-mediated interactions is inevitable.

References

1. Turkle, S. (1995). *Life on the Screen: Identity in the Age of the Internet.* New York: Simon & Schuster.
2. Schroeder, R. (2002). Social interaction in virtual environments: Key issues, common themes, and a framework for research, in R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*, (London: Springer), pp. 1–18.
3. Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. (2002). Immersive virtual environment technology: Not just another research tool for social psychology. *Psychological Inquiry, 13*: 103–124.
4. Slater, A. Sadagic, M. Usoh, R., & Schroeder, R. (2000). Small group behaviour in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments, 9*(1): 37–51.
5. Normand, V., Babski, C., Benford, S., Bullock, A., Carion, S., Chrysanthou, Y., *et al.* (1999). The COVEN project: Exploring applicative, technical and usage dimensions of collaborative virtual environments. *Presence: Teleoperators and Virtual Environments, 8*(2): 218–236.
6. Yee, N., chapter in this volume.
7. Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J., & Turk, M. (2004). Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments, 13*(4): 428–444.
8. Davis, M.H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology, 70*: 713–726.
9. Taylor, S.E. & Fiske, S.T. (1975). Point of view and perception of causality, *Journal of Personality and Social Psychology, 32*: 439–445.
10. Benford, S., Bowers, J., Fahlen, L., Greenhalgh, C., & Snowdon, D. (1995). User embodiment in collaborative virtual environments. In *Proceedings of CHI'95*, New York, ACM Press, pp. 242–249.
11. Loomis, J.M., Blascovich, J., & Beall, A.C. (1999). Immersive virtual environments as a basic research tool in psychology. *Behavior Research Methods, Instruments, and Computers, 31*(4): 557–564.
12. Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication Online, 3*(2). Available at http://www.ascusc.org/jcmc/vol3/issue2/-biocca2.html
13. Rheingold, H. (2000). *The Virtual Community: Homesteading on the Electronic Frontier*. Revised Edition. Cambridge: MIT Press.
14. Gibson, W. (1999). *All Tomorrow's Parties*. Ace Books.
15. Shanteau, J. & Nagy, G. (1979). Probability of acceptance in dating choice. *Journal of Personality and Social Psychology, 37*: 522–533.
16. Byrne, D. (1971). *The Attraction Paradigm*. New York: Academic Press.
17. Brock, T.C. (1965). Communicator-recipient similarity and decision change. *Journal of Personality and Social Psychology, 1*: 650–654.
18. Gaertner, S.L. & Dovidio, J.F. (1977). The subtlety of white racism, arousal and helping behavior. *Journal of Personality and Social Psychology, 35*: 691–707.

19. Bargh, J.A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype priming on action. *Journal of Personality and Social Psychology, 71*: 230–244.
20. Blanz, V. & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *SIGGRAPH'99 Conference Proceedings*, pp. 187–194.
21. Busey, T.A. (1988). Physical and psychological representations of faces: Evidence from morphing. *Psychological Science, 9*: 476–483.
22. Bailenson, J.N., Garland, P., Iyengar, S., & Yee, N. (2004). The effects of morphing similarity onto the faces of political candidates. Manuscript under review.
23. Zajonc, R.B. (1971). Brainwash: Familiarity breeds comfort. *Psychology Today, 3*(9): 60–64.
24. Bailenson, J.N. & Yee, N. (2004). *Transformed Social Interaction and the Behavioral and Photographic Capture of Self*. Stanford Technical Report.
25. Petty, R.E. & Cacioppo, J.T. (1986). *The Elaboration Likelihood Model of Persuasion*. New York: Academic Press.
26. Bailenson, J.N., Beall, A.C., Blascovich, J., Raimundo, M., & Weisbuch, M. (2001). Intelligents agents who wear your face: User's reactions to the virtual self. In A. de Antonio, R. Aylett, D. Ballin (Eds.), *Lecture Notes in Artificial Intelligence*, 2190: 86–99.
27. 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.
28. Guernsey, L. (2001). Software is called capable of copying any human voice. *New York Times*. July 31: Section A, Page 1, Column 1.
29. Morton, G. (1980). Effect of eye contact and distance on the verbal reinforcement of attitude. *Journal of Social Psychology, 111*: 73–78.
30. Sherwood, J.V. (1987). Facilitative effects of gaze upon learning. *Perceptual and Motor Skills, 64*: 1275–1278.
31. Otteson, J.P. & Otteson, C.R. (1979). Effect of teacher's gaze on children's story recall. *Perceptual and Motor Skills, 50*: 35–42.
32. Fry, R. & Smith, G.F. (1975). The effects of feedback and eye contact on performance of a digit-encoding task. *Journal of Social Psychology, 96*: 145–146.
33. Wellens, A.R. (1987). Heart-rate changes in response to shifts in interpersonal gaze from liked and disliked others. *Perceptual & Motor Skills, 64*: 595–598.
34. Argyle, M. (1988). *Bodily Communication*. (2nd ed.). London, UK: Methuen.
35. Garau, M., Slater, M., Bee, S., & Sasse, M.A. (2001). The impact of eye gaze on communication using humanoid avatars. *Proceedings of the SIG-CHI Conference on Human Factors in Computing Systems*, March 31–April 5, Seattle, WA, USA, pp. 309–316.
36. Beall, A.C., Bailenson, J.N., Loomis, J., Blascovich, J., & Rex, C. (2003). Non-zero-sum mutual gaze in immersive virtual environments. *Proceedings of HCI International 2003*, Crete.
37. Bailenson, J.N., Beall, A.C., Blascovich, J., Loomis, J., & Turk, M. (2004). Non-Zero-Sum Gaze and Persuasion. Paper presented in the Top Papers in Communication and Technology Session at the 54th Annual Conference of the International Communication Association, New Orleans, LA.
38. Chartrand, T.L. & Bargh, J. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality & Social Psychology, 76*(6): 893–910.
39. Suzuki, N., Takeuchi, Y., Ishii, K., & Okada, M. (2003). Effects of echoic mimicry using hummed sounds on human-computer interaction. *Speech Communication, 40*(4): 559–573.
40. Kendon, A. (1977). *Studies in the Behavior of Social Interaction*. Indiana University: Bloomington.

Chapter 2

SELECTIVE FIDELITY: INVESTIGATING PRIORITIES FOR THE CREATION OF EXPRESSIVE AVATARS

Maia Garau

Recent works of cyberfiction have depicted a not-so-distant future where the Internet has developed into a fully three-dimensional and immersive datascape simultaneously accessible by millions of networked users. This virtual world is described as having spatial properties similar to the physical world and its virtual cities are populated by digital proxies of people, called avatars. The multisensory sophistication of this shared space is such that it supports interpersonal communication on a level of richness interchangeable with face-to-face interaction. The vision presented encapsulates two of the central goals not only of collaborative virtual environments (CVEs), but also of any communication medium. First, to enable groups of people to collaborate and interact socially in an efficient and enjoyable way, and second, to foster the illusion that people are together when in reality they are in distinct physical locations.

CVEs have the makings of a potentially powerful medium of communication that heralds new promises and challenges. It is their inherently spatial property that sets them apart from other collaborative media. Though videoconferencing and groupware systems allow users to interact visually, the 3D context of each person's physical environment is lost. This can pose difficulties in small group interaction where conversation management can be disrupted by ambiguous eye gaze cues. The loss of 3D context can also be particularly problematic in tasks for which it is essential to preserve spatial relationships, such as remote acting rehearsals. CVEs can begin to address these concerns by placing geographically dispersed users in a shared, computer-generated space where they can interact with the environment and with other users represented by avatars. Immersive interfaces can also offer multimodal, surrounding experiences that can create a strong sense of being inside that artificial space (presence), and sometimes of being there with others (copresence). As mediators of users' actions and

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 17–38.

appearance, avatars are likely to play a significant role in social interaction in CVEs.

One of the central challenges in the development of CVEs is the creation of expressive avatars capable of representing users' actions and intentions in real time. This chapter focuses on the issue of avatar fidelity, arguing for the need to explore priorities by investigating the impact of avatar appearance and behaviour on the experience of interaction. It presents research on minimal fidelity, and discusses its implications for the future development of CVEs as a viable communications medium.

1. CVEs as a Communication Medium

CVEs are networked, computer-generated environments capable of supporting human-to-human communication by allowing users to interact with the space and with each other via graphical embodiments called avatars. CVEs can be used explicitly for work-related purposes, but also for social interaction and play; applications can range from conferencing, simulation and training, shared visualisation and collaborative design, to social communities and multiplayer games. Avatars play a significant role in all of these contexts because they embody the user in a shared space, opening multiple possibilities for interaction.

Virtual environments (VEs) can be experienced non-immersively using a desktop, or immersively using a head-mounted display (HMD) or Cave (CAVE™is a trademark of the University of Illinois at Chicago, but the term "Cave" is used here to describe the generic technology as described in [1] rather than to the specific commercial product). Non-immersive desktop VEs can suffer from the same limitations in field of view as videoconferences. Immersive VEs (IVEs), however, combine stereoscopic images with head-tracking to produce a sense of being surrounded by the virtual world [2]. In IVEs, avatars representing interaction partners are experienced not as 2D images on a screen, but as life-size, 3D entities occupying a shared, surrounding mediated space (figure 2-1).

CVEs have several properties that make them suited to group interaction. They are:

- *multi-user*, supporting multiple, geographically dispersed users;
- *synchronous*, enabling people to interact with each other in real time;
- *navigable*, allowing users to freely navigate the 3D space;
- *embodied*, representing users by digital proxies called "avatars";
- *spatial*, providing a shared 3D interaction context.

It is their inherent *spatiality* that sets CVEs apart from other groupware systems such as video-mediated communication (VMC) and media spaces.

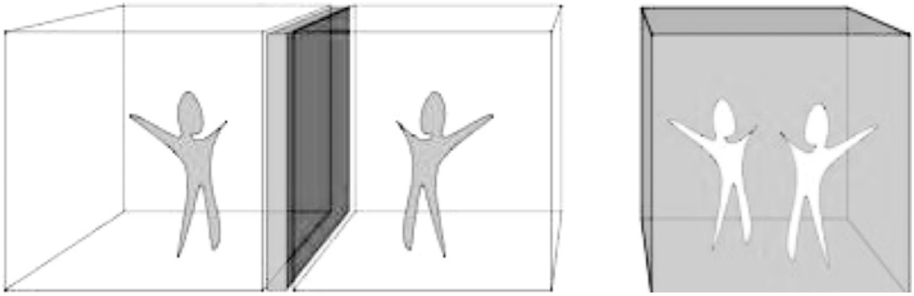


Figure 2-1. Using groupware systems such as VMC, people remain in separate physical contexts and interact with each other via video projection. Using IVEs, people interact in a shared, computer-generated 3D context where they are represented by digital proxies called "avatars".

Though media spaces enable people to share visual information from their physical environment [3], they fail to preserve the spatial context of each user's physical environment [4]. The portrayal of space in CVEs has two practical advantages for remote collaboration: the provision of a shared interaction context for geographically dispersed users, and the portrayal of directed attention.

While it is not the aim of this chapter to compare the relative merits of video and avatar-mediated communication, three key distinctions help to highlight some potential strengths of CVEs as a medium (figure 2-2). Videoconferencing portrays participants' real appearance and actions as well as views of their real environment, and is therefore high in *fidelity*; however, it is experienced on a 2D screen and is therefore low in *spatiality* and *immersiveness*. Conversely, IVEs provide a 3D surrounding experience and are high in *spatiality* and *immersiveness*. However, they are lower in *fidelity* because they portray artificial, computer-generated scenes as opposed to real scenes captured from the physical world. In the context of group interaction, the degree of fidelity of a CVE hinges on its capacity to portray a convincing context and process for collaboration. The ambiguous relationship between an avatar and the person represented therefore poses complex challenges in terms of creating expressive embodiments that contribute meaningfully to the ongoing interaction. One key aim of CVE research is to increase fidelity with a view to bridging the gap between virtual and face-to-face interaction.

2. The Need for Avatar Fidelity: Goals for Expressive Avatars

One of the underlying assumptions behind research in both VMC and CVEs has been that the inclusion of visual information can improve mediated

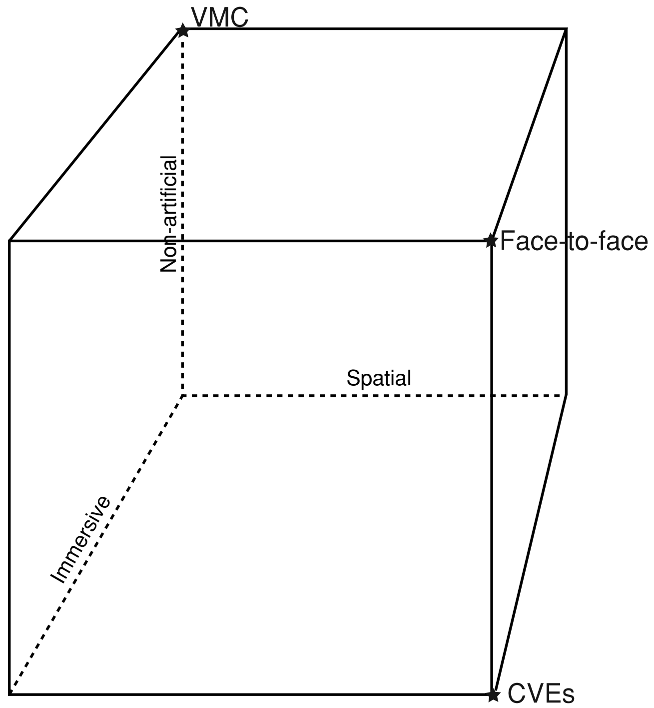


Figure 2-2. Comparison between VMC and IVEs along the dimensions of fidelity, spatiality, and immersion.

interaction by harnessing our natural ability to read meaning into the human form. Short, Williams and Christie have argued that all attempts at producing visual communications media are "primarily directed at remedying what is the most obvious defect of the simple telephone—the fact that one cannot see the other person or group" [5, p. 43]. The question that arises with the advent of CVEs is, what happens when both the environment and the people in it are not portrayals of the real world, but computer-generated representations? One significant barrier to interaction in current CVEs is in the paucity of avatar expression compared with live video of real people. One of the challenges in developing CVEs as a communications medium is therefore the creation of expressive avatars that enrich, rather than hinder, communication between remote participants.

In face-to-face interaction people rely heavily on nonverbal cues such as eye gaze, facial expression, posture, gesture and interpersonal distance to supplement the verbal content of conversation [6]. Indeed some argue that nonverbal signals not only constitute a separate channel of communication, but that they often override verbal content [7]; in other words "how" something is said can be more important than "what" is said.

Nonverbal behaviours serve at least two central functions in face-to-face interaction: conversation management and the communication of emotion. Conversation management concerns the use of paralinguistic cues to ensure the smooth flow of conversation. Movements such as eyebrow raises, head nods and posture shifts give structure and rhythm to the conversation and are essential to maintaining a sense of mutual understanding. The communication of emotion is itself integral to the regulation of communication and interaction [8, 9]. Picard explains that in addition to enriching the quality of interaction, emotion is crucial in the communication of understanding, and speakers continually monitor listeners' body language and facial expression for confirmation that they are being understood [8].

Given the central function played by nonverbal behaviours in face-to-face conversation, avatars' ability to convey such nonverbal cues is likely to affect how they are perceived as well as their contribution to social interaction. In works of cyberfiction such as Neal Stephenson's *Snow Crash* [10], avatars are both highly photorealistic and expressive. They perform seamlessly in real time, and are so reliable in conveying intended behaviour that businessmen happily substitute face-to-face meetings with interactions in the "Metaverse". In comparison, avatars in today's CVEs are extremely limited in their expressive potential.

3. Constraints on Avatar Fidelity

There are key technical constraints and theoretical concerns affecting the degree of avatar fidelity possible in current CVEs. The first consideration, in terms of the avatar's static appearance (visual fidelity), is the tension between *realism* and *real time*. The second, in terms of its dynamic animation (behavioural fidelity), is the tension between *control* and *cognitive* load.

3.1. The Tension between Realism and Real Time

Visual fidelity concerns not only the avatar's morphology and level of photorealism, but also the degree to which it resembles the person represented (referred to by Benford *et al.* as "truthfulness" [11]). Figure 2-3 illustrates three key dimensions of visual fidelity.

This chapter is concerned exclusively with humanoid avatars, and the issue of "truthfulness" is beyond the scope of the present discussion. For simplicity, visual fidelity will refer here to the avatar's level of photorealism. Typically, avatars used for communication purposes are relatively cartoonish. Cheng *et al.* [12] suggest that this may be partly dictated by user preference. However, restrictions related to rendering and bandwidth also mean that there is a tension

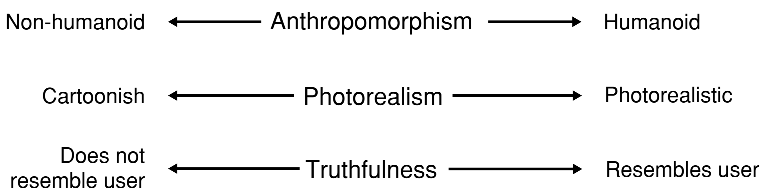


Figure 2-3. The dimensions of visual fidelity include anthropomorphism, photorealism and truthfulness.

between real-time performance and the level of realism achievable. Increased photorealism introduces computational complexity, resulting in significant and unwanted delays to real-time communication. Morningstar and Farmer cite this as a particular concern in the design of graphical chats [13]; for the same performance-related reasons, Hindmarsh *et al.* advocate using recognisable but simplistic humanoid avatars for small group communication purposes [3].

3.2. The Tension between Control and Cognitive Load

Being computer-generated, avatars afford control not only over appearance but also over behavioural expression, thereby potentially avoiding the pitfalls of nonverbal leakage that can occur in both face-to-face and video-mediated communication. However, avatars in existing graphical chats have been widely critiqued for their insufficient and sometimes misleading behaviours [14].

Avatar behaviours can be driven in a variety of ways. Manual driving through menu selection, mouse movement, pen gesture [15] and hand gesture [16] afford control over the avatar's actions but require continuous attendance to its state. Several alternative approaches have been proposed in response to the problem of enriching avatar communication while reducing cognitive load. Cuddihy and Walters [17] suggest a solution involving high-level control through a dynamic interface that clarifies what actions are available to users at any given time. This would make it possible to direct a "waving" action at an approaching avatar rather than manually orienting the avatar and then raising its arm, as was the case in Slater *et al.*'s acting rehearsal experiment [18]. A similar high-level approach is taken by Vilhjálmsson and Cassell in the BodyChat system [19]. Here, users choose whether to be available for conversation, and their avatars automate appropriate cues such as smiles, eyebrow raises and glances to indicate a willingness to approach or depart. Analogously, Tromp and Snowdon suggest the use of automated behaviours to enhance group interaction, for instance locking gaze to the speaking avatar to denote attention [20]. However, the drawback is that automation may result in misleading behaviours.

A radically different approach involves mapping the person's real-life expression onto the avatar's. Durlach and Slater indicate two possible approaches: the use of "direct, pass-through video of the participants" [21, p. 216], or using tracking data to manipulate the avatar's 3D mesh. Body and facial tracking makes it possible to animate an avatar using motion data from a real person. Tracking equipment can, however, be expensive as well as intrusive for users. On a theoretical level, it is also questionable whether full tracking will be desirable in a medium that is prized for the control it offers users over their own embodiment.

Overall, there are significant challenges in driving appropriate behaviours for avatars. In addition to technical challenges, there remain open questions about the appropriateness of tracking or automating behaviours in the quest to reduce cognitive load without sacrificing users' control over avatar actions.

4. Setting Priorities: The Trade-off between Visual and Behavioural Fidelity

Combined, these technical and theoretical concerns mean there is a need to make trade-offs and establish priorities for avatar fidelity. Fraser *et al.* have stated that many designers of CVEs and virtual characters operate on the premise that more realistic environments and avatars should result in qualitatively better experiences in CVEs: "virtual environments—models, avatars, interfaces and so on—are often designed with realism in mind" [22, p. 30].

The need for literal portrayals in VEs is, however, a matter of debate. As Zeltzer argues, given current technical limitations, the priority is to develop selective fidelity based on contextual needs, and further research is needed to understand how to measure selective fidelity. Similarly, Fraser *et al.* propose a shift in priorities away from literalism and realism, particularly given the crudeness of current interfaces for conveying human movement [22]. Benford *et al.* argue that improving avatar expressiveness necessarily involves compromises [23], later adding that the streamlining of avatars and the use of more "abstract" approaches to their design may be more appropriate [11]. They therefore advocate incremental, context-driven improvements to fidelity rather than an absolutist drive towards photorealism.

Several authors share the assumption that rather than attempting to maximize realism, the priority is to focus on improving behavioural fidelity for communication purposes. For instance, Sallnäs argues that in collaborative tasks realistic appearance is secondary to the support of body positioning, pointing and object manipulation [24] . Similarly, Swinth and Blascovich reason that both anthropomorphism and photorealism are separate from, and secondary to, behavioural realism, which they define as "the extent to which avatars and other objects in an virtual environment behave like their counterparts in the physical world" [25, p. 329].

The assumption that visual fidelity is secondary to behavioural fidelity is partly supported by lessons from animation. Disney animators translated films of actors' body language and facial expression into simple line drawings and discovered it was possible to achieve effective emotional portrayals in visually simplistic characters, provided the movement was convincing [26]. More recently, Katsikitis and Innes' study on line drawings of a smile illustrated that even a cartoonish representation of an expression can be decoded accurately down to its five phases of development [27].

Recent studies on the transmission of nonverbal cues in mediated communication add further support to the argument favouring behavioural fidelity. Ehrlich *et al.* [28] point out that the same bandwidth restrictions constraining CVEs also apply to VMC. They suggest that the standard approach of preserving spatial and colour resolution at the expense of temporal degradation is counterproductive. Their experimental findings indicate that preserving motion information is critical to the recognition of facial expression and may compensate for significant losses in image resolution.

Considering that the transmission of nonverbal cues can be severely affected by temporal delays and inconsistencies, they suggest that "if a bandwidth trade-off is required, one should consider preserving high-fidelity motion information at the expense of image realism, not the other way around" [28, p. 252]. In a separate study on facial affect recognition, Schiano, Ehrlich, Krisnawan, and Sheridan [29] compared a low-fidelity robot enacting the six "basic" emotions with video of human actors enacting the same emotions. Though scores for the robot were lower, the expressions were decoded in a pattern that closely followed the human faces. This further supports the argument prioritising behaviour over accurate appearance in the transmission of nonverbal cues.

Bente and Kramer [30] describe a related study on person perception, this time comparing silent video clips of dyadic interactions between human actors with equivalent clips of identically animated computer-generated agents. Their findings indicate a remarkable correspondence in responses to both conditions, despite the lower-fidelity appearance of the agents. In summary, technical limitations have forced the need to set priorities in avatar design. Findings from different media experiences partially support the notion that behavioural fidelity may be more pressing than visual fidelity for communication purposes.

5. Exploring the Impact of Minimal Fidelity

The argument for exploring the lower boundaries of fidelity is not born exclusively out of technical necessity. Reeves and Nass [31] document a series of studies suggesting that people respond to media as social actors, and tend to anthropomorphise even the simplest of text-based interfaces. This theory of the “medium as social actor” is of direct interest to avatar design because it suggests that minimal cues can elicit social responses.

Biocca, Harms and Burgoon [32] maintain that interaction in CVEs may be built on minimal cues because the automatic interpretation of humanoid forms and nonverbal behaviour can lead people to attribute a degree of sentience to virtual humans. This tension between automatic social responses and the rational knowledge that virtual humans are artificial entities represents a fundamental and engaging issue that has been addressed in a selection of studies in different research institutions.

Studies on fear of public speaking [33,34] and spatial interaction with humanoid agents [35] support this notion that people can respond socially to virtual humans even in the absence of two-way verbal interaction, and despite knowing rationally that they are not “real”. In our research we sought to explore the impact of minimal fidelity on communication experiences in CVEs, investigating one key behaviour, eye gaze, in the context of dyadic interaction.

6. Experiments on Eye Gaze and Photorealism

One of the central problems in mediated communication is the portrayal of directed attention. The advantage of CVEs is that participants’ embodiments can be seen in spatial relation to each other and to the objects they are interacting with. Unlike videoconferencing and media spaces where camera positions are fixed, participants in CVEs are free to control their point of view (POV) by navigating through the environment. As Bowers, Pycock and O’Brien point out [36], this alone allows a degree of awareness of the others’ focus of attention. However, the granularity of this understanding depends largely on the fidelity of the embodiment, on its level of visual detail (photorealism) and behavioural accuracy. There are significant challenges involved in portraying accurate eye gaze in CVEs, particularly in an immersive setting where participants’ faces are partially obscured by stereoscopic goggles, making tracking more problematic.

Gaze is a richly informative behaviour in face-to-face interaction. It serves at least five distinct communicative functions [37, 6]: regulating conversation flow, providing feedback, communicating emotional information and the nature of interpersonal relationships, and avoiding distraction by restricting visual input. Research on gaze in mediated communication has been concerned mainly with issues of conversation management in multiparty interaction. One of the perceived limitations of telephony-based videoconferencing systems is that they do not support selective gaze [38–40]. Various media space systems have attempted to address this limitation by distributing individual audiovisual units in physical space to represent each user (see [40] for a review).

Studies in CVEs have attempted to address the problem of how to support selective gaze in multiparty interaction within a shared 3D space. The GAZE groupware system [40] is designed to ease turntaking by conveying gaze direction in a shared virtual space using VRML2. This system uses an advanced desk-mounted eye-tracking system to measure where each person is looking. The gaze information is then represented metaphorically in the form of a 2D texture-mapped "persona" that moves about its own x- and y-axes in the 3D environment.

Taylor and Rowe [4] argue that the GAZE groupware system is problematic for two reasons. First, using a snapshot instead of video precludes any possibility of expressing other nonverbal cues through the persona. Second, the use of a plane makes it difficult to generate the kinds of profile views useful in multiparty communication. They address these limitations by rendering video of the facial region on a generic 3D model of a face. Their system animates the head movement by tracking the two earphones and microphone to obtain head position information for each user. The eye movement is contained in the video image. Their system renders avatars from an asymmetric viewpoint that corresponds to the position of the real participant, who typically sits 20 inches away from a 14-inch desktop screen. They conclude that this system improves group interaction by preserving the semantic significance of gaze. However, integrating video as a part of gaze animation fails to address the needs of users who prefer to remain visually anonymous behind a synthetic avatar.

Both of the above systems are concerned with supporting selective gaze in groups of three or more. In terms of two-person (dyadic) communication, Colburn, Cohen and Drucker [41] present findings from an experiment comparing visual attention to the screen during 20 conversations using an avatar. Participants were presented with three 3-minute visual stimuli in random order: a blank screen, a fixed-gaze avatar and an avatar with a functioning eye gaze model, based on who was speaking and whether or not the participant was looking at the screen. Participants looked at the screen more when the avatar was present and most of all when the gaze model was active. The experiments presented in the remainder of this chapter extend this research by investigating the impact of eye animations on a range of subjective responses, including perceived quality of communication and copresence.

6.1. Behavioural Fidelity: Exploring the Impact of Eye Gaze

Our first 100-person experiment was designed to investigate the importance of eye gaze in humanoid avatars representing people engaged in conversation. The experiment was conducted using a video-tunnel setup and was therefore not immersive (figure 2-4); this was done deliberately to isolate gaze behaviour from any other factors, such as spatial, gestural or postural cues

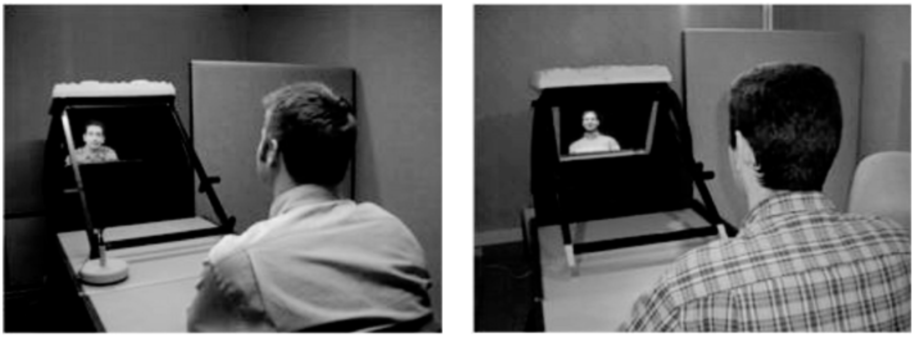


Figure 2-4. Participants in the video condition speaking with each other via the video-tunnel.

that might have confounded results. In the avatar condition, participants saw a face-on head-and-shoulders view of the avatar representing their conversation partner.

We compared responses to dyadic conversation in four mediated conditions. An avatar with "random" head and eye movements was compared to a visually identical "inferred-gaze" avatar that combined simple head-tracking with "while speaking" and "while listening" eye animations inferred from the audio stream. The design of these eye animations was informed by social psychology research on the differences in gaze patterns while speaking and while listening in face-to-face interaction [37, 42, 6]. Both avatar conditions were then compared to video (with audio) and audio-only baseline conditions. The impact of each condition on the perceived quality of communication was assessed by comparing participants' subjective responses along four dimensions: how natural the conversation felt, their degree of involvement in the conversation, their sense of copresence, and positive or negative evaluation of the conversation partner.

The goal of the experiment was two-fold: firstly, to test whether an avatar with minimal behavioural fidelity could contribute to the perceived quality of communication between two remote users. The second, more specific goal was to examine the role of gaze: when the avatar's gaze was directly related to the conversation, would this improve the quality of communication compared to a visually identical avatar with random gaze?

The perception of eye gaze depends on a combination of head and eye orientation [43, 6]. In the random-gaze condition, both the head and eye movements were designed to appear natural but were in no way tied to the content or flow of conversation. For the inferred-gaze avatar, the head movements were tracked using a single Polhemus sensor attached to the headphones. The eye movements were inferred from the audio stream. One of the fundamental rules for gaze behaviour in face-to-face dyadic interaction is that people gaze at their communication partner more while listening than while speaking [44, 42, 6].

Figure 2-5. Male and female avatars looking “at” and “away”.

Drawing on this principle, “while speaking” and “while listening” eye animations were implemented based on timing and frequency information taken from face-to-face dyadic studies. Figure 2-5 shows the male and female avatars in “at” and “away” gaze.

In order to assess the avatar’s impact on perceived quality of communication, a task was needed in which participants would be sensitive to visual feedback. It has been suggested [45, 39] that users benefit most from having visual feedback when performing equivocal tasks that have no single “correct” outcome but require negotiation. Short, Williams and Christie [5] argue that tasks involving conflict and negotiation are particularly suited to testing perceptions of communications media. A role-playing negotiation task was developed specifically for the study, requiring participants to come to a mutually acceptable agreement to avoid a family scandal breaking out in a small town. It was also thought that the emotional content of the scenario combined with the negotiation requirements of the task would mean that results could speak both to social and business contexts.

Participants’ responses were elicited by means of the post-experiment questionnaire, each response being on a 9-point Likert scale.

The results showed that the random-gaze avatar did not provide a significant improvement over pure audio, suggesting that the simple introduction of an avatar does not automatically improve participants’ perception of communication. Rather, the avatar must have certain behaviour characteristics in order to be useful. The inferred-gaze avatar outperformed the random-gaze avatar and the audio-only condition on several response measures. This suggests that an avatar whose behaviours reflect an aspect of conversational flow can indeed make a contribution to improving remote communication. Finally, the inferred-gaze avatar significantly outperformed the random-gaze avatar on all measures, indicating that an avatar whose behaviours are related to the conversation can present a marked improvement over an avatar that merely exhibits liveliness.

These findings had encouraging implications for inexpensive approaches to improving avatar fidelity. However, a central question remained unanswered. In the inferred-gaze condition, the avatar's gaze behaviour was being driven by two separate channels of information: its eye movement was based on inference from the audio stream, while its head movement was based on tracking the participant's real head movement. The open question was whether the significant impact on participants' perceptions was due to tracked motion data or from inferences about the eye movement based on research from face-to-face interaction. Answering this question would have significant implications for providing inexpensive ways to improve eye gaze based on information readily available from the audio stream.

Moreover, the experiment was conducted in a non-immersive setting. The question remained of how the inferred-gaze model would perform in a more demanding immersive setting, where participants were free to wander about a shared 3D space. A second experiment was therefore designed to address these concerns.

Between the publication of results from the first eye gaze experiment [46] and the second experiment [47], Lee, Badler and Badler published a similar study comparing subjective responses to a humanoid agent with static, random and inferred gaze [48]. Their agents' inferred-gaze animations were consistent with the timings from the face-to-face literature detailed above, but were refined using a statistical model developed from their own gaze tracking analysis of real people engaged in dyadic interaction. Their results from a 12-person evaluation are consistent with those from our first study, in that the inferred-gaze model results in more positive perceptions. The inferred-gaze agent significantly outperforms the visually identical random-gaze agent in terms of perceived interest, engagement, friendliness, and liveliness. However, it is not clear whether participants were engaged in two-way verbal communication with the agent, or whether they simply viewed the animations on a screen.

In terms of eye gaze and photorealism, two studies by Fukayama *et al.* are also directly relevant [50]. The first is a 13-person study concerning the impact of eye animations on the impressions participants formed of an interface agent [50]. Their gaze model consists of three parameters: amount of gaze, mean duration of gaze and gaze points while averted. Their comparative analysis of responses to nine different gaze patterns suggests that agent gaze can reliably influence impression formation. For this particular study they isolated the agent's eyes from any other facial geometry. In a related study, they investigate whether the impact of the gaze patterns is affected by the photorealism of the agent's face [49]. Their findings suggest that varying the appearance from visually simplistic to more realistic has no effect on the impressions produced. The interaction is one-way, with participants viewing a pre-recorded agent animation. It is therefore difficult to know whether the findings would generalise to a sustained verbal interaction.

6.2. Visual and Behavioural Fidelity: Exploring the Impact of Eye Gaze and Photorealism

One aspect of the studies described above was that participants were shown a limited, head-and-shoulders view of the virtual human, and that the spatial relationship was fixed by the 2D nature of the interaction. They therefore left open the question of how these gaze models might hold up in an immersive situation where participants are able to wander freely around a shared space, and where they can interact with a full-body, life-size avatar. Our follow-up experiment was designed with these questions in mind.

The goal for this second experiment was threefold. Firstly, to disambiguate between the effect of inferred eye movements and head-tracking, both of which may have contributed to the results reported in the first study. Secondly, to test how the inferred-gaze model performs in a less forgiving immersive setting where participants were free navigate in the 3D IVE. Finally, to explore the relative impact of two logically distinct aspects of avatar fidelity: appearance and behaviour.

As previously discussed, one assumption made by several researchers is that convincing behaviour is a higher priority than realistic appearance in the development of expressive avatars. We wished to test this assumption by investigating the impact of the (higher-fidelity) inferred-gaze model with the (lower-fidelity) random-gaze model on avatars whose appearance represented different levels of photorealism. The initial hypothesis was that behavioural realism would be independent in its effects on perceived quality of communication from the impact of visual realism, and that it would be of greater importance. The inferred-gaze model was expected to outperform the random-gaze one for both the higher-realism and lower-realism avatar. One open question concerned the extent to which the gaze animations would impact on the lower-realism avatars, or how the two avatars would perform in comparison with each other.

Participants were represented to their conversation partner as a life-size avatar, as illustrated in figure 2-6. Both participants in each pair were represented by a visually identical avatar to avoid differences in facial geometry affecting the impact of the animations.

Since one of the central aims of this experiment was to disambiguate the impact of head-tracking and the inferred eye animations, participants' heads were tracked in all conditions, and only the eye animations were varied. The

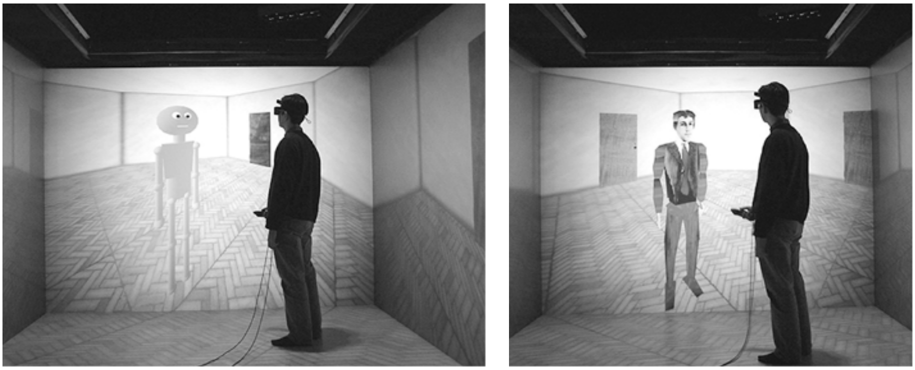


Figure 2-6. Participants saw their conversation partner as a life-size avatar. The avatar was either lower realism (left) or higher realism (right).

random-gaze eye animations were identical to the first eye gaze experiment, but the inferred-gaze animations were refined based on newly published information [48].

Our previous four indicators of perceived communication quality of communication were considered in the analysis, along with a number of additional responses including:

- *Gaze fidelity*: The sense of mutual gaze with the conversation partner;
- *Avatar fidelity*: The degree to which the avatar's appearance and behaviour were seen to be realistic;
- *Social-copresence*: Consisting of the following subcomponents:
- *General copresence*: The sense of being "in the company of" another person;
- *Spatial copresence*: The sense of being in the same space as the conversation partner;
- *Personal contact*: The degree of personal contact experienced with the partner.

Again, participants' subjective responses were elicited by means of a post-experience questionnaire. Our analysis revealed a very strong interaction effect between the *type of avatar* and the *type of gaze*. In other words, the impact of the gaze model is different depending on which type of avatar is used. For the higher-realism avatar the (more realistic) inferred-gaze behaviour *increases* perceived effectiveness for several response measures. For the lower-realism avatar, the (more realistic) inferred-gaze behaviour *reduces* effectiveness. This seems to indicate a need for consistency between the visual appearance of the avatar and the type of behaviour that it exhibits. With respect to eye gaze, low-fidelity appearance demands low-fidelity behaviour, and correspondingly higher-fidelity appearance demands a more realistic behaviour model.

The findings clear up the ambiguity from the first experiment regarding whether the significant differences in performance between the gaze models were due to head-tracking or eye animations inferred from the audio stream. They indicate that independent of head-tracking, inferred eye animations can have a significant positive effect on participants' responses to an immersive interaction. The caveat is that there should be some consistency between visual and behavioural realism, since the lower-realism avatar did not appear to benefit from the higher-realism, inferred-gaze model. This finding has implications for inexpensive ways of improving avatar expressiveness using information from the audio stream and suggests avenues for interim solutions for the difficult problem of providing robust eyetracking in Cave-like systems.

Findings from an in-depth qualitative analysis of the interviews from the first eye gaze study indicate that avatar fidelity does not work in isolation in shaping interaction experiences. Both communication context and personal characteristics such as everyday social anxiety, prior media experiences, and technical expertise shape perceptions of the avatar's role in interaction. The quantitative findings suggest that simply adding "liveliness" to the avatar's behavioural repertoire does not add value to the interaction. They further suggest that users would benefit from information about how "truthful" the animations are, because in the absence of priming, people are likely to interpret the significance of the avatar's behaviour according to their own technical assumptions about how it is driven. These assumptions are sometimes illogical and uninformed, and may work to the detriment of the avatar by leading users to discard as insignificant even those selected behaviours that are in fact informative. Overall this analysis signals the importance of educating users about the behavioural capabilities of the avatar.

Interviews with participants indicated that it is possible to rationally think of avatars as computer-generated and therefore not "real", while simultaneously exhibiting social responses towards them. This is consistent with Reeves and Nass' [31] theory of the medium as social actor that predicts people will tend to anthropomorphise media and treat them as social entities. This finding also supports Blascovich and colleagues' hypothesis that there are at least two levels of response to virtual humans: higher-level rational responses, and lower-level involuntary responses [51]. This opens avenues for future research on the complex multi-level responses of people to virtual humans in the context of small group communication in CVEs.

7. Conclusions and Future Work

One key barrier to effective communication in current CVEs is the relative paucity of avatar expressiveness as compared to live video. Increasing the expressive potential of avatars involves significant challenges. In terms of their appearance, the tension between realism and real time means that photorealism comes at the expense of unwanted delays to real-time communication. Visual fidelity must therefore be traded-off against available computing resources. In terms of behaviour, the tension between control and cognitive load underlines the difficulty of transparently driving avatar behaviours that appropriately represent the user. Full manual control of avatar behaviour would entail an unacceptable level of cognitive load. On the other hand, full tracking can be expensive and invasive, and may not be desirable in a medium that is prized for the control it affords over visual identity.

Given these constraints, the approach taken in our research was to explore the lower boundaries of avatar fidelity. The overarching goal was to investigate whether minimal increments in fidelity could contribute to participants' interaction experience in CVEs. Fidelity was considered in terms of both dynamic behaviour (behavioural fidelity) and static appearance (visual fidelity).

The findings from the first experiment on eye gaze indicate that an avatar with minimal fidelity can make a positive contribution to interaction. However, simply adding "liveliness" to its behavioural repertoire does not add value to the interaction. In the case of gaze, the avatar's animation should reflect some aspect of the ongoing conversation, even something as simple as turntaking. Concerning the relationship between visual and behavioural fidelity, we discovered a significant and overwhelmingly consistent interaction effect between visual and behavioural realism. The findings from the experiment on eye gaze and photorealism indicate that the effect of identical eye animations changes in relation to the avatar's appearance. The higher-realism avatar benefited from the higher-fidelity inferred-gaze animations, whereas the opposite was true for the lower-realism avatar. This suggests the need to align behavioural fidelity with avatar appearance. The conclusion is that the impact of behaviour is not independent of appearance and points to a more complex picture than was previously envisaged. Simply privileging behavioural over visual fidelity may not necessarily lead to optimal improvements for expressive avatars.

Bowers, Pycock and O'Brien suggest that "a viable and systematic research strategy for developing useful CVEs is to incrementally add further sophistication to very simple embodiments *as and when* analysis reveals that it is called for in the support of social interaction" [36, p. 65]. Varying individual dimensions allows researchers to explore their individual impact on the perceptual and social impact of avatars. This is a logical approach given the need to prioritise which aspects of fidelity might be traded-off against available computing resources.

Our experiments attempted to focus on the impact of a single nonverbal behaviour. We chose eye gaze because of its central function both in conversation management and in the directed expression of emotion. The isolation of a single behaviour is potentially problematic, however, because it compromises the "gestalt" of nonverbal expression that characterises face-to-face interaction.

Short, Williams and Christie point out the danger of treating behaviours individually: "In attempting to assess the functions of the visual channel, it is dangerous to confine attention at any one time to individual cues such as posture, eye-gaze, proximity and the like. The channels do interact. . . . Studies of media must look at relevant combinations of channels. Important overall properties of communication may be missed if attention is restricted to individual channels" [5, p. 7–8]. Additional behaviours should therefore be investigated in conjunction with gaze and with each other, with a view to exploring their interdependencies.

Our research focused on participants' subjective responses to their interaction experience. Although interviews and in-depth qualitative analysis go a step further than questionnaires in understanding users' perceptions, they necessarily only capture the rational level of experience. The findings reported here indicate that it is important to further explore both higher-level and lower-level social responses to virtual humans. Previous research has indicated that minimal fidelity can affect lower-level involuntary responses such as spatial behaviour in response to an agent [35]. Potentially fruitful avenues for research include the observation of involuntary behaviours, as well as the use of psychophysiological measures to study objective responses. Future work will build on these findings by combining subjective and objective approaches to understand how avatars can be further improved for richer multiparty interaction in CVEs.

There are numerous application areas for VEs, from simulation to training to the treatment of phobias. At present, CVEs are primarily used for research and entertainment purposes and have yet to come into mainstream use as a communication medium. They have the potential to extend beyond their present usage to address the practical collaborative needs of geographically dispersed users. One of their chief attractions lies in their ability to combine 3D spatial interaction with a high degree of multisensory immersion. In *Simulacra and Simulation* [52], Baudrillard argues that science fiction is nothing more than an exaggeration of the possibilities inherent in present circumstances. If recent films such as the Matrix [53] and works of cyberpunk literature [10] are anything to go by, then we face a future where communication in CVEs will become part of the fabric of our everyday lives. Avatars, the visual representations of people in CVEs, therefore deserve careful consideration as they will play a pivotal role in enriching the communicative possibilities of this medium.

It is possible to imagine a day when avatars will, like those in the novel *Snow Crash* [10], communicate users' intentions so reliably that people will willingly use CVEs for social interaction and for serious collaborative purposes. In 1998 Allbeck and Badler argued that every aspect of avatar design, motion and appearance described in Snow Crash had already begun to be tackled by different research groups throughout the world [54]. It is encouraging to think that all of these various branches of research may converge to create truly compelling communicative avatars.

Avatars are computer-generated and therefore, unlike video, free us from the need to present a faithful visual replica of real places or real people. Like masks, they preserve our visual anonymity and open up the possibility for new and potentially different forms of interaction. Existing technical limitations have meant that it is not currently possible to model human appearance and behaviour in all its complexity for real-time interaction. This has dictated a need to explore minimal and selective fidelity with a view to gradually increasing the expressive potential of avatars.

The ability of humans to decode caricature and cartoons indicates that we do not require exhaustive photorealistic depictions to decipher the human form. The research and ideas discussed in this chapter have rested on the assumption that our common goal is to enhance avatar animation to harness our natural ability to decode nonverbal behaviours. It is conceivable, however, that a move away from photorealism might be accompanied by a parallel move away from behavioural literalism. Just as the lack of sound in silent films made actors instinctively "turn up the volume" on their visual performances, perhaps current constraints will lead to ways of "emoting" through avatars that do not precisely mirror everyday nonverbal communication. Freed from the need to mimic the real world, avatars can distort, change colour and morph into new forms to express emotional states in non-literal ways. As the medium matures it will be interesting to see how current constraints will give rise to creative emergent solutions, and how people will choose to use their avatars to express themselves in new and possibly more magical ways.

Acknowledgements

My warmest thanks to my advisors Mel Slater, Angela Sasse and Simon Bee for their precious guidance over the past few years. Special thanks to my collaborators at University College London (UCL) for making the research possible: David-Paul Pertaub, Anthony Steed, Andrea Brogni, Vinoba Vinayagamoorthy, Pip Bull and David Swapp. I am also very grateful to Tim Child, Sanja Abbott and Marcus Tutt for giving me use of their Televirtual avatars for the first study on eye gaze, and to the people at BT Exact who helped in so many ways with the study: Dickon Povey, Alex Bourret, Michelle and Jim Tasker, Paul Bowman, Tim Stevens, Mike Hollier, Dan Ballin and Dan Argent.

References

1. Cruz-Neira, C., Sandin, D.J., & DeFanti, T.A. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*. Anaheim, CA: ACM Press, pp. 135–142.

2. Barfield, W. & Weghorst, S. (1993). The sense of presence within virtual environments: A conceptual framework. In G. Salvendy & M. Smith (Eds.), *Human Computer Interaction: Software and Hardware Interfaces.* Amsterdam: Elsevier, pp. 699–704.
3. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (1998). Fragmented interaction: Establishing mutual orientation in virtual environments. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. Seattle, WA: ACM Press, pp. 217–226.
4. Taylor, M.J. & Rowe, S.R. (2000). Gaze communication using semantically consistent spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. The Hague, The Netherlands: ACM Press, pp. 400–407.
5. Short, J., Williams, E., & Christie, B. (1976). *The Social Psychology of Telecommunications*. London: John Wiley & Sons.
6. Argyle, M. & Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge: Cambridge University Press.
7. Forgas, J.P. (1985). *Interpersonal Behaviour: The Psychology of Interpersonal Interaction*. Oxford: Pergamon Press.
8. Picard, R. (1997). *Affective Computing*. Cambridge, MA: MIT Press.
9. Goleman, D. (1996). *Emotional Intelligence*. London: Bloomsbury Publishing Plc.
10. Stephenson, N. (1992). *Snow Crash*. London: ROC.
11. Benford, S., Bowers, J., Fahlén, L.E., Greenhalgh, C., & Snowdon, D. (1995). User embodiment in collaborative virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Denver, CO: ACM Press, pp. 242–249.
12. Cheng, L., Farnham, S., & Stone, L. (2002). Lessons learned: Building and deploying shared virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 90–111.
13. Morningstar, C. & Farmer, R. (1990). The lessons of Lucasfilm's Habitat. In M. Benedikt (Ed.), *Cyberspace: First Steps*. Cambridge, MA: MIT Press, pp. 273–302.
14. Vilhjálmsson, H. & Cassell, J. (1999). Fully embodied conversational avatars: Making communicative behaviours autonomous. *Autonomous Agents and Multi-agent Systems, 2*: 45–64.
15. Barrientos, F. & Canny, J.F. (2002). Cursive: Controlling expressive avatar gesture using pen gesture. In *CVE '02: Proceedings of the 4th International Conference on Collaborative Virtual Environments*. Bonn, Germany: ACM Press, pp. 113–119.
16. Lee, C., Ghyme, S., Park, C., & Wohn, K. (1998). The control of avatar motion using hand gesture. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. Taipei, Taiwan: ACM Press, pp. 59–65.
17. Cuddihy, E. & Walters, D. (2000). Embodied interaction in social virtual environments. In *Proceedings of the Third International Conference on Collaborative Virtual Environments*. San Francisco, CA: ACM Press, pp. 181–188.
18. Slater, M., Howell, J., Steed, A., Pertaub, D-P., Garau, M., & Springel, S. (2000). Acting in virtual reality. In *Proceedings of the Third International Conference on Collaborative Virtual Environments*. San Francisco, CA: ACM Press, pp. 103–110.
19. Vilhjálmsson, H. & Cassell, J. (1998). Bodychat: Autonomous communicative behaviors in avatars. In *Proceedings of the 2nd Annual ACM International Conference on Autonomous Agents*. Minneapolis, MN: ACM Press, pp. 269–276.
20. Tromp, J. & Snowdon, D. (1997). Virtual body language: Providing appropriate user interfaces in collaborative virtual environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST)*. Lausanne, Switzerland: ACM Press, pp. 37–44.
21. Durlach, N. & Slater, M. (2000). Presence in shared virtual environments and virtual togetherness. *Presence: Teleoperators and Virtual Environments, 9*(2): 214–217.

22. Fraser, M., Glover, I., Vaghi, I., Benford, S., Greenhalgh, C., Hindmarsh, J., & Heath, C. (2000). Revealing the realities of collaborative virtual reality. In *Proceedings of the Third International Conference on Collaborative Virtual Environments*. San Francisco, CA: ACM Press, pp. 29–37.
23. Benford, S., Bowers, J., Fahlén, L.E., Mariani, J., & Rodden, T. (1994). Supporting cooperative work in virtual environments. *The Computer Journal, 37*(8): 653–668.
24. Sallnäs, E.L. (2002). Collaboration in multi-modal virtual worlds: Comparing touch, text, voice and video. In R. Schroeder (Ed.). *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 172–187.
25. Swinth, K. & Blascovich, J. (2002). Perceiving and responding to others: Human-human and human-computer social interaction in collaborative virtual environments. In *Fifth Annual International Workshop on Presence.* Porto, Portugal: ACM Press, pp. 310–340.
26. Thomas, F. & Johnston, O. (1981). *Disney Animation: The Illusion of Life*. New York: Abberville Press Publishers.
27. Katsikitis, M. & Innes, J.M. (1997). Encoding and decoding of facial expression. *The Journal of General Psychology, 124*(4): 357–370.
28. Ehrlich, S.M., Schiano, D., & Sheridan, K. (2000). Communicating facial affect: It's not the realism, it's the motion. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. The Hague, The Netherlands: ACM Press, pp. 251–252.
29. Schiano, D., Ehrlich, S.M., Krisnawan, R., & Sheridan, K. (2000). Face to interface: Facial affect in (hu)man and machine. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. The Hague, The Netherlands: ACM Press, pp. 193–200.
30. Bente, G. & Kramer, N. (2002). Virtual gestures: Analysing social presence effects of computer-mediated and computer-generated nonverbal behaviour. In *Fifth Annual International Workshop on Presence*. Porto, Portugal: ACM Press, pp. 233–244.
31. Reeves, B. & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places*. Cambridge: Cambridge University Press.
32. Biocca, F., Harms, C., & Burgoon, J.K. Criteria for a theory and measure of social presence. (Submitted). *Presence: Teleoperators and Virtual Environments*. Available at: http://www.mindlab.org/networkedminds/
33. Pertaub, D-P., Slater, M., & Barker, C. (2001a). An experiment on fear of public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments, 11*(1): 68–78.
34. Pertaub, D-P., Slater, M., & Barker, C. (2001b). An experiment on fear of public speaking in virtual reality. In D. Stredney, J.D. Westwood, G.T. Mogel and H.M. Hoffman (Eds.), *Medicine Meets Virtual Reality 2001: Outer Space, Inner Space, Virtual Space*. Newport Beach, CA: IOS Press, pp. 372–378.
35. Bailenson, J.N., Blascovich, J., Beall, A.C., & Loomis, J.M. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators and Virtual Environments, 10*(6): 583–598.
36. Bowers, J., Pycock, J., & O'Brien, J. (1996). Talk and embodiment in collaborative virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, Canada: ACM Press, pp. 58–65.
37. Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26*: 22–63.
38. Buxton, W. (1992). Telepresence: Integrating shared task and person spaces. In *Proceedings of the Conference on Graphics Interface*. Vancouver, Canada: ACM Press, pp. 123–129.
39. Sellen, A. (1995). Remote conversations: The effect of mediating talk with technology. *Human-Computer Interaction, 10*(4): 401–444.

40. Vertegaal, R. (1999). The GAZE groupware system: Mediating joint attention in multiparty communication and collaboration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Pittsburgh, PA: ACM Press, pp. 294–301.
41. Colburn, A., Cohen, M., & Drucker, S. (2000). The role of eye gaze in avatar-mediated conversational interfaces. Technical Report MSR-TR-2000-81, Microsoft Research.
42. Argyle, M. & Ingham, R. (1972). Mutual gaze and proximity. *Semiotica, 6*: 32–49.
43. Gibson, J.J. & Pick, A. (1963). The perception of another person's looking behaviour. *American Journal of Psychology, 46*: 386–394.
44. Argyle, M., Ingham, R., Alkema, F., & McCallin, M. (1973). The different functions of gaze. *Semiotica, 7*: 10–32.
45. Straus, S. & McGrath, J.E. (1994). Does the medium matter? The interaction of task and technology on group performance and member reactions. *Journal of Applied Psychology, 79*: 87–97.
46. Garau, M., Slater, M., Bee, S., & Sasse, A. (2001). The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Seattle, WA: ACM Press, pp. 309–316.
47. Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., & Sasse, A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Ft. Lauderdale, FL: ACM Press, pp. 529–536.
48. Lee, S.H., Badler, J.B., & Badler, N.I. (2002). Eyes alive. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*. San Antonio, TX: ACM Press, pp. 637–644.
49. Fukayama, A., Sawaki, M., Ohno, T., Murase, H., Hagita, N., & Mukawa, N. (2001). Expressing personality of interface agents by gaze. In *Proceedings of INTERACT Conference on Human-Computer Interaction*. Tokyo, Japan: IOS Press, pp. 793–794.
50. Fukayama, A., Takehiko, O., Mukawa, N., Sawaki, M., & Hagita, N. (2002). Messages embedded in gaze of interface agents: Impression management with agent's gaze. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Minneapolis, MN: ACM Press, pp. 41–48.
51. Blascovich, J. (2002). Social influence within immersive virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 127–145.
52. Baudrillard, J. (1994). *Simulacra and Simulation*. (Translated by S.F. Glaser *Simulacres et Simulation*. English). Ann Arbor: University of Michigan Press.
53. The Matrix. 16 mm, 136 min. Burbank: Warner Bros., 1999.
54. Allbeck, N. & Badler, N. (1998). Avatars à la Snow Crash. In *Proceedings of Computer Animation*. Philadelphia, PA: IEEE Computer Society, pp. 19–24.

Chapter 3

ANALYSIS AND VISUALIZATION OF SOCIAL DIFFUSION PATTERNS IN THREE-DIMENSIONAL VIRTUAL WORLDS

Shashikant Penumarthy and Katy Börner

1. Introduction

In order to indicate how difficult it is to orient oneself and navigate in virtual worlds, let us begin with an example—the story of John. Just having finished an hour of e-mail, John looks at his watch and realizes that it is time to switch identities. A few mouse clicks, the brief appearance of a splash screen and faster than one can say "Avatar!", the monitor screen fills up and one world is replaced by another.

Exit John, enter PringleCrow. As LinkWorld (one of several three-dimensional virtual worlds that John has access to) fades into view, PringleCrow waits to orient herself (today John has chosen a female avatar). She waits for her world to manifest itself, one object at a time. The network isn't very fast and PringleCrow is suddenly "lost", the connection to her life-force severed, as John's consciousness comes back to the real world wondering why the graphics are so jerky. Back again to PringleCrow, the life-force restored! This time, though, enough of the world has loaded ensuring that PringleCrow won't be lost again soon. An extension of consciousness takes over, as the human and his avatar are unified. Today, PringleCrow is going to be part of a treasure hunt that is designed to test how well the inhabitants of the virtual world understand the principles of physics. When she teleports to the meeting point for the participants, she realizes that the other avatars are already there, the chatter almost "deafening". She spots a chat utterance by CyberDog, her team-mate for the treasure hunt, who has been waiting for PringleCrow for a few minutes. A few more chat utterances later ("I am a blue puppy", "Let's meet near the waterfall", "It's at coordinates 10 west 56 north I think"), they finally meet. At this point neither avatar is aware of where the rest of the avatars are or

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 39–61.

even how many of them are in the world. They walk together to the podium where the avatar StrazyFoure, who is the moderator for the hunt, is reviewing the rules. After the rules are done, StrazyFoure raises his hand to signal the start of the treasure hunt and PringleCrow and CyberDog begin looking for clues. They are careful to communicate through "whispers", rather than chat to ensure that everything they say is "inaudible" to the avatars from competing teams.

At the end of the day, StrazyFoure goes to each team to find out what they think about the hunt and is a little concerned to discover that not everyone has succeeded to find all answers. While some found the clues very difficult to decipher, others complain that they sometimes could not figure out if they were going the right way. The virtual world is so big that some lost sense of how far they had come from the starting point.

At what point did the unsuccessful avatars get lost? Which clues were easy and which ones were difficult? Did the avatars interpret the clues correctly? What did the avatars talk about during the whole period? Could some of the clues have been placed in locations that were easier to access? These are just a few of the interesting questions that could be asked in order to determine how successful the treasure hunt was and how to support information foraging in 3D virtual worlds.

The avatars of the virtual world as well as researchers that study them need to be able to quickly obtain a broad overview of activities in the virtual world, as well as examine local details. Researchers must also be able to discover patterns of behavior, of movement, chat and interaction on a local and global scale. Some questions are short term such as "What paths did the avatars take?" or "Which avatars' movement trails deviated most from the average of the entire group?" Other questions are: "Which areas of the world are most used?", "What topics are the inhabitants talking about and where?", or "Who talked/interacted with whom?"

In this chapter we describe how advanced data mining and information visualization techniques [1–4] can be advantageously applied to augment, evaluate, optimize, and study collaborative 3D virtual worlds and to study their evolving communities [5]. We begin by discussing related work in the area of analysis and visualization of spatial and social data. After briefly commenting on the nature of virtual world data, we outline types of user groups and their tasks in virtual worlds. The use of patterns for analysis of virtual world behavior is detailed. We then show how techniques developed in areas as diverse as data analysis [6, 7] and information visualization can be combined to create simple yet powerful means for answering some of the questions posed above. A toolkit [8] for the analysis and visualization of social diffusion patterns in virtual worlds is presented. The chapter concludes with directions for future research in analysis and visualization in virtual worlds.

2. Related Work

Several highly diverse research areas have developed theories, techniques and systems for analysis and visualization of data possessing a distinct spatial component. The well-developed area of scientific visualization deals with applying computer graphics to scientific data "for purposes of gaining insight, testing hypothesis, and general elucidation" [9]. Scientific data usually takes the form of measured values obtained from a system or a phenomenon, in addition to three-dimensional coordinates associated with each measurement. Scientific visualization typically consists of data that can be accurately represented using three-dimensional geometry and is being used in such diverse areas as designing aircraft and constructing complete three-dimensional views of the human anatomy [10].

Another mature field that makes extensive use of visualization of spatial patterns is geography. Today's Geographic Information Systems (GIS) contain a large number of spatial analysis methods that enable one to analyze, model and query spatial data. Cartographers, in particular, have developed a number of techniques to visualize spatio-temporal diffusion patterns. Pioneering work by Dorigo and Tobler [11] represents and visualizes diffusion potentials and gradients as vector fields computed using a continuous spatial gravity model. By overlaying population density information over a geographic map, they have shown that New York, being densely populated, exhibits the highest outward pressure, i.e., it acts as a source, while Florida has the highest inward pressure, i.e., it acts as a sink.

The work that relates most closely to ours comes from the area of social visualization. Social visualizations are a special type of information visualization that focus on analysis of social behavior. For example, lifeline visualizations reveal migrations, transitions and trajectories of users or user groups [12]. Other research aims at the visualization of large-scale conversations, such as those that take place on the Usenet [13] or visualization of access patterns of users on the Web [14]. There also exists work on visualizing and supporting social interactions in text-based or 2D graphical systems. For example, Chat Circles [15] is a 2D graphical interface for synchronous conversation that visualizes the non-textual components of online chatting, such as pauses and turn-taking behavior, which are typically not available in chat log files. Another piece of work analyzes gestures and movement of users in VChat, a graphical chat system [16], in which a comparison is made between average distance and orientation of users in relation to users targeted in their chat and randomly selected users. Naper [17] was among the first to analyze chat text logged in a 3D virtual worlds and called for the use of visual semiotics for the analysis of computer mediated communication.

A very interesting body of work that analyzes social patterns exists in the area of urban studies. Whyte [18] determined the influence of steps, fountains, green spaces, sitting places, building arrangement, etc. on the crowd flow and social interaction in New York plaza. The company Space Syntax Limited (http://www.spacesyntax.com/) undertakes projects that aim to quantify the degree to which urban planning influences socio-economic factors such as pedestrian flow, crime patterns and land value. Results augment strategic design as well as the selection of design alternatives that best serve the needs of a particular segment of population.

Work on mapping MUDs and 3D virtual worlds is a relatively new area. The section on MUDs and Virtual Worlds in Dodge's and Kitchin's "Atlas of Cyberspaces" [19] provides a beautifully illustrated overview of this area. The well-known AlphaWorld Mapper (http://mapper.activeworlds.com/aw/) by Greg Roelofs and Pieter van der Meulen provides access to an impressive zoomable 2D map of a virtual world that is roughly the size of the state of California (429,025 km^2) [20]. However, the scale of hardware and other resources required to generate such a map is beyond the reach of most virtual world researchers. Until recently, it was not possible for researchers and users of virtual worlds to create a map of a virtual world or to analyze and visualize user interaction data collected in virtual worlds. The toolkit described in this chapter allows researchers to do this on a regular basis and in a consistent manner. The results have been shown to provide new insights about the evolution and usage of virtual worlds and the activities of their users [8, 21, 22].

3. Data in Virtual Worlds

It has been estimated that 80% of all data that exists in our world today has a spatial component [23]. This suggests that using spatial analysis and visualization techniques is essential in order to gain a fuller understanding of social patterns in real as well as virtual spaces. However, in the real world, collecting information about pedestrian flow, traffic or building utilization is difficult and often requires invasive and expensive methods such as the placement of sensors or cameras at various locations. The task of data collection is greatly simplified in virtual environments due to the ease with which user behavior can be recorded in a manner that does not interfere with the activities of inhabitants of the virtual world. The monitoring methods used in virtual worlds also have the inherent advantage that they result in data that is clean and consistent across time and space. The data therefore readily lends itself to analysis using data mining methods. In particular, the ActiveWorlds SDK (http://www.activeworlds.com/sdk/) offers ways to collect information such as movement, orientation, chat text, interactions such as clicks and even gestures using programmed software agents known as "bots". Information on the structure of virtual worlds is obtained using what are known as "propdump" and "registry" files, which respectively describe relative and absolute geometry of objects in the virtual world. The availability of such rich data eases the study of the interplay among space and social behavior. However, before one can begin addressing these questions, one must first identify the subjects of such a study, i.e., the user groups and the tasks that they perform within or related to the environment under consideration. This is the focus of the next section.

4. User Groups and Tasks

Our work in virtual worlds is embedded primarily in educational settings. Our users consist of kids from primary schools exploring information in the virtual world, performing scientific experiments, taking part in team-based games and engaging in other social activities—always with *learning* as the underlying goal. In this context, we aim to support three user groups:

Inhabitants: These are the core users who actually use the virtual world for activities such as exploring an art exhibition or participating in a treasure hunt. In our case, this usually consists of school kids who take part in the learning activity and their teachers who act as moderators or guides. This user group has all the needs and problems associated with using virtual worlds for collaborative work such as navigation, communication and coordination.

Designers: These are professionals who design the virtual world, set up the "infrastructure" and mentoring strategies to perform various kinds of activities. They also configure the environment so that it is suitable for diverse events to be held in the virtual world. This user group consists of technical designers who use computer and artistic skills to create the virtual world as well as teachers who aid the designers in creating an environment that supports a certain learning goal. This user group defines the overall design of the virtual world; placement of objects, locations of teleports, colors, textures, and behaviors of objects in response to user actions such as clicking or movement. This user group can benefit from answers to questions pertaining to patterns of usage in virtual worlds, such as the influence of the age of buildings in the virtual world on their usage, characteristics of objects that attract a high amount of interaction versus those that are ignored, etc.

Researchers: This user group is primarily concerned with asking and answering questions about social and technological aspects of virtual worlds such as "What are the effects of the environment on user behavior?", "What is the extent of influence of interface devices on the navigation capabilities of users?" and "What role does spatial reference play in the interaction between avatars?".

As is evident, each of these user groups has distinct roles to play either inside virtual worlds, outside or both. The specific nature of problems that each user group encounters varies widely. However, one can see that answering any of these questions needs information that synthesizes data about local and global patterns of activity. For example, in order to understand how often particular areas of the world are used, one needs to step back and look at the global pattern of density of usage of the virtual world. At the same time, to understand why such a pattern exists, one needs to zoom-in to specific regions of the world and examine interaction data or chat text to identify what the avatars in those regions have been exploring or talking about. Observations about such local and global activity can then be combined with the analysis of learning outcome but also pre and post-test questionnaires to obtain a comprehensive picture of a virtual world event.

5. Patterns of Social Activity

The term "pattern" can be defined as: a perceptual structure, a customary way of operation, a model worthy of imitation, a decorative or artistic work. Patterns occur in nature due a variety of physical, chemical and biological mechanisms. Many of these patterns manifest themselves as concrete physical phenomena and lend themselves to immediate perception in a sensory manner. At the same time, we find a number of patterns that are hidden, sometimes due to their existence on a plane other than the physical or sensorial, requiring a serious cognitive effort to uncover them, at other times due to their recursive nature, which reveals only the gross manifestation, while the intricacies are hidden away within the structure. Curiously, patterns of social activity in virtual worlds may be found to be of both types—visible as well as obscure. Geospatial patterns such as patterns of user trails are highly visible and fairly simple to visualize, while patterns in semantic space such as those found in chat activity are more difficult to uncover.

5.1. Social Diffusion Patterns

In a virtual world, avatars perform activities in isolation or in the immediate presence of (or collaboration with) other avatars. We use the term *group* to denote a set of users that perform an activity together. The characteristics that define a group are similar to those necessary for the existence of a community [24], including "a common interest" and "being rooted in the same geographical space". However, in educational settings, a group is usually a short-term congregation of avatars who share a common goal. For example, the group of all researchers exploring ways to improve a virtual world for education form a community, while the set of students that take part in a short-term science experiment, form a group within the context of the experiment. A lone user exploring a virtual world in a random manner does not share a common purpose with other avatars and hence does not form part of a group in our sense. Note that this notion of a *group* is different from the way we refer to it in the term *user groups*.

We define social activities as activities that are carried out by a group of avatars as part of a larger objective. In our research, this larger objective is usually *learning* and the activities may be a play, a college-recruitment event, a treasure hunt, etc. The patterns that emerge as a result of these activities, hence, are also social in nature.

The Webster's New Millennium Dictionary of English defines the term *diffusion* in several ways, one of them being "the process by which a cultural trait, material object, idea, or behavior pattern is spread from one society to another". For such a diffusion process, two things are necessary: (1) the diffusing element and (2) the diffusion medium. In the above definition, the cultural trait, material object or idea is the diffusing element. This is the element that actually propagates through society possibly changing the structure of society as it diffuses. The diffusion medium in the above case is society itself, which consists of a large number of discrete units (people) that all share some common characteristic that enables the diffusion of ideas or material objects. Although homogeneity of the diffusion medium is not a pre-requisite for a diffusion process, a heterogeneous diffusion medium presents a complicated case and hence, here we consider diffusion only in a homogenous diffusion medium. We can then classify types of diffusion on the basis of whether the avatars of the virtual world act as (1) the diffusing element, or (2) the diffusion medium.

5.1.1. Users as Diffusing Element

Users in a virtual world move in diverse ways in accordance with their personal goals or the requirements of their group. Members of a group of users move towards areas that are most suitable to their tasks. For short-term activities, this process is not evident, since the quick succession of movements and teleports can hardly be termed diffusion. However, in the case of large worlds such as the *Avatars* world (see http://ccon.org), which is 2 kilometers long and 2 kilometers wide, we see that over time, there is a non-uniform distribution of users through the world due to certain areas of the world being used more often than others. Such a pattern is also reflected in the evolution of virtual world buildings over time, where we see distinct building patterns that seem to suggest that certain parts of the world are more preferred than others (figure 3-1).

5.1.2. Users as Diffusion Medium

An interesting way of looking at user behavior in virtual worlds is obtained by viewing users of virtual worlds as actors and a medium for the diffusion of ideas. Users in a virtual world communicate with each other in a variety of ways,

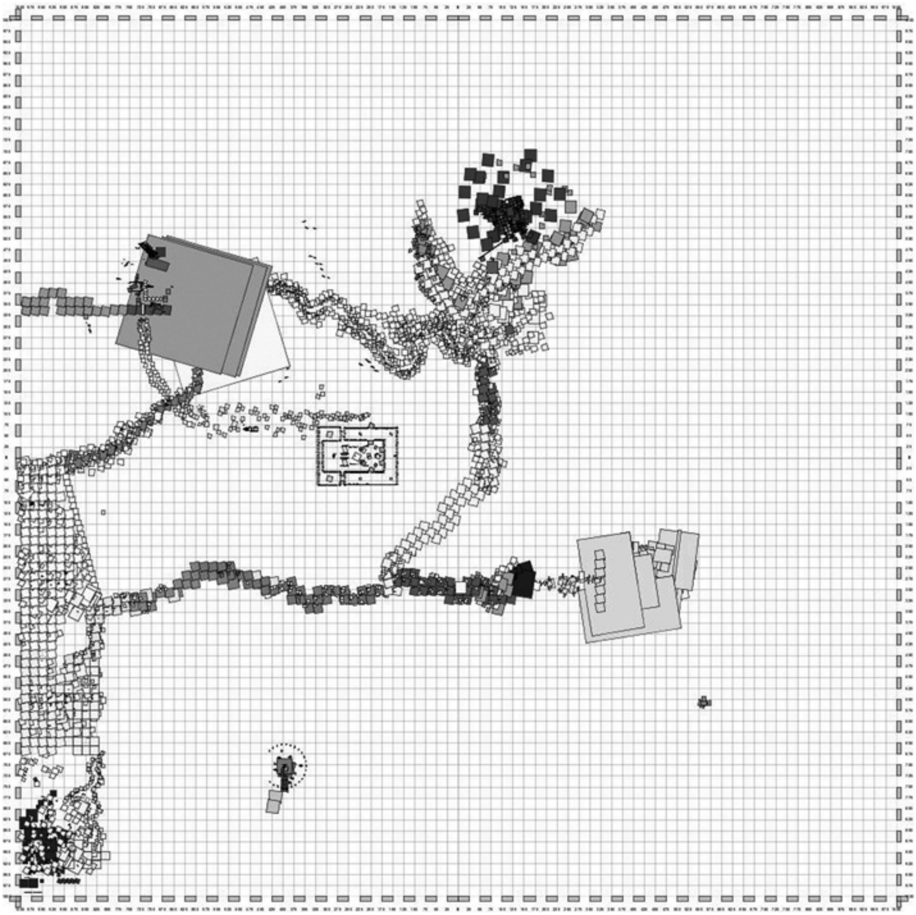


Figure 3-1. Map showing the structure of the "Avatars" world in the ActiveWorlds universe. Grid lines are used to indicate the size of the virtual world: the distance between two grid lines corresponds to 25 meters in the virtual world. The rectangles represent buildings and other structures in the virtual world and are color-coded by age. Older buildings are shown in darker color while younger buildings are shown in lighter color.

including whispering, chatting, gesturing, etc. with chat being the dominant method of communication. In fact, many users show very little movement during the entire duration of their visit to a virtual world and restrict their activity almost exclusively to chatting.

Today, virtual worlds offer a variety of ways to control the effect of the environment on the users and their activities; among the most common being configurable limits on the "visibility" of objects and the "hearing range" of chat. Visibility places a limit on how far an avatar can see in the virtual world. Objects that are farther away can be made less visible (by blurring them or hiding them by fog) or not visible at all, while objects that are closer to the avatar are made clearly visible. Hearing range limits how far away an avatar can move from a sound source (or an avatar uttering chat phrases) before it stops "hearing" sounds or receiving chat text. An avatar must be within a particular distance from another avatar to be able to receive chat-text from the latter. These limits ensure that awareness of events in a virtual world is not *broadcast* to all the users; there is clearly a diversity of awareness or knowledge of events among users. Thus, if an idea or a concept (or simply a word) appears in the chat text of one user, it spreads to the other users subject to the limitations posed by visibility and hearing range. This spread will depend—as with any diffusion process—on a variety of factors including the density and distribution of users, the "strength" of an idea or its relevance to subsets of users. Here the idea becomes the diffusing element, while the avatars become the diffusion medium.

5.2. Emergence of Patterns in Group Situations

There is a great deal of diversity in the kinds of groups that are found in virtual worlds and they vary from short-term task based groups such as the ones we encounter in our research—to longer-term communities such as the "E-Church" world studied by Schroeder, Heather and Lee [25]. Research on groups in virtual worlds is in its infancy and much research is needed before we will achieve a comprehensive picture of groups, roles and interactions in virtual worlds. This task is not simple because groups in virtual worlds have different dynamics compared with groups in the real world. Therefore, one cannot assume that communities in virtual environments are a mere projection of communities in the real world [26]. Keeping this in mind, we caution the reader to view the subsequently discussed types of user groups as a personal observation coming out of looking at large amounts of virtual world data, rather than a complete list.

Leader-Follower: This type of group is characterized firstly by the presence of a single avatar that is automatically accepted as the leader and secondly by the existence of prescribed paths and actions. Members of the group (not including the leader) are the followers which obey orders, follow instructions given by the leader or simply copy her actions. Some examples of such a group include a group taking a guided tour (the tour guide being the leader) or a group of students working on a building project (the leader of the group being their teacher). The followers in the group are not expected to significantly deviate from the pattern set by the leader.

Moderated: This type of group is different from the one outlined above in that there are no real followers in the group. The group consists of autonomous members free to do as they wish within the scope of the activity which the group is involved in. In a way, the moderator *could* be considered the leader of the group, since the moderator holds the right to decide upon the range and variety of actions that can fall within the scope of that particular group activity. An example of such a group is a research discussion group on virtual worlds led by the organizer of the event. In this particular group, researchers are free to discuss issues related to research in virtual worlds. However, if a heated debate gets underway or if the group begins to digress away from the main theme of discussion, the moderator can step in and restore order and focus. Typically such groups also possess the trait that its members all share a personal goal that is a subset of or is related to the common goal of the group. For example, every virtual world researcher has his/her own research agenda. However, a common thread of interest in virtual worlds binds the group together.

Competitive: This type of group is characterized by the presence of sub-groups within the main group, each of which competes against the other for some reward. All members of the sub-group share with each other a sense of belonging and a group identity. At the same time, each sub-group member maintains a sense of separation from the competing sub-groups, while they may all belong to the same super-group. In the geospatial sense, this group is usually seen to have fixed destinations but variable paths as each sub-group takes varying approaches to achieve their goals. A treasure hunt played by kids of a class is an example of a situation where such groups are encountered.

Ad-hoc: This group is characterized by its short-term existence. Members of this group do not usually possess a shared objective outside the context of the group activity being performed. For example, in the case of a group of school kids taking part in a role play, the kids may or may not have a common goal related to stage-performance. In this particular case, the group is brought together for the sake of the play and is usually dissolved as soon as the task ends. Each child in the group associates himself/herself with the group activity only as long as the activity is being carried out.

The traits described above are not mutually exclusive. Two avatars can collaborate in order to perform a task and this group might be competing against other similar groups, all of which belong to the one bigger group that is led by a single avatar. Also one might notice that the above listed traits result in progressively more decentralized groups. The level of decentralization and more importantly the concurrence of objectives of the members of the group can play a significant role in the paths taken by members of the group during an activity. Geospatially speaking, the trail of a member of a group or the group as a whole can vary depending on how coherent the behavior of the members of the group is with respect to each other and the group as a whole. Looking at the patterns of avatar trails in virtual worlds with this idea in mind reveals that we can classify group trails into the following four types in terms of the variation of their geospatial distribution over time [8] (see figure 3-2):

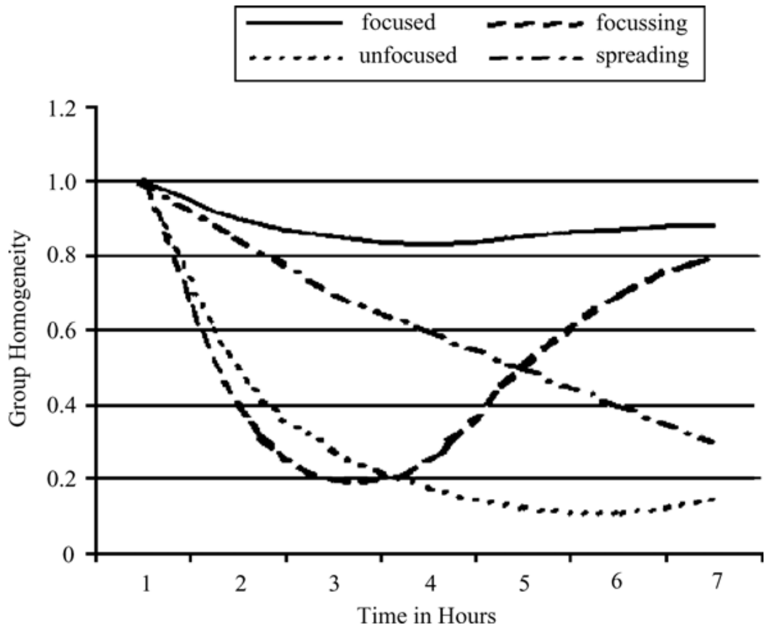


Figure 3-2. A hypothetical example showing the diffusion behavior of groups of avatars. The four types of behavior can be represented by a function that captures the variation of group homogeny over time.

- *Focused*: This group remains tightly knit and the members of the group do not deviate much from each other's paths.
- *Unfocused*: The members of this group move as they wish, resulting in quite random trail patterns.
- *Focusing*: This group starts out dispersed all over space but over time tends to congregate as the trails of the group members become increasingly similar to each other.
- *Spreading*: This group starts out with all its members on a common trajectory, but over time, the members of the group diverge, resulting in very diverse trail patterns.

One can now see that specific kinds of events can result in specific types of diffusion behavior. An art exhibition where all visitors follow pre-set paths around the exhibits results in focused group trails. A random group of users in a public virtual world results in unfocused diffusion behavior, as each member explores the world individually. An event where members of the group are assigned set regions of the world which they scan results in focusing or spreading group trails. Identifying the type of group trail resulting from an event can give researchers a good idea of whether the event went as planned. This idea is elaborated in the next section.

6. Analysis and Visualization

Any study on patterns requires tools for repeated and consistent analysis of data. A good toolkit can facilitate this by providing researchers with a wide variety of perspectives from which to analyze and view a data set. An effective visualization can summarize the data, highlight anomalies or outstanding features and even reveal patterns that are impossible to identify by looking at data in its raw form. In this section we describe how visualization can support the study of social behavior.

6.1. The ActiveWorld Toolkit

The ActiveWorld Toolkit (available for download at http://ella.slis.indiana.edu/~sprao/research/virtualworlds/) is a free toolkit that helps researchers to analyze and visualize world structure and user behavior data in virtual worlds built with ActiveWorlds technology. The toolkit allows one to load world data (*propdump* and *registry* files) and data about activity of avatars in the virtual world (bot log files). The world data is usually obtained by using an administrative utility for the virtual world which can export data about location, orientation, size and information on how the objects react in response to user actions such as movements or mouse clicks. User data is collected in the virtual world using bots which continuously monitor user activity such as movements, chats, clicks, teleports, etc.

To analyze how this toolkit can help researchers analyze virtual world events, consider an educational treasure hunt in a virtual world where the idea is to let kids follow clues to find treasures, picking up bits of information about issues such as environmental pollution, conservation of natural resources, etc. This requires designing the clues for the hunt in a way that is appropriate to the level of the users' cognitive abilities. One cannot incorporate clues so complex that kids get frustrated and quickly lose interest. On the other hand, the clues must be complicated enough to ensure that not all kids immediately figure them out. Hence, the ideal set of clues for the hunt would help the kids "stay in the flow" [27] of the hunt.

Also critical to the success of such an event is the ability to successfully navigate through the world. Navigational clues must be placed in strategic locations to ensure that the kids can find their way around the world, but at the same time they should not give away the locations of the clues themselves. Some questions about the treasure hunt that can be asked are "Were most of the users able to follow the clues and find their way through the world?", "Which group of users was most focused?", "Which users were lost and at what point?", "Which parts of the treasure hunt were most or least challenging?" etc. These are some questions which can be answered using visualization of user data.

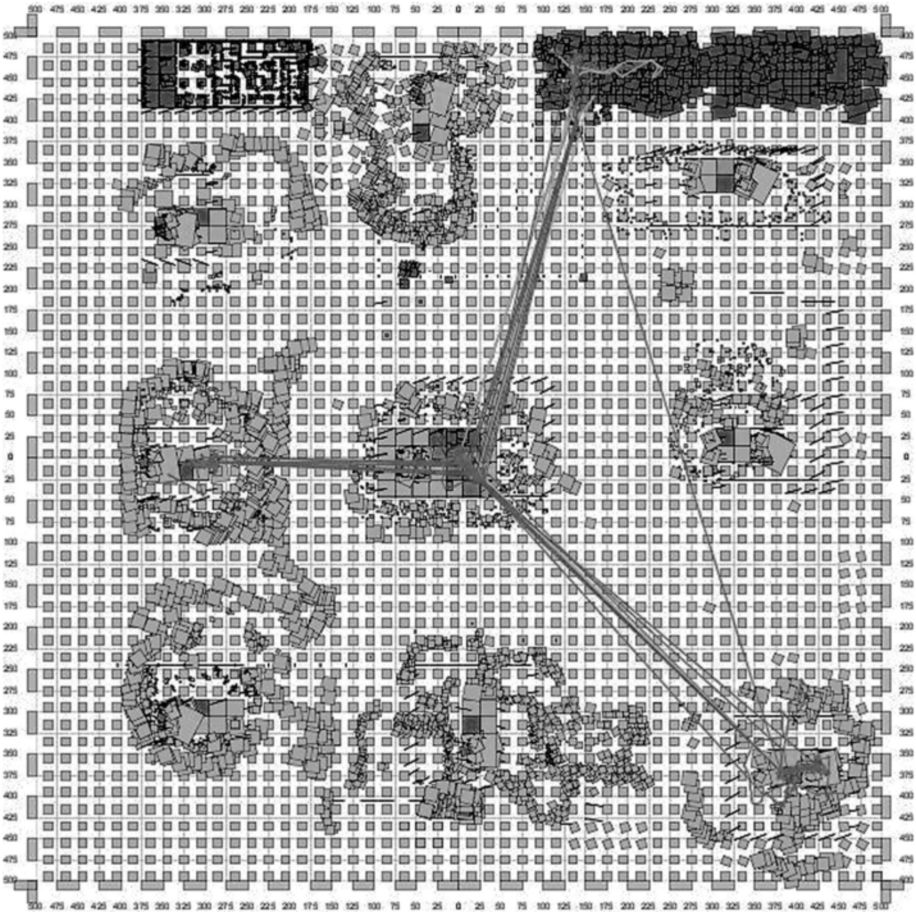


Figure 3-3. Visualization of the trails of the participants of three parallel Spanish learning sessions in LinkWorld. The three sessions are taking place in the north, south-east and west. The trails of the participants (avatars) are represented by lines. The rectangles represent buildings and other structures in the virtual world.

An example of such a visualization is shown in figure 3-3. This visualization shows the trails of users that took part in three parallel Spanish learning sessions in LinkWorld, with the treasure hunt being held in the northern part of the world. This map gives us an overview of where avatars went and where they have been most active, i.e., where they have chatted the most or where they have clicked and interacted with objects the most. In order to answer specific questions, however, one needs to be able to examine the interactions of each individual user. This is facilitated in the toolkit using the zoom function, which allows one to focus on the area of interest. The toolkit allows one to smoothly move between overview and detail mode, thus helping the user maintain her orientation while working with the toolkit.

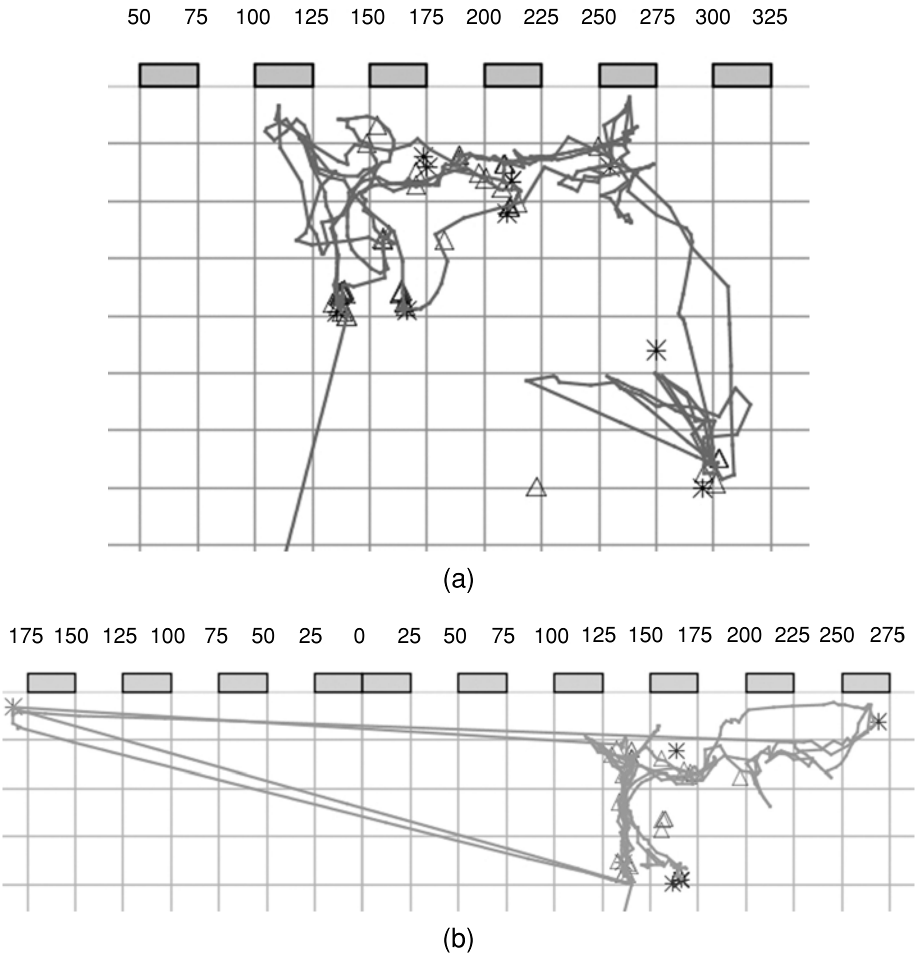
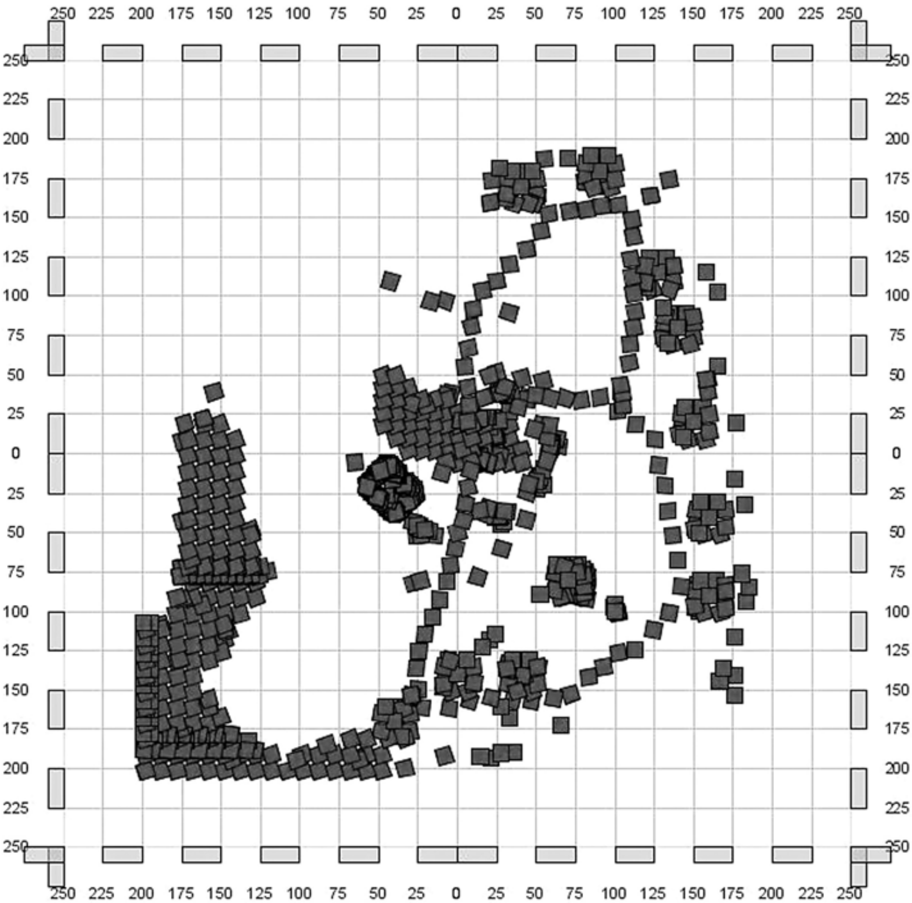


Figure 3-4. Two participants who have strayed away from the main treasure hunt area during the event. Triangles show locations where participants chatted, stars show locations where they clicked objects in the world, while trails show the path that these two participants took. The numbers at the periphery indicate the distance in meters from the centre of the virtual world along the *x*-axis.

6.2. Analysis of Geo-Spatial Patterns

Figure 3-4a and b show zoomed-in views of the trails and chats (triangles) of two users who seem to have strayed significantly from the area where the treasure hunt was being held. In both cases an examination of the chat text revealed that the moderator of the treasure hunt realized that two participants were lost and made an attempt to bring them back. One can also look at the chat text of the participants to understand when they themselves realized that they were not in the right area. There are several reasons why such confusion could have arisen: the clues for the hunt were not right, or perhaps the placement of

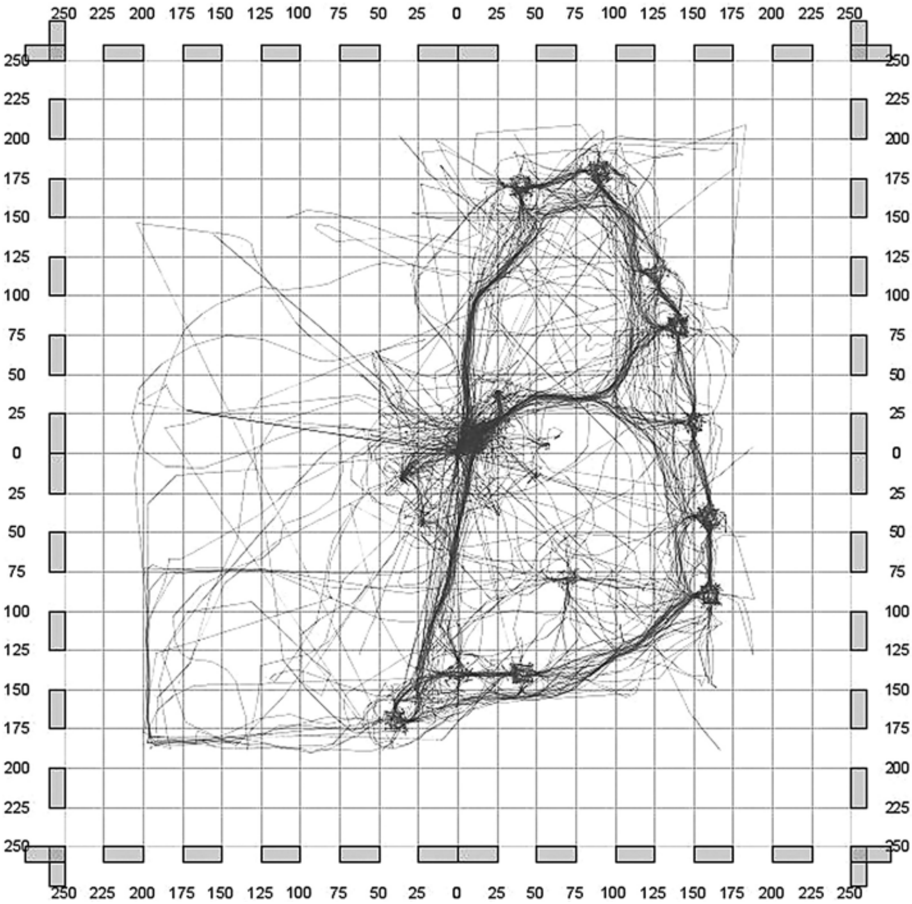


(a)

Figure 3-5. Three views of an art show in the Avatars world. (a) Structure of the world showing positions of objects. (b) Trails of avatars viewing the art show (represented by lines). (c) Locations where users chatted (triangles) and clicked objects (stars).

objects was such that the participants were misled into following a path that was out of bounds for the event, or maybe the participants were curious to explore other areas in the world. A conclusion of the latter type is especially useful when a statistical analysis is planned; one can quickly decide whether or not the behavior of the user in question is an anomaly and accordingly exclude or include him/her in the analysis. A quick peek at the chat text of these avatars revealed that one of them (figure 3-4a) spent a considerable amount of time in an area out of bounds of the actual hunt before coming back to the main area, while the other avatar (figure 3-4b) was brought back to the main area almost immediately after they left the group.

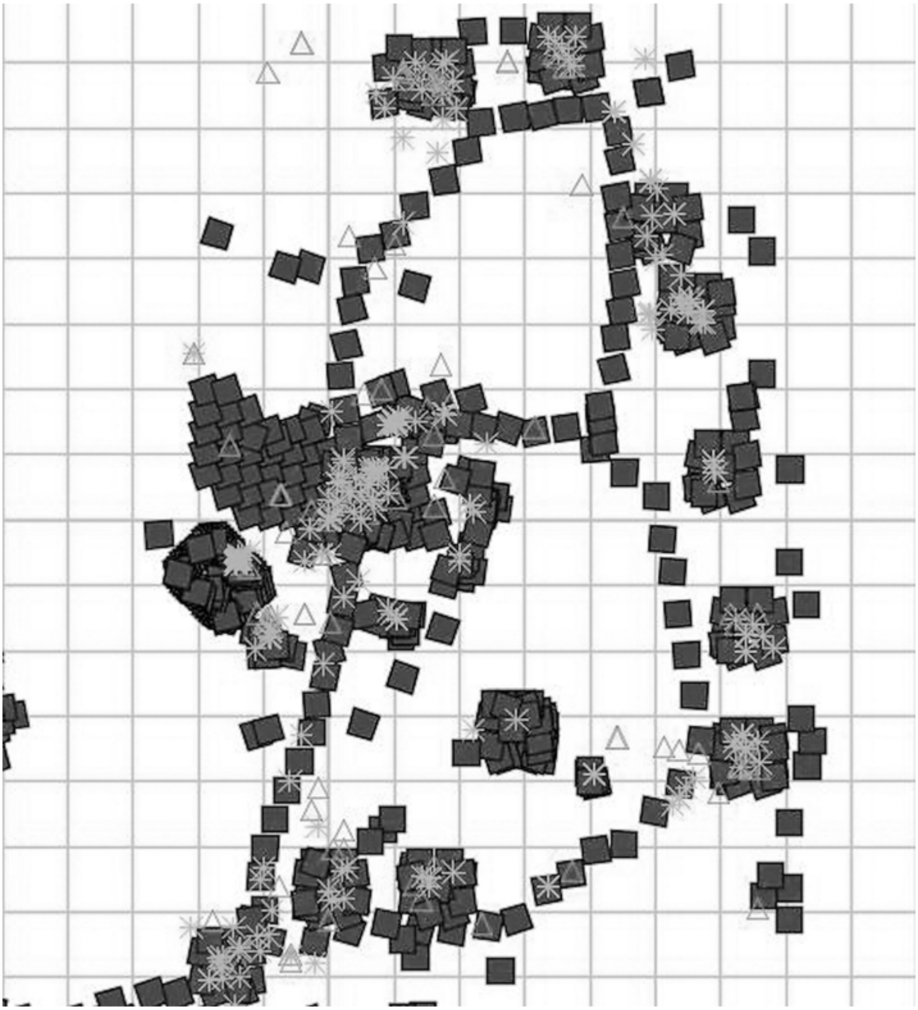
Figure 3-5 shows another set of visualizations that uncover behavioral patterns in virtual worlds. These are visualizations of an art show held during the



(b)

Figure 3-5. (Continued)

Avatars! Conference in 2002 (see http://www.ccon.org). Figure 3-5a shows the structure of the world. The art exhibits are placed roughly in the shape of the letter "B". Figure 3-5b shows the trails of the avatars as they explored the exhibition. From these two figures, it is clear that the avatars followed the path of exhibits very well. The trails which deviate from the center were found to belong to avatars who, after seeing all the exhibits, decided to explore the world a bit more and then exit. Finally, figure 3-5c shows the locations of chat (triangles) and clicks (stars) overlaid over the structure of the world (dark objects). This final figure gives us a clear picture of the exhibits near which people chatted with each other or interacted with objects. This information can be used along with post-event questionnaires to determine what aspects of those locations or exhibits encouraged people to talk to each other.



(c)

Figure 3-5. (Continued)

6.3. *Analysis of Chat Data*

Chat data in virtual worlds is especially interesting because it can potentially reveal patterns about users' thought processes while they performed activities inside the virtual world. Chat data collected over a long period of time give us a picture of the evolution of the topics discussed among the inhabitants of the virtual world. A particularly interesting analysis is the detection of changes in the frequency of usage of particular words. The results enable us to make inferences about the emergence of topics in a conversation among avatars.

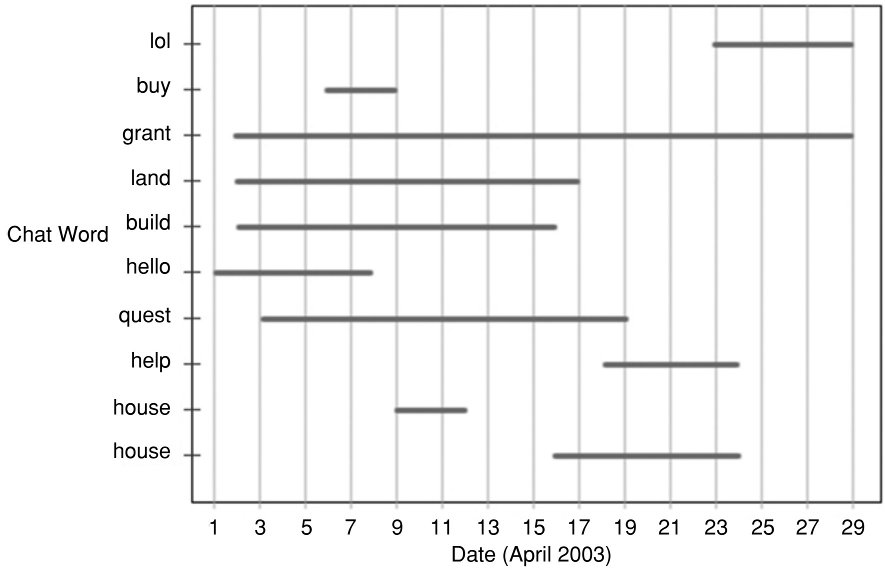


Figure 3-6. Top ten bursty words over the month of April 2003 in the Culture world of the Quest Atlantis universe. The vertical (*Y*) axis shows the burst words, while the horizontal (*X*) axis shows the dates during which the bursts occurred. For each word, a horizontal line is drawn starting from the date that the burst began till the date that burst ended.

Such an analysis demands that we do not restrict our search to any particular set of words, but rather that we let the analysis itself determine the set of most popular words. Specifically, we are looking for words which suddenly "burst on the scene", generate a lot of activity and then die down. We believe such *bursty* words are good indicators of popular topics in chat.

In order to identify bursty words, we can use a predictive model such as Kleinberg's burst detection algorithm [7]. Simply speaking, this algorithm considers a word to be bursty if it appears with a high frequency over a time period that is short compared to the total time span of the complete data set. Note that the set of words identified using this analysis are not simply the set of words which have high frequency of usage; in fact a word that constantly keeps appearing in chat will not be recognized as a bursty word.

Figure 3-6 shows the result of such an analysis performed on chat data collected over the month of April 2003 for the Culture world in Quest Atlantis universe (http://questatlantis.uni.activeworlds.com). The "*X*" axis shows the bursty words. The "*Y*" axis represents the dates in the month of April. Each horizontal line begins at the date at which the word started to burst and ends at the date at which the usage of that word died down. Table 3-1 shows the strength of the burst for these words. The "strength of burst" indicates the amplitude of the burst of that word, i.e., how intense the usage of that word was. Table 3-1 shows us that the word "lol" (an acronym for Laugh Out Loud), appears with very high

Table 3-1. Strength of burst of the top ten bursty words in chat data over the month of April 2003 in Culture world in the Quest Atlantis universe.

Word	**lol**	**Buy**	**Grant**	**Land**	**Build**	**Hello**	**Quest**	**Help**	**House**
Strength of burst	88.213	50.387	45.257	33.25	31.555	30.492	29.879	29.481	28.694

amplitude compared to other words such as "grant" or "hello". However, the latter category of words have a longer life-span: the word "lol" remains in the conversation for 7 days but the word "land" and "build" remain dominant for 15 and 16 days, respectively. Note that the word "house" appears twice, meaning that this word appeared as the dominant word twice in one month—the first time between April 9 and 12 and the second time between April 16 and 24.

From figure 3-6, we know exactly what day the burst of a particular word occurred and therefore we can go back to the original data and determine the cause for that burst. In this case, the reason for the sudden burst of the word "lol" was traced back to a kid who kept welcoming every new visitor into the virtual world with strange and funny welcome messages. The messages were written in a way that made them look as if they were coming from an automated bot. Every time this kid made a fool of a new user, everyone assembled in the virtual world would burst out in laughter. The other words have a more serious tone and clearly indicate building and trading activities going on very actively for a period of about 15 days. One can also take advantage of the starting and ending dates of the burst and make inferences about the chain of events that took place. So we see that "hello" is one of the first words that burst—probably a time for introductions—after which users got down to business talking about "quest", "build" and "land", ending with "lol". One can see that simple analyses like these can lead to interesting inferences that can then be verified using questionnaires and interviews.

The sequence of chat data can be visualized to show trends in the rise and fall of the word frequency over time. One of the most frequently used words in chat text which did not appear in the analysis outlined above was the word "you". This word was consistently used with very high frequency and hence was ignored by the burst algorithm. Figure 3-7 shows the frequency of usage of the word "you" over time in the same data set analyzed above visualized using TimeSearcher [28]. From the figure, the pattern of usage over time is immediately evident: the word started off at a low frequency, its average increased slowly over time, it took a big jump after the half-way point through the time period of observation and maintained a relatively high average frequency until finally the frequency dropped very suddenly. This pattern of frequency can then be compared to changes in the user group, environment or activity in the virtual world to yield additional clues.

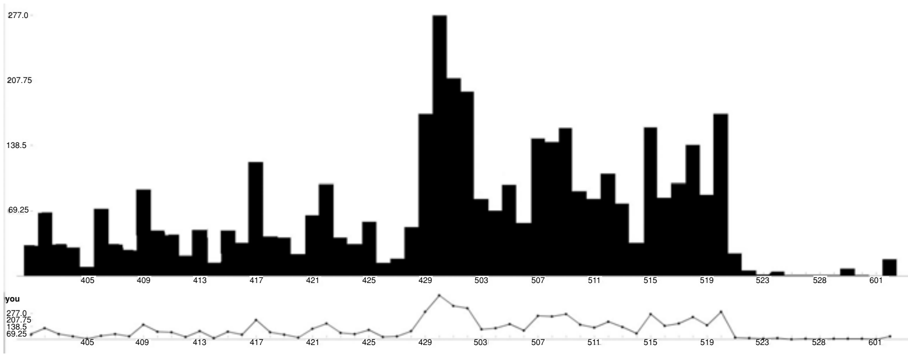


Figure 3-7. Visualization of the frequency of usage of the word "you" in a two-month sample of chat data in Culture world in the Quest Atlantis universe. The vertical axis (Y) indicates the frequency of usage of this word, while the horizontal (X) axis indicates time. The bar graph (top) and the line graph (bottom) provide the same information. However, in the TimeSearcher application, the bar graph allows users to zoom in and out as well as query for details over a particular time period, while the line graph acts as a constant reference.

7. Outlook

In this chapter, we have seen how one can combine diverse methods of analysis and visualization in order to create powerful ways to analyze user behavior data and discover patterns in them. The use of patterns for social research in virtual worlds is not only powerful, but necessary in order to truly understand the impact of space and time on the behavior of individual users and groups in virtual worlds. Visualizations provide researchers with effective means to discover these patterns and enable efficient communication of analysis results to peers.

In the future, we hope to apply these analysis and visualization techniques to gain an understanding of virtual world behavior in real-time. One can imagine a tool that provides continuous real-time summaries of activity patterns of users as they perform tasks in a virtual world. Such a real-time map could provide navigational support to users, present a real-time summary of chat topics emerging in different areas of the virtual world, or even display a dynamically-generated social network of avatars as they interact in the virtual world. We believe that such tools will prove to be essential for determining the influence of spatial, semantic, and social factors on dynamic group behavior.

Another line of research that looks promising is the development of techniques to automatically identify roles that avatars take in a group based on spatial patterns, chat utterances and interaction data. For example, we can ask the question, "From the patterns of social behavior, is it possible to identify automatically which avatar emerged as the leader?" Such a question calls for an intelligent synthesis of theories and techniques developed in computer supported cooperative work, information visualization, spatial analysis, clustering methods, computer-mediated communication [29], social network analysis [30], and others.

Acknowledgements

We would like to thank Bonnie DeVarco of the BorderLink project (http://www. borderlink.org) for her support throughout this research and for access to the LinkWorld data. Thanks also to Sasha Barab and Bob Carteaux of the Quest Atlantis project (http://questatlantis.uni.activeworlds.com) for their valuable comments and for permitting us to log Quest Atlantis log data and Bruce Damer of Digital Space Commons (http://www.digitalspace.com) for granting us permission to log the "Avatars!" Conference and generate maps of the Avatars world. Thanks also to Rich Bernstein of Cornell University whose helpful suggestions led to several improvements to the ActiveWorld Toolkit. This work is supported by a National Science Foundation grant under Role-0411846.

Figure 3-2 reprinted with permission from "Social Diffusion Patterns in 3D Virtual Worlds", Information Visualization Journal, Vol. 2, Issue 3, Sept. 2003. Copyright Palgrave-Macmillan, UK.

References

1. Chen, C. (1999). *Information Visualisation and Virtual Environments*. London: Springer Verlag.
2. Card, S., Mackinlay, J., & Shneiderman, B. (Eds.) (1999). *Readings in Information Visualization: Using Vision to Think.* San Francisco: Morgan Kaufmann.
3. Ware, C. (2000). *Information Visualization: Perception for Design*. Morgan Kaufmann Interactive Technologies Series. San Francisco: Morgan Kaufmann.
4. Spence, B. (2000). *Information Visualization*. Reading, MA: Addison-Wesley.
5. Börner, K. (2002). Twin worlds: Augmenting, evaluating, and studying three-dimensional digital cities and their evolving communities. In M. Tanabe, P. van den Besselaar, & T. Ishida (Eds.), *Digital Cities II: Computational and Sociological Approaches*. Japan: Springer Verlag, pp. 256–269.
6. Rabiner, L.R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. In *Proceedings of the IEEE, 77*(2), February, pp. 257–286.
7. Kleinberg, J. (2002). Bursty and hierarchical structure in streams. In *The 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, July 23–26.
8. Börner, K. & Penumarthy, S. (2003). Social diffusion patterns in three-dimensional virtual worlds. *Information Visualization, 2*(3): 182–198.
9. Aref, H., Charles, R.D., & Elvins, T.T. (1994). Scientific visualization of fluid flow. In C.A. Pickover & S.K. Tewksbury (Eds.), *Frontiers of Scientific Visualization*. Hoboken, NJ: Wiley Interscience.

10. Senger, S. (1996). Digital cadavers (TM): An environment for the study and visualization of anatomic data. In *Proceedings of the Visible Human Conference*, National Institutes of Health, Bethesda, Maryland, p. 127.
11. Dorigo, G. & Tobler, W. (1983). Push pull migration laws. *Annals Association of American Geographers, 73*(1): 1–17.
12. Plaisant, C., Milash, B., Rose, A., Widoff, S., & Shneiderman, B. (1996). *Life lines: Visualizing personal histories*. In the *ACM CHI '96 Conference Proceedings*. New York: ACM Press, pp. 221–227.
13. Donath, J.S., Karahalios, K. & Viegas, F. (1999). Visualizing conversation. *Journal of Computer Mediated Communication, 4*(4). Available at http://www.ascusc.org/jcmc/vol4/issue4/donath.html
14. Hochheiser, H. & Shneiderman, B. (2001). Using interactive visualizations of WWW log data to characterize access patterns and inform site design. *American Society for Information Science, 52*(4): 331–343.
15. Donath, J.S. (1995). Visual Who: Animating the affinities and activities of an electronic community. *In Proceedings of the Third ACM International Conference on Multimedia*, San Francisco, CA, pp. 99–107.
16. Smith, M., Farnham, S., & Drucker, S. The social life of small graphical chat spaces. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 205–220.
17. Naper, I. (2000). System features of an inhabited 3-D virtual environment supporting multimodality in communication. In *Proceedings of the 34th Annual Hawaii International Conference on Systems Sciences*. Maui, Hawaii: IEEE Computer Society.
18. Whyte, W.H. (1980). *The Social Life of Small Urban Spaces*. Washington, DC: The Conservation Foundation.
19. Dodge, M. & Kitchin, R. (2001). *An Atlas of Cyberspaces*. Reading, MA: Addison Wesley.
20. Schroeder, R., Smith, A., & Huxor, A. (2001). Activeworlds: Geography and social interaction in virtual reality. *Futures, 33*: 569–587.
21. Börner, K., Penumarthy, S., DeVarco, B. J., & Kerney, C. (2004). Visualizing social patterns in virtual environments on a local and global scale. In *Digital Cities 3: Local information and communication infrastructures: Experiences and challenges*. To be published by Springer Verlag.
22. Börner, K., Jun Lee, G., Penumarthy, S., & Jones, R. J. (2004). Visualizing the VLearn3D 2002 conference in space and time. In *Visualization and Data Analysis*, San Jose, CA. SPIE-IS&T, Vol. 5295: 24–32.
23. *Federal Geographic Data Committee, Geospatial Standards.* Available at http://www.fgdc.gov/publications/documents/standards/standards.html
24. Becker, B. & Mark, G. (2002). Social conventions in computer-mediated communication: A comparison of three online virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 19–62.
25. Schroeder, R., Heather, N., & Lee, R. (1999). The sacred and the virtual: Religion in multi-user virtual reality. *Journal of Computer-Mediated Communication, 4*(2). Available at http://www.ascusc.org/jcmc/vol4/issue2/schroeder.html
26. Wellman, B., Salaff, J., Dimitrova, D., Garton, L., Gulia, M., & Haythornthwaite, C. (1996). Computer networks as social networks: Collaborative work, telework, and virtual community. *Annual Review of Sociology, 22*: 213–238.
27. Bederson, B. (2004). Interfaces for staying in the flow. *Ubiquity, (5)*27, Sept. 1–7. Available at http://www.acm.org/ubiquity/views/v5i27_bederson.html

28. Keogh, E., Hochheiser, H., & Shneiderman, B. (2002). An augmented visual query mechanism for finding patterns in time series data. In *Proceedings of the Fifth International Conference on Flexible Query Answering Systems*, October 27–29, Copenhagen, Denmark. *Lecture Notes in Artificial Intelligence*. London: Springer.
29. Herring, S.C. (2002). Computer-mediated communication on the Internet. *Annual Review of Information Science and Technology, 36*: 109–168.
30. Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Chapter 4

COLLABORATIVE VIRTUAL ENVIRONMENTS FOR SCIENTIFIC COLLABORATION: TECHNICAL AND ORGANIZATIONAL DESIGN FRAMEWORKS

Diane H. Sonnenwald

1. Introduction

Collaboration among scientists is human behavior that facilitates the sharing of meaning and completion of tasks with respect to a mutually shared scientific goal, and which takes place in social settings. Scientific collaboration across geographic distances began centuries ago, when scientists began utilizing postal and shipping services to exchange ideas as well as samples of plants and animals. More recently, collaborative virtual environments have brought new opportunities—and challenges—for scientific collaboration across distances. This chapter discusses two different challenges: designing collaborative virtual environment software tools, and designing organizational structures and practices to facilitate collaboration across geographical distances. To address the first challenge, a technical design framework that focuses on supporting situation awareness is proposed. The framework is based on research conducted by designing and evaluating a scientific collaboratory system, called the nanoManipulator Collaboratory, that allowed scientists to synchronously conduct experiments, collecting and analyzing data from an atomic force microscope. To address the second challenge, an organizational design framework based on a two-year case study of a distributed scientific organization is proposed. These two frameworks may be relevant for a wide range of distributed scientific work.

Today, many scientists often use multiple collaborative virtual environment (CVE) tools in various combinations to facilitate collaboration. Examples of tools that support synchronous collaboration are the telephone, video conferencing, instant messaging, chat, shared electronic whiteboards, virtual networked computing, shared access to scientific instruments, and

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 63–96.

shared applications, such as shared data visualization programs. Examples of tools that support asynchronous collaboration are e-mail, file transfer programs, WIKIs, electronic lab notebooks, project management tools, and listservs. Examples of tools that support individual access to information or a device include web pages that provide information about scientific research and outcomes, digital libraries and search programs that allow scientists to contribute, find and/or comment on scientific publications and data, and single-user remote access to scientific instrumentation and data.

No one tool today provides all features needed to fully support collaboration across distances during the entire scientific research life cycle. One challenge facing collaborative virtual environment tools is the ability to fully support shared situation awareness. Situation awareness has been defined as: "continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture in directing further perception and anticipating future events" [1, p. 11].

It is a general sense of knowing about things that are happening in the immediate environment and includes having both an accurate understanding of the situation and the knowledge to respond appropriately as the situation evolves [2]. Based on previous research and our empirical studies, the types of information needed to develop and maintain situation awareness include contextual, task and process, and socio-emotional information. Research in virtual reality systems suggests that control, sensory, distraction and realism attributes of technology contribute to a sense of presence [3]. Consideration of these attributes with respect to contextual, task and process, and socio-emotional information provides insights to guide technical design decisions. The resulting framework was used when designing a CVE for scientific collaboration [4]. Results from a controlled experimental evaluation of the collaboratory system help illustrate the framework's utility.

It has long been known that organizational structures and practices influence technology adoption and use [e.g., 5, 6]. In addition to supporting situation awareness through technology, scientific collaboration in collaborative virtual environments may often require new organizational practices, especially when larger numbers of scientists need to collaborate across distances. Thus providing a framework to help design CVE tools should be augmented by a framework to help design organizational structures and practices in which the tools will, ideally, be embedded.

The organizational design framework to facilitate scientific collaboration proposed in this chapter is based on a two-year case study of a group of over 100 scientists. The scientists were primarily chemists and chemical engineers at four universities in the USA who were members of a research centre. They collaborated during all phases of the scientific process, including sharing knowledge during proposal development, scientific instrument design and construction, experiment design, data collection and analysis, and dissemination of results through papers and presentations. The case study began during the beginning stages of the centre and continued for 2 years. While conducting the case study, the author was a participant observer, having both complete and peripheral membership roles, collecting case study data and providing advice regarding collaboration technology and collaborative practices. Case study data included interviews, observations of meetings, sociometric surveys, and centre documents. These data were analyzed in the ethnographic and grounded theory traditions [7, 8], and the organizational design framework discussed in this chapter emerged from this analysis.

The organizational design framework introduces a new concept, called the conceptual organization. A conceptual organization has characteristics in common with traditional organizations, invisible colleges, collaboratories and virtual teams. However, it uniquely combines a management structure that is interwoven across other organizations, collaborative work practices based on collaboration technology, and use of integrative, economic and destructive organizational power. It has few employees in the traditional sense; most members are scientists who join the organization because they wish to contribute to its vision and goals. Benefits of a conceptual organization include its ability to discover solutions, quickly and effectively contributing to relevant dynamic knowledge bases and meeting diverse stakeholder needs with minimum capitalization and start-up costs [9].

This chapter first presents the technology design framework and, second, the organizational design framework. In presenting the technology design framework, the type of information needed to develop and maintain situation awareness in scientific collaboration, and features of technology to support the acquisition and sharing of these types of information are discussed. The application and evaluation of the framework in the context of the nanoManipulator collaboratory is also discussed. In presenting the organizational design framework, it is defined and compared to other organizational designs. Details regarding management structure, organizational membership, stakeholders, and collaborative knowledge management practices and use of technology are presented. Potential benefits and an evaluation of the organizational design framework are also presented. In sum, this chapter is a first effort to consider both technology and organizational design relevant for a wide range of future distributed scientific work.

2. A Technical Design Framework Focusing on Supporting Situation Awareness

2.1. Information to Develop and Maintain Situation Awareness

To develop an understanding of situation awareness in scientific collaborative work, 27 interviews with faculty, postdoctoral and graduate student scientists actively collaborating with one another were conducted. The average length of these interviews was 1.5 hours, with a minimum duration of an hour and a maximum duration of 2.75 hours. Study participants were also observed on nine occasions as they conducted experiments while working alone and with others. Analysis of the data and previous research suggest that situation awareness in scientific research collaboration requires several types of information, including contextual, task and process, and socio-emotional information. Distinguishing between these types of information facilitates our understanding of situation awareness and technical requirements for collaborative virtual environments.

C*ontextual information* is a broad sense of the context in which things are happening. Context can be defined as a "framework of meaning" [10, p. 8] or a "framework of understanding" [11, p. 52]. Contextual information includes information regarding norms of scientific practice, research goals, organizational culture and work environment. Contextual information can vary between collaborators, as one scientist described:

> The person that we were collaborating with was so much into "let's hurry up and publish this before so-and-so beats us and we won't get credit if they beat us." . . . And in the end, we found out the results weren't reproducible. I resisted all of the ideas this collaborator had to publish and in the end, it was the best thing I ever did because if we had published it we would have been wrong.

In a sense, contextual information includes the "rules of the game" and the "players in the game," and how to apply the rules. A scientist stated:

> I don't mind the political games necessary to see a few things come together . . . To get an idea . . . going . . . you have to get the blessings of various people.

When collaborators work face-to-face in the same context, they may already know most of the contextual information relevant to the situation, reducing the amount of contextual information that must be mediated by technology. However, when collaborators come from different contexts, they need to be able to discover differences and similarities in their understanding of the context, and possibly discuss or negotiate those differences.

Task and process information is defined as information about current and relevant task activities and work processes. It includes information about tasks currently being performed and who is performing them. It also includes information about what tasks should be performed, how they can be performed, who can perform them and where and when they can be performed. Task and process information assists an individual in understanding what collaborators are doing. It also assists in creating expectations regarding what collaborators might do. When two collaborators share an in-depth understanding of processes, they may appear to function as one with work responsibilities passing smoothly between them.

An individual may increase his or her task and process information by observing the sequence of tasks another person or group of people are performing and by discussing tasks and processes with them. As one scientist explained:

> Every now and then, she would look at us over the shoulder, I guess, and see how the experiment was going. And we talked with her too, every now and then.

Collaborators may have different task and process information, especially when collaborators come from different disciplines or have different expertise. For example, a scientist told us:

> [My collaborators] will generate data and then they'll go "we're going to run this through our computer software to see blah, blah, blah" .... I have no sense of what that involves. Is that a week of running a mainframe or is this something you put on there and click it and it comes back and says here's your picture?

Socio-emotional information is interpersonal information about collaborators. It includes information about their skills, work styles, approach to science, likes and dislikes, personality and emotional state. Several scientists discussed the important role socio-emotional information plays in their collaborative work:

> The best collaborations I have are the ones where the person I'm collaborating with thinks differently than I do ... [this] is much more important than just getting experiments done more quickly.
>
> Interviewer: How do you judge whether somebody would be a good collaborator? What criteria do you use?
>
> Scientist: [The] kind of behavior they have towards other people. Do they behave ethically? Are they forthright? Do they openly discuss their research or are they secretive? Are they, it may sound silly, but do I like them? ... Do I find them interesting people with sort of the same kind of values that I have towards the science?

Bales [12] and Nardi, Whittaker and Bradner [13] have shown that groups, working both face-to-face and remotely, communicate socio-emotional information. They show tension, tension release, antagonism, enthusiasm, solidarity, agreement, disagreement and empathy through a variety of mechanisms including jokes, questions, assertions and body language.

Contextual, process and task, and socio-emotional information can be interrelated. Information or a lack of information of one type can enhance or limit one's understanding of other information. For example, a scientist described collaborating with a professional and not understanding why the professional did not complete several tasks. The scientist lacked contextual and socio-emotional information about his collaborator, and could not understand the task and process information at hand:

> [He] is a clinician that deals with children who have this lethal disease . . . I just cannot seem to ever get him to come over or respond to e-mails . . . [is] he so inundated with clinical stuff that he can't carve out of his day what he needs to do the scientific? . . . I don't understand that . . . He can treat these patients for his whole career. Here's an opportunity to potentially bring a cure to them, and I don't understand why [he] can't say this is a priority.

This lack of contextual and socio-emotional information not only hindered the immediate collaboration but also future collaborations.

In summary, research suggests that situation awareness is built on a foundation of contextual, task and process, and socio-emotional information from previous situations. Poorly designed collaborative virtual environments that do not adequately support the development and maintenance of situation awareness may not only reduce the quality of current work but also of future work.

2.2. Technology Features to Enable Situation Awareness

When scientists collaborate face-to-face, they share an immediate environment and can develop situation awareness using contextual, process and task information gained through exploring, and experiencing, the (local) environment independently and/or collaboratively. However, when collaborating across distances, this exploration must occur across multiple environments (local and remote). The exploration of the remote environments is no longer a direct experience, but is mediated by technology. It is important to design systems that enable scientists to obtain contextual, task and process, and socio-emotional information about the remote environments independently and/or collaboratively.

Substituting "virtual" for "remote" in the previous sentence makes an obvious link to virtual reality (VR) technology. For CVE tools, the goal is to enable users to create and maintain situation awareness at the remote and local sites; in virtual reality technology, the goal is for users to create and maintain a sense of presence, or "being there" in a place other than where she is physically.

Virtual reality research suggests several attributes, or factors, of virtual reality systems that contribute to providing a sense of presence [3, 14–16]. More recent research [e.g., 17, 18] investigates the impact of these types of system attributes on task effectiveness.

Witmer and Singer [3] organize VR system attributes into four groups: control, sensory, distraction, and realism. *Control attributes* describe how well the user can interact with and change the virtual or remote environment. *Sensory attributes* are concerned with delivering information about the remote environment to the remote user, allowing the user to move through the remote

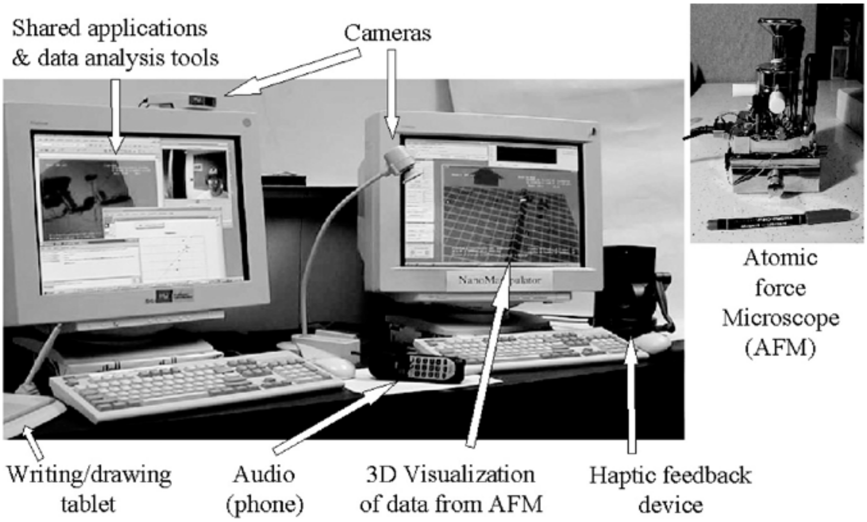


Figure 4-1. The nanoManipulator Collaboratory System.

environment and to actively and purposefully explore it. Just as systems must provide appropriate sensory stimuli, they must also minimize irrelevant external stimuli, or *distraction attributes*, that are not a part of, and particularly are inconsistent with, the stimuli from a remote environment. *Realism attributes* concern how much the remote world is like the natural world, i.e., the degree of consistency between the users' experience of the real world and their experience of the remote place. [Two of Witmer and Singer's realism attributes, meaningfulness of experience and separation anxiety/disorientation, are not dependent on technology design, but rather the application domain. Thus these factors are excluded from the design framework.]

We propose that, when designing a system, these attributes should be considered with respect to their ability to facilitate access to contextual task and process, and socio-emotional information. A table, with each row representing an attribute and each column representing a type of information, can be created to assist in this process (table 4-1). Each blank cell in the table represents something to consider during the design process.

2.3. *Applying the Framework*

2.3.1. *The nanoManipulator Collaboratory*

The framework was applied to the design of a CVE tool, the nanoManipulator collaboratory (figure 4-1). The goal of the nanoManipulator collaboratory is to provide shared remote access to a specialized scientific instrument, called a nanoManipulator (nM), and to support small groups of scientists as they

Table 4-1. Technical design framework.

Technology attributes		Information needed for situation awareness: Contextual	Task & process	Socio-emotional
Control	Degree of control			
	Immediacy of control			
	Anticipation			
	Mode of control			
	Physical environment modifiability			
Sensory	Modality			
	Environmental richness			
	Multimodal presentation			
	Consistency of multimodal presentation			
	Degree of movement perception			
	Active search			
Distraction	Isolation			
	Selective attention			
	Interface awareness			
Realism	Scene realism			
	Information consistent with objective world			

conduct research that utilizes the nM instrument. The single-user nM provides haptic and 3D visualization interfaces to a local (collocated) atomic force microscope (AFM), providing a natural scientist with the ability to interact directly with physical samples ranging in size from DNA to single cells [19, 20].

Hardware elements in the collaboratory system include two PCs. One PC is equipped with a Sensable Devices Phantom™ force-feedback device. This PC and its associated software provide haptic and 3D visualization interfaces to a local or remote atomic force microscope (AFM) and support collaborative manipulation and exploration of scientific data. Scientists can dynamically switch between working together in *shared* mode and working independently in *private* mode. In shared mode, remote, i.e., non-collocated, collaborators view and analyze the same (scientific) data. Mutual awareness is supported via multiple pointers, each showing the focus of attention and interaction state for one collaborator. Collaborators can perform almost all operations synchronously. Because of the risk of damage to an AFM, control of the microscope tip is explicitly passed between collaborators. In private mode, each collaborator can independently analyze the same or different data from stream files previously generated. When switching back to private from shared mode, collaborators return to the exact data they were previously using.

Another PC supports shared application functionality and video conferencing (via Microsoft NetMeeting™) and an electronic writing/drawing tablet. This PC allows collaborators to work together synchronously using a variety of domain-specific and off-the-shelf applications, including specialized data analysis, word processing and whiteboard applications. Video conferencing is supported by two cameras. One camera is mounted on a gooseneck stand so it can be pointed at the scientist's hands, sketches, or other physical artifacts scientists may use during experiments; the other is positioned to primarily capture a head and shoulders view of the user. Collaborators have software control of which camera view is broadcast from their site. A wireless telephone headset and speakerphone connected to a commercial telephone network provides high quality audio communications for collaborators.

2.3.2. Making Design Decisions

Following is a discussion of the technology attributes, and their importance to contextual, task and process, and socio-emotional information. Examples of design decisions that were made using the framework are provided. However, due to space limitations not all design decisions made based on the framework are presented here.

2.3.2.1. Control Attributes

Degree of control: Degree of control refers to the number of elements in the remote and local environments that the user can control and the extent of that control. The more control that collaborators have over the remote environment, the greater their situation awareness. The more control that collaborators have over the local environment, the easier it is to proactively provide contextual, task and process, or socio-emotional information to a remote collaborator. For example, controlling the focus of a local camera on task activities can help increase a remote collaborator's task and process information. The capability to reserve a system for an experiment and to learn about what experiments are planned will increase contextual information. Thus, the nM system was designed with these features.

Immediacy of control: Immediacy of control focuses on system responsiveness. The smaller the delay between initiating a system function (in the local or remote environment) and seeing the function's impact, the greater the sense of presence afforded. This enhances situation awareness by providing the means by which scientists can confirm actions. Software mechanisms that provide feedback as well as efficient algorithms and high network transmission speeds with low latency are typically needed to support this feature.

Anticipation: Anticipation is supported through media richness and consistency. When scientists conduct experiments while working face-to-face, they can recognize the activities being done and the status of those activities, and anticipate subsequent activities. For example, they can gather socio-emotional information, such as frustration and excitement, and anticipate responses, such as encouragement. This was addressed by providing multiple (high resolution and low latency) video camera views of the remote scientists.

Mode of control: In a collaborative system, a person may need to perform an action in one environment in order to cause a responding action in the other environment. Situation awareness is facilitated when that action is natural and similar across environments. For example, actions to change parameters for a data visualization should be exactly the same at all locations.

Physical environment modifiability: Situation awareness is enhanced when collaborators are able to modify the same artifacts in the remote environment as they could if co-located. For the nM system, this includes actions such as pushing, touching and/or modifying a sample during an experiment. In the stand-alone nM system, a scientist could perform these actions locally. The nM collaboratory system allows a scientist to perform the same actions when working remotely.

2.3.2.2. Sensory Attributes

Sensory modality: This attribute implies that systems should avoid forcing users to substitute one sensory mode for another. For example, although audio communication is closely related to collaborative task accomplishment, visual information is a strong source of socio-emotional, task and contextual knowledge [e.g., 21], and its absence in a collaborative virtual environment system would diminish the effectiveness of the system. In the nM system, haptic feedback is also provided when the microscope tip is pushed against a sample.

Environmental richness: Environmental richness implies that systems should gather and display a variety of contextual, task and process, and socio-emotional information at adequate resolution and update rates. This implies a need for high quality video connections that show facial expressions, gestures and local objects; high quality audio connections; and shared applications to increase the richness of the environment. In addition, a "window within a window" to view a collaborator's remote screen while still viewing your local screen may enhance contextual and process knowledge.

Multimodal presentation: When the senses of sight, hearing, smell and touch are stimulated in an integrated and complete manner, situation awareness may be increased. Our observations of scientists engaged in audio–video conferences reveal that the more artifacts brought into the discussion (such as shared drawing tools or shared documents), the more the participants become engaged in the discussions. In a collaborative virtual environment, such senses may include touch integrated with sight and sound. In the nM collaboratory system, haptic feedback is integrated with visual information coming from the visualization screen, contributing task and process knowledge.

Consistency of multimodal presentation: When visual, audio and haptic information are consistent and synchronized, people can more easily understand information. This can increase their confidence in, and use of, the information.

Time synchrony across data presentation modes is an important component of consistent multimodal presentation.

Degree of movement perception: This attribute focuses on self-motion within an environment. Zahorik and Jenison [22] believe that presence is enhanced when one understands the result of an action in an environment, whether that environment is virtual or real. Ideally, a collaborator should be able to clearly see and hear actions that occur in the remote location as a result of a local action.

Active search: Active search capabilities allow users to control sensors at remote locations to obtain desired information. When collaborators can modify sensors to effectively search the remote environment, their socio-emotional, task and process, and contextual information can increase. The scientists interviewed indicated that they would like to have multiple, pre-set video views of a remote collaborator's environment and the ability to modify those views dynamically using remotely controlled pan-tilt-zoom camera mounts and/or automatic tracking cameras. Previous research [e.g., 23, 24] has also illustrated the importance of providing the ability to switch between multiple camera views, as well as repositioning and refocusing cameras.

2.3.2.3. Distraction Attributes

Isolation: Isolation refers to the extent that the user is physically shielded from non-relevant, or distracting, information or activities in the local and remote environments. For example, devices that isolate a scientist from non-relevant aspects of the local environment can enhance his or her ability to gather and understand information from a remote environment. An example is the use of headphones to reduce ambient noise in the local environment so that a user may fully concentrate on the interaction with a remote collaborator. However, observations of scientists using desktop audio–video conferencing tools such as NetMeeting™ also show that they like to hear auxiliary conversations arising in the local and remote environment to increase their contextual situation awareness. A solution is to provide options regarding audio headsets and audio speakers.

Selective attention: This attribute focuses on the extent users ignore non-relevant information. For example, a collaborator's willingness to ignore distractions in the local environment should enhance their awareness of the remote environment. This is a psychological issue.

One method to capture and focus the attention of another person, particularly with respect to information on a monitor, is through pointing. We frequently observed scientists pointing to computer screens with their mouse pointer, fingers and pens to selectively focus a collaborator's attention. Thus, the nM system was designed to enable each scientist's pointer to be viewable by all collaborators.

Interface awareness: This attribute focuses on human–computer interface design. The human–computer interface for all types of information should be natural and easy to use. This has been previously proposed for all types of systems, and widely discussed in the human computer interaction literature [e.g., 25].

2.3.2.4. Realism Attributes

Scene realism: Scene realism, or the realistic rendering of the remote environment, addresses the validity of information from the remote environment used to develop situation awareness. Scene realism can be developed using real-world content, such as video, and simulated content, such as computer animation or graphical representations. It is affected by camera resolution, light sources, field of view, as well as the connectedness and continuity of information being presented [3]. Emerging technology, such as 3D telepresence, has the potential to increase scene realism.

Consistency of information with the natural world: Information about the remote environment provided by the system should be consistent with information learned through first-hand experiences. For example, if a scientist had previously visited a collaborator's lab then information about the lab, e.g., a floor plan, provided by the system should be consistent with the scientist's existing knowledge of the lab. Even when scientists have not had the opportunity to visit a collaborator's environment, they have expectations regarding that environment based on their previous experiences. Information provided by the system should be consistent with these expectations.

2.4. Evaluating the Design Framework

2.4.1. Controlled Experiment Evaluation

To investigate the validity and utility of the framework, we can evaluate the systems it helps to create. This is an indirect measure, and conclusions from the evaluation should be interpreted with caution.

The evaluation conducted was a repeated measures, or within-subjects, controlled experiment comparing working face-to-face and working remotely, with the order of conditions counterbalanced. The hypotheses followed previous research [e.g., 26] and would predict that task performance and perceptions of the system when using it to collaborate across distances would be lower because collaborators would lack the richness of collocation and face-to-face interaction, including multiple and redundant communication channels, implicit cues, and spatial co-references. This lack of richness is often thought to impair situation awareness and subsequently have a negative impact on task performance and perceptions of technology.

In the evaluation, twenty pairs of study participants (upper level undergraduate natural science students) conducted two realistic scientific research activities each requiring 2–3 hours to complete. Ten pairs of study participants worked face-to-face first and, on a different day, worked remotely (in different locations). Another ten pairs worked remotely first and, on a different day, face-to-face. When face-to-face, the participants shared a single nM collaboratory system; when collaborating remotely, each location was equipped with a complete nM collaboratory system.

The scientific research activities completed by the participants were designed in collaboration with natural scientists. The tasks were activities the scientists actually completed and documented during the course of their research. To complete the tasks the participants had to engage in the following activities typical of scientific research: operate the scientific equipment properly; capture and record data in their (electronic) notebook; perform analysis using scientific data analysis software applications and include the results of that analysis in their notebooks; draw conclusions, create hypotheses and support those hypotheses based on their data and analysis; and prepare a formal report of their work.

Task performance was measured through graded lab reports. The information participants were asked to provide in the reports mirrored the information scientists record in their lab notebooks. Each pair of study participants collaboratively created a lab report under each condition, generating a total of 40 lab reports; 20 created working remotely and 20 created working face-to-face. In addition, each participant was interviewed after each session. The post-interviews focused on participant's perceptions of the collaboratory system and their work patterns. The lab reports were graded blindly, and the post-interviews were analyzed using both open and axial coding [27]. During open coding, the interviews were read thoroughly and carefully and coding categories, or frames, were identified. During axial coding, the final step, all interviews were re-read and analyzed using the coding categories. Additional details regarding the experiment can be found in Sonnenwald *et al.* [28, 29].

2.5. Evaluation Results

The average lab report scores for the first task session were identical (70/100) for both the face-to-face and remote condition (table 4-2). Although a null result statistically, the comparable scores between the two conditions on the first task are encouraging.

Data analysis further indicated that in this study collaborating remotely first had a positive effect on the second, face-to-face interaction. Using a multivariate analysis of variance (MANOVA) test, the differences in scores for the face-to-face and remote conditions were not statistically significant at the

Table 4-2. Graded lab report statistics.

	Graded lab report scores (max. score = 100)									
	Collaborated FtF first ($n = 20$)					Collaborated remotely first ($n = 20$)				
Condition	**Mean**	**SD**	**Max**	**Min**	**Range**	**Mean**	**SD**	**Max**	**Min**	**Range**
Face-to-face	70.0	16.75	88	42	46	86.4	10.52	98	70	28
Remote	75.1	10.49	89	56	33	70.0	8.89	80	55	25

$p \leq 0.05$ level. However, when order is taken into account, participants who collaborated remotely first scored significantly higher on task 2 than did those who collaborated face-to-face first (df = 1, $F = 9.66$, $p < 0.01$). Due to available resources, we did not study the cases where participants completed the two task sessions under the same condition, e.g., both face-to-face or both remotely, and thus we are unable to eliminate the possible effect of task differences between the two sessions. However, these results suggest that the scientific tasks conducted remotely were of similar quality as those conducted face-to-face. The results further suggest that working remotely, each scientist having full access to the system at all times, may facilitate their learning about the system and scientific tasks at hand, and this possibly influenced subsequent scientific work in a positive way.

Participants' perceptions regarding control, sensory, distraction and realism attributes of the nM collaboratory system emerged from the analysis of the interview data. Participants requested several features the framework predicted would be important but were not implemented due to technical constraints; they reported negative perceptions of features that did not conform to the framework; and they reported positive perceptions regarding features the framework predicted would be important and which were implemented.

An example of features suggested by the framework and not implemented are automatic tracking or remotely controlled pan-tilt-zoom camera capabilities. These features were originally suggested when considering the active search (sensory) attribute. In post-interviews, several participants requested these features, e.g., one participant commented: "We didn't want to waste our time always adjusting the camera . . . have the camera follow you."

Similarly, a participant requested the capability to view a collaborator's remote screen while viewing their local screen, a feature suggested by the environmental richness (sensory) attribute. The participant explained:

> It would be good . . . if you're both in your private state [if] you could each see what the other's doing . . . if you have two different ideas of how to go about something, then you each can try it and see if you get to the same point without having to flip back and forth between [states.]

Study participants also reported negative perceptions of features that did not conform to the framework. For example, the consistency of multimodal presentation (sensory) attribute emerged as problematic for study participants. In particular, the video would “freeze” and be out of sync with the audio. Participants commented: “The video window froze and that was slightly aggravating.” And, “[the video] kept stopping . . . his picture would freeze . . . the audio would be far ahead of where the video was.” Other participants commented: “[The video] was extremely helpful . . . I couldn’t really describe [a scientific phenomenon] as well as I could just move my hands . . . in front of the camera.” Also, “I liked the video conferencing . . . I like seeing people as I interact with them and they react.”

Thus the video had utility, but the multimodal presentation of the video and audio was not effective. This particular problem can be addressed through improvements in networking infrastructure and algorithms that provide faster and more reliable video transmission and coding and decoding. These types of issues are typically outside the scope of any single collaborative virtual environment, yet they impact users’ perceptions of the environment.

Participants also reported positive perceptions of features that were suggested by the framework and implemented. This was particularly evident with respect to the mode of control attribute. In the nM visualization software component, all users can execute system functions concurrently. Thus the model of control is identical when working individually and collaboratively. Participants commented: “The best thing was . . . the ability to work on the same thing at the same time with the nM.” And, “[we] never fought over the nM because . . . both of us [could] use it at the same time.”

In comparison, the mode of control differed when using the off-the-shelf shared application software, NetMeeting™. In NetMeeting users were required to explicitly take control of a shared application by double-clicking on the application window. One participant explained:

> [It] became exceedingly frustrating . . . to share control . . . When I wanted to do something and my partner wanted to do something at the same time, we . . . went back and forth double-clicking to gain control, and . . . it took us a few seconds to even acknowledge that. Essentially . . . we were fighting over control.

Task performance as measured by graded lab reports and perceptions of the system as discussed in post-interviews help demonstrate the appropriateness of the features suggested by the framework and provide some insights regarding the validity and utility of the framework.

2.6. Limitations of the Technical Design Framework

It may not be possible or necessary for a CVE to equally support the acquisition and dissemination of contextual, task and process, and socio-emotional information. Emphasizing one or two types of information may help prioritize design and implementation decisions. The framework presented in this chapter identifies types of information to support situation awareness but it does not prioritize them. For example, due to technical and budget constraints, some design decisions suggested by the framework could not be implemented.

Furthermore, it is not known if the list of attributes in the framework is exhaustive. Additional categories of attributes and additional attributes within categories may emerge as technology and our understanding of human information processing evolves. Future research is needed to investigate guidelines for prioritization.

3. An Organizational Design Framework

As mentioned earlier, technology adoption and use is not solely influenced by features of the technology but also by organizational structures and practices. These organizational factors become increasingly complex when a large group of scientists needs to collaborate across distances, not only by sharing data and instruments but also by actively discussing ideas, identifying problems and solutions. Increasingly scientific problems require such collaboration among groups of scientists, many of who often work in different organizations, geographic locations and disciplines. How can such scientific research be best organized?

To address this issue, a two-year field study of a large, distributed group of scientists was conducted. Analysis of organizational documentation, sociometric surveys, interviews and observation data suggests that a new type of research and development (R&D) organization, called the conceptual organization, is one solution [9, 30].

3.1. Research Methodology

The conceptual organization framework is based on an in-depth two-year case study of an R&D centre in the US. Initially, the Centre had approximately 30 faculty scientists and 82 students and postdoctoral fellows, and three full-time staff members physically located at four different universities in the US. Membership has changed over the years, and after 3 years there were approximately 45 faculty scientists, 70 students and postdoctoral fellows and three full-time staff members located at five US universities. The R&D centre was first funded late 1999, with a five-year $15 million dollar commitment from a national funding agency with matching support from several participating universities, corporations and a non-profit foundation.

The case study began during the beginning stages of the centre and continued for 2 years. While conducting the case study, the author was a participant observer, having both complete and peripheral membership roles. As a complete member, the author had functional, in addition to research, roles in the research setting. The author served as the Centre Coordinator of Social Science Research Efforts and a member of the Centre management team. She actively participated in the management meetings, contributing to discussions and participating in decision-making. However, when the meetings and decision-making focused on research in natural science and engineering topics, topics not in the author's areas of expertise, she assumed the role of a peripheral member. She observed the activity, taking notes and audio-recordings, and occasionally discussed events and outcomes with meeting participants but she did not actively participate in the discussions and decision-making. Seventy-three management team meetings were held during the two-year study, and the author observed and participated in these meeting. The author was also a peripheral member participant in centre-wide weekly research meetings, generally observing discussions and only completely participating when discussions regarding collaboration and collaboration technology took place. Centre members were made aware of the author's roles.

Observation data included transcribed audio-recordings of meetings, video-recordings of videoconferences, meeting and centre documentation and researcher notes. These data were analyzed in the ethnographic and grounded theory traditions [7, 8]. Using semantic content analysis [31] patterns and meanings behind the observations were sought. That is, a theoretical framework was not imposed on the data a priori but rather the data were thoroughly analyzed for patterns within the data and the meaning of those patterns. Results were subsequently shared with several centre members (informants) and their feedback was incorporated.

Two sociometric surveys were also conducted to provide quantitative data regarding collaboration within the centre. The surveys investigated current collaboration among centre members, and took place approximately 12 and 24 months after the centre was established. Response rates for the two surveys were 68% and 73%, respectively. The data were analyzed using sociometric techniques [32] to investigate the number of collaborations among scientists and students, collaborations across universities and changes in collaborations over time. To further investigate collaboration and organization effectiveness, co-authorship of journal publications and research funding data over 3.5 years was collected and analyzed.

Although the technology used in the case study was not a CVE in the strictest sense, components of the technology used had features, such as synchronous applications and voice over IP, that overlap with CVEs. Furthermore, we know that issues regarding the use of CVEs are often organizational in nature. Getting people to participate in CVEs, sharing information equally, are typically more serious and harder to resolve than technical issues, and thus this case study has relevance for CVEs.

3.2. The Conceptual Organization

3.2.1. Definition of a Conceptual Organization

A conceptual organization is a new type of research and development (R&D) organization that has emerged to facilitate collaboration among large groups of geographically distributed scientists in order to tackle large, complex and challenging problems of national and global importance. Its purpose is to discover solutions, quickly and effectively contributing to relevant dynamic knowledge bases and meeting diverse stakeholder needs with minimum capitalization and start-up costs. It has a conceptual organizational structure in addition to a physical structure, both of which are interwoven across other external organizational structures. It has few employees in the traditional sense; most members are scientists who join the organization because they wish to contribute to its vision and goals. These scientists are typically employed full time by universities or an R&D laboratory, and have a part-time affiliation with the conceptual organization. The conceptual organization provides a management structure and organizational practices that facilitate collaboration among members working towards its vision and goals. The power of the conceptual organization is primarily integrative in nature.

3.2.2. Synthesis of Multiple Organizational Forms

The conceptual organization has characteristics, or features, in common with traditional organizations, invisible colleges, scientific collaboratories and virtual teams. For example, similar to traditional R&D organizations, conceptual organizations need physical space, including offices for researchers and staff as well as laboratories to house specialized scientific equipment and conduct scientific experiments. For the conceptual organization, however, these needs are often negotiated and met through relationships with other organizations, such as universities, with which their members are affiliated. Conceptual organizations and traditional R&D organizations also have aspects of management in common, such as a management team that includes directors and an external advisory board who reviews the organization's progress. However, as discussed below the management structure of a conceptual organization has a more diversified membership.

Similar to invisible colleges [33], members elect and are selected to participate in a conceptual organization based on their knowledge and expertise. However, in a conceptual organization the selection and participation process is more formal than in an invisible college. Participation in an invisible college is often a matter of knowing its members and thereby gaining entry and acceptance through interaction with them. In a conceptual organization, there is a formal invitation or application process in addition to the informal process.

This is because conceptual organizations typically provide research funding or other costly resources for its members whereas invisible colleges do not.

A scientific collaboratory is a laboratory without walls [34]. A conceptual organization has many characteristics in common with a collaboratory, e.g., a conceptual organization may provide remote (electronic) access to data sources, artifacts, tools and experts. However, the primary goal of a conceptual organization is to address a specific, complex and challenging research issue, while the primary goal of a typical collaboratory is to provide remote access to data sources, artifacts, tools and experts to facilitate scientists' individual research initiatives. The nature and emphasis of these goals are slightly different, although the implementation of these may have aspects in common. For example, a conceptual organization and collaboratory may use similar technology, such as a CVE, to facilitate collaboration across geographic distances. However, a conceptual organization focuses on, and is evaluated with respect to, the results of its research and educational activities; whereas a collaboratory typically focuses on, and is evaluated with respect to, the utilization of its resources.

Virtual teams are groups of individuals who may not meet face-to-face but work together towards a common goal at a distance. Often the team is brought together to address a specific goal and disbanded after that goal is met or when the goal is no longer deemed important [35]. In corporate settings, these teams may cross organizational boundaries and include individuals from different corporations. A conceptual organization may encourage teams to form to address goals related to the vision, and some of these teams may be virtual. For example, a virtual team could be formed to help coordinate all proposed research efforts going on in two locations on a particular topic. However, a virtual team is more limited in scope and size than a conceptual organization.

Thus, a conceptual organization has characteristics in common with traditional organizations, invisible colleges, collaboratories and virtual teams. However, it also appears to be a unique organizational form. As described below, its management structure, use of organizational power, types of stakeholders, benefits and challenges combine to represent a new organizational form that facilitates collaboration across organizations and geographic distances.

3.3. Description of a Conceptual Organization

As discussed previously, the conceptual organization has characteristics in common with other types of R&D organizations, and it employs an innovative combination of organizational practices found in them. These organizational practices are presented in this section to provide a detailed portrayal of a conceptual organization.

3.3.1. Management Structure and Organizational Membership

The management structure of a conceptual organization includes a director who sets the overall prioritization for the centre and is responsible for leading the strategic vision and planning process. In addition, the director takes a lead in organizing the research as well as the dissemination of the research in "real time" by organizing the centre-wide group meetings. As director, this person also leads the interactions with the external stakeholder groups, such as the national funding agency, an external advisory board, affiliate university administrations and the media. In addition to these responsibilities, the director teaches and conducts research.

A conceptual organization typically also has a co-director and a deputy director. The co-director's primary responsibility is financial leadership and leadership in strategic planning. The co-director is also the leader of the external industrial affiliates group and conducts research. In the organization studied, the co-director was a close research collaborator to the director and was essentially interchangeable with the director in many functions. The deputy director is a position created explicitly to help with the numerous administration requirements associated with the centre. The deputy director plays an organizational lead position for the strategic plan and its implementation and accountability. The deputy director is also responsible for leading the generation of the annual report and overall compliance with the cooperative agreement between the universities and the funding agency. In a supporting role, the deputy director also assists with the numerous outreach programs and represents the organization at external venues on numerous occasions.

Thus the directors share in the responsibility of creating and communicating the vision of the organization, as well as administrative tasks. This helps to alleviate common burnout, which often leads to a degradation of management's ability to create and maintain a strategic vision and vibrant research program.

To further broaden participation in organization management, the directors are assisted by a management team that includes a site coordinator for each participating university, a coordinator of collaborative efforts, a higher education outreach coordinator, a kindergarten through 12th grade (K-12) education outreach coordinator, a scientific program committee and an office manager. Site coordinators handle location-specific administrative issues, ranging from reserving a videoconference room for weekly meetings to distributing allocated budget funds. The coordinator of collaborative efforts manages socio-technical activities to support collaboration within the centre and coordinates social science research done in the centre. The higher education and K-12 outreach coordinators oversee the educational outreach activities done by centre members and their staff. The scientific program committee provides input regarding natural science research and development.

The participation of site coordinators, i.e., representatives from each physical location, provides ongoing dialog about challenges, progress, perceptions and ways of working at each location. It is a way to interweave the conceptual organization among multiple physical locations and the external organizations at those locations. It eliminates the need for individual scientists to take sole responsibility of coordination and cooperation between their local and remote organizations (in this case, between their local university and a conceptual organization). It also facilitates learning about different ways of working and collaborative problem solving when members from different locations suggest how practices at their location may solve problems at other locations. For example, one team member suggested a possible solution to a colleague at a different location:

> Another thing you can do . . . to magnify your undergraduate help is that you can have undergraduates getting paid for a certain amount of their research but then getting credit for a certain amount, so that you only have to pay for part of it . . . We pay [our undergraduate students], but . . . we also want them to take two semesters of [research credits].

Similarly, the participation of K-12 outreach, social science, minority and technical program coordinators on the management facilitated coordination and collaboration among these diverse domains.

Scientists and students in a conceptual organization typically have a primary affiliation with the university at which they are physically located. They become members by proposing research projects and activities that would help the conceptual organization achieve its vision, mission and goals. Faculty scientists (current and potential members) may submit proposals that outline research projects that, ideally, support the conceptual organization's vision and mission. The proposals are typically reviewed and discussed by members of the conceptual organization's management team. Primary evaluation criteria may include: fit to strategic plan, potential impact and scientific merit. Secondary evaluation criteria may include: collaboration plan, K-12 outreach record and plan, and outside funds attracted.

3.3.2. Power within the Conceptual Organization

Boulding [36] describes three types of organizational power: destructive, economic and integrative. Destructive power, the power to destroy things, can be used for carrying out a threat and as a prelude to production, where things are destroyed or altered to make way for production. An example of destructive organizational power is the firing of employees who are seen as resisting change in an organization. Economic power is used in all organizations. It involves the creation and acquisition of economic goods, including intellectual property, through production, exchange, taxation or theft. Integrative power involves the capacity to build organizations, inspire loyalty, to bind people together and to develop legitimacy. It has a productive and destructive aspect. In a negative sense it can create enemies and alienate people. All organizations have some integrative power or they could not survive. Some, however, rely on integrative power more than others; these include religious organizations, political movements, volunteer organizations and clubs. Their existence and growth are influenced by the extent to which the objectives of these organizations match the dynamic value structures within a larger population.

The conceptual organization uses a combination of integrative, economic and destructive power; however, its primary source of power appears to be integrative. It solicits funding and participation based on its vision, mission and goals. That is, it attracts funding from corporations, government agencies and other institutions by convincing them that its vision, mission and goals are valid and achievable. A conceptual organization cannot promise an economic return on investment although it may offer some hope to funding corporations that it will effectively educate students who may become future employees and generate patents and other knowledge that may have economic value. Conceptual organizations attract scientists and students similarly, i.e., by convincing them that the organization's vision, mission and goals are exciting and participation in the organization can provide great personal satisfaction.

A conceptual organization may use integrative power in developing its vision, mission and goals. For example, when describing the process of developing a vision, the executive director commented:

> It's intended to be an inclusive process. We've included most of the [faculty] here in the centre in this process. Certainly our external advisory board had a part to play. It's iterative . . . We made our first draft of the vision, mission and goals, and reviewed those with [the faculty] . . . We then reviewed those with [industrial partners] and with our external advisory board. We got their input, what they thought we should be doing in a strategic direction . . . we integrated these comments.

A conceptual organization augments integrative power with economic power, providing some funding to scientists and students. For example, in the conceptual organization studied scientists typically receive one month's summer salary, funding for one graduate student or 50% funding for a postdoctoral fellow, up to $4,000.00 for supplies, and $500 for travel. However, these amounts are by themselves not necessarily sufficient to attract and retain high-caliber scientists who often receive government and corporate funding in much larger amounts. A vision that scientists believe in is also required.

As in any organization, destructive power is used when members do not meet expectations or keep commitments. This may be implemented in the conceptual organization through decisions not to continue funding scientists whose work is judged not in alignment with the conceptual organization's vision, mission and goals. For example, during a meeting deciding funding, participants supported and criticized proposals using comments such as: "This [proposed project] was not the lowest on my list, but I really miss the connection to objectives, goals, mission, etc. here. I could not see where this is going to lead."

These decisions, however, should be reached through integrative power. In the conceptual organization studied, the review was done collaboratively with the scientific program committee, consisting of a lead scientist from each location and the conceptual organization's director, co-director and deputy director. This group also developed the call for proposals. The call included the conceptual organization's vision, mission, goals and critical needs as well as the proposal process and evaluation criteria. The process included a preliminary proposal in which faculty were requested to provide a title and a brief statement of research objectives (six to eight lines in length.) The committee provided feedback to the faculty on their preliminary proposals. The preliminary proposals were: "A mechanism for earlier dialogue . . . The benefits are . . . to attempt to avoid excess overlap [between projects]; . . . to identify opportunities for collaboration . . . not only within a given university, but also between universities; . . . to identify any unmet needs."

Thus, through interaction with faculty and collaboration among management team members, integrative and destructive powers were used.

3.3.3. Stakeholders of a Conceptual Organization

All organizations, including conceptual organizations, have stakeholders, i.e., individuals or organizations who have a stake in a given organization's success. The case study suggests that stakeholders in a conceptual organization include society, scientific disciplines or paradigms, government funding agencies, businesses and academic institutions.

It appears that society is a primary stakeholder of a conceptual organization's vision in that society legitimizes the government, corporations and institutions that ultimately fund the conceptual organization. For example, the vision of the conceptual organization in the case study supports green chemistry. Green chemistry in general is currently valued by American society. The need to develop new processes and products that do not pollute the environment are recognized as important throughout American society. Even with this general support, results and justification of the government's investment is needed. For example, the conceptual organization directors have made presentations to the US Congress and met with Senators and Representatives. These activities are necessary in part because if citizens in democratic societies do not approve of a conceptual organization's goal, they may organize to limit its funding.

Scientific disciplines are also stakeholders interested in the mission of a conceptual organization. Disciplines typically wish to see knowledge created and students trained in certain scientific areas. This is motivated by collectively held belief systems and yearning for self-preservation and perpetuation of a discipline or scientific paradigm [37], and the mission of a conceptual organization has the potential to contribute to the growth of knowledge in particular scientific disciplines and/or paradigms.

Government funding agencies, businesses and academic institutions are stakeholders who are typically interested in a conceptual organization's vision, mission and goals. For these stakeholders the vision and mission is necessary but not necessarily sufficient. They are also interested in how the vision and mission will be achieved and measured, i.e., the organization's goals. They are typically concerned about justifying their investment in the conceptual organization to their stakeholders, e.g., federal and state governments, and upper management. For example, the conceptual organization studied produced a 226-page report detailing its activities and accomplishments during the preceding 12 months to help justify its government funding. Quantitative measures reported included publications, presentations, patents, supplemental funding, students supported, students graduated, K-12 and minority students reached through outreach activities, and K-12 teachers reached.

Businesses do not appear to seek a return on investment from a conceptual organization in the same way when investing in a company because they anticipate other benefits. For example, in a survey of 249 corporations who participated in industrial-university research centers, Gray, Lindblad and Rudolph [38] found that professional networking, including enhanced student recruitment and improved cooperation with scientists, was the primary factor influencing corporate decisions to maintain their relationship with and support of an industry-academic centre. Quality of the research and technical benefits, such as commercialization impact, were not found to impact corporate support of the centers. A conceptual organization should hold bi-annual or annual meetings that showcase students for its external industrial affiliates group.

3.3.4. Knowledge Management: Interaction among Members

A conceptual organization must utilize CVE and other technology as mechanisms to support its vision and mission, or incur expensive monetary and temporal travel costs. For example, in the conceptual organization studied video conferencing and shared electronic whiteboards were used for organization-wide meetings, weekly centre-wide research meetings, and weekly project team meetings. Organization-wide meetings were held relatively infrequently (e.g. once every 6–8 months); these meetings included all members at all universities and have been used to share information among all members. For example, a conceptual organization-wide meeting was held that introduced the organization's mission, management structure and conceptual organization-wide activities several months after the organization was established.

In the organization studied, centre-wide research meetings were held weekly; all members were invited to attend these meetings, however, students were required to attend. Each meeting typically lasted 1.5–2 hours, and at each university participating in the centre, small groups of 2–25 students and faculty would be in attendance. During the meetings, students and postdoctoral fellows presented and discussed their work at least once per year, responding to questions and comments from other participants.

The format and technology used in these meetings evolved over time. New social protocols, including the introduction of sharing interpersonal information, were introduced to compensate for constraints imposed by the technology. New operations protocols to help reduce technical problems were developed and implemented working with centre members and technical staff [39].

The video-conference technology used for centre-wide meetings included: a large electronic whiteboard and PC running shared application software to display slides and create and capture notes in real time; two large (120 inch) display screens that showed an overview shot of participants in each location and multiple views of one or several individuals in each location; microphones for each participant to capture and broadcast anything they wish to say; stereo audio speakers to enable each participant to clearly hear what is said by others; multiple cameras at each location to capture views of the audience, especially the person currently speaking; and a combination of networks such as ISDN/H.320, local state government analog video network, and video over IP (internet protocol) and required muxes. The technology used for project team meetings was similar but smaller in scale, e.g., it included the large electronic whiteboard but not the large display screens. Although these are not CVE tools in the strict sense, there are many similarities with CVEs if used for large meetings.

Initially, the technology increased the formality of the meetings. Students were concerned about using technology that was new to them and discussing their work with such a large audience; thus they initially prepared more before the meetings and gave formal talks. The initial meetings were also plagued with technical problems and this frustrated many participants. However, after these issues were resolved through new social and technical protocols [39], the meetings became very interactive and increased members' awareness of one another's work. In particular, students received important feedback on their work, and faculty learned about ongoing research efforts. The latter was achieved through minimal effort. If a presentation and discussion was not relevant to a member's work, the member could unobtrusively do other work during the discussion. However, problems originating from a lack of trust among members can still occur and need to be managed. Discussions regarding this issue can be found in Sonnenwald [30].

3.4. Benefits of a Conceptual Organization

A benefit of a conceptual organization is its ability to contribute to and respond to dynamic needs for new knowledge. This is achieved through multiple mechanisms. One such mechanism used is the dynamic incorporation of scientific experts in emerging relevant areas. For example, the centre has a call for proposals on a two-year cycle. This enables the incorporation of new scientists and research topics every other year. Another mechanism that supports the dynamic incorporation of scientific experts and emerging relevant areas is "seed funding" which is available on a yearly basis. In other R&D organizations, such efforts have been called "skunk works" but these are limited to existing organizational members and are hidden from other parts of the organization. In conceptual organizations, such efforts can be proposed by existing or potential members. These efforts are not hidden from view, and may be fully integrated in the organization through activities such as review meetings. Thus, all results are shared among centre members so everyone can learn from them. A third mechanism is matching funding. Scientists can use funding from the conceptual organization as matching funds in other grant proposals that may include additional scientists and students as well as emerging relevant research topics. This brings additional resources to bear in addressing the vision, mission and goals of the conceptual organization.

An additional benefit provided by the conceptual organization appears to be lower capitalization or start-up costs. These lower costs are achieved through the re-use of existing physical spaces and equipment at the associated universities and organizations, limited term and partial commitment to members and the inclusion of students and postdoctoral fellows. A conceptual organization may rely on space and equipment at its associated universities to support the research being conducted by its members, scientists and students. In return, the organization may purchase new equipment that scientists and students at the universities but not associated with the conceptual organization may also access. The conceptual organization also provides funding to enable students to attend the universities. A limited (2 or 1 year) and partial commitment to scientists (only one month summer salary is typically provided to scientists) further reduces the start-up costs of a conceptual organization. Of course, the inclusion of students and postdoctoral fellows who are by definition limited term also reduces or limits start-up costs for it as well.

A further benefit of a conceptual organization may be found in its ability to meet diverse stakeholders' and members' needs. As discussed previously, a conceptual organization's stakeholders can include society, scientific disciplines or paradigms, government funding agencies, corporations and academic institutions. This diverse and important set is an outgrowth of a variety of political, social and economic forces; no other type of R&D entity has a similar broad set of stakeholders. Furthermore, the infrastructure at academic institutions is typically based on department and disciplinary boundaries with fierce competition for resources, authority and territory [40]. This is often a barrier when addressing large complex and challenging problems of national and global importance where the best scientists irrespective of discipline, department or institution affiliation are required.

3.5. Evaluation of the Conceptual Organizational Framework

To evaluate the effectiveness of the conceptual organizational design framework, data regarding collaborations, co-authorship, and funding in the organization studied were collected and analyzed.

As previously mentioned, two sociometric surveys were conducted asking organization members to identify other members they were currently collaborating with. The first survey took place 1 year after the conceptual organization was established; the second took place 1 year later. The number of collaborations reported among faculty scientists increased from an average of 2.37 per scientist to 3.36 per scientist; a 41.7% increase from year 1 to year 2 (see table 4-3.) A larger increase was seen in the growth of collaborations among scientists at different universities than among scientists at the same university (61.1% versus 27.6%). This indicates that collaboration among scientists within the organization developed across universities (and distances).

Another effectiveness measure is co-authorship of journal publications. Table 4-4 shows the number of co-authored and single-authored journal articles published over 3.5 years. Not surprisingly, there were fewer articles published in year 1 due to the start-up time lag that naturally occurs in research. From year 2 on, there were more articles published by co-authors from different universities than published by authors at the same university or published by single authors. On average, 48% of the total articles published were by authors from different universities. These data further suggest that the organizational structure and practices within the conceptual organization studied facilitated collaboration.

A third measure of research effectiveness is the ability to attract funding. The centre studied was initially successful in obtaining research funding from a government agency. Over the next 3 years, the organization also procured $1 million per year in funding from the participating universities, $1 million per year from other sources, e.g., corporations and non-profit organizations. The participating faculty also procured an additional 128 grants for a total of $47 million. These data combined with data regarding the quantity of self-reported collaborations and co-authorship trends provide insights regarding the utility of the conceptual organizational framework.

Table 4-3. Reported collaborations in the centre.

Type of collaboration	Collaborations: After 1 year — Total	After 1 year — Per person	After 2 years — Total	After 2 years — Per person	Change between 1st and 2nd year — Total	Change — Per person	Change — % change per person
Among all scientists	71	2.37	148	3.36	+77	+0.99	+41.7
Among scientists at the same university	37	1.23	69	1.57	+32	+0.34	+27.6
Among scientists at different universities	34	1.13	80	1.82	+44	+0.69	+61.1
Among all scientists & students	191	1.71	223	1.96	+32	+0.25	+14.6
Among scientists & students at the same university	42	0.38	68	0.60	+26	+0.22	+57.9
Among scientists & students at different universities	139	1.24	155	1.36	+16	+0.12	+9.7

Table 4-4. Co-authorship of journal articles by centre members.

Publication year	Co-authors from: Same university #	Same university %	Different universities #	Different universities %	Single author #	Single author %	Total
Year 1	3	75	1	25	0	0	4
Year 2	2	14	10	71	2	14	14
Year 3	10	34	12	41	7	24	29
First 6 months of year 4	8*	28	16**	55	5	17	29
Yearly averages	*38*		*48*		*14*		

* Includes 3 published and 5 submitted.
** Includes 5 published and 11 submitted.

3.6. Challenges for a Conceptual Organization

One challenge for a conceptual organization involves reconciliation with existing academic and disciplinary cultures. As discussed, a conceptual organization is embedded within existing academic and disciplinary cultures; its members must also be active and accepted participants in their university departments and disciplines. Conflict among these can emerge with respect to job performance evaluation and career paths.

For example, one critical job performance evaluation in research universities in the US occurs when an assistant professor is reviewed for tenure and promotion to associate professor. Typically, an assistant professor is required to leave the university where they are employed if tenure is not granted. Decisions regarding tenure are initially decided by colleagues in the same department and discipline (who may not be members of the conceptual organization). These decisions are based on several evaluation criteria, including: an individual's ability to establish a research agenda or vision; an individual's record of research funding; and recognition of the individual's research contributions in the larger academic community. All of these may be negatively perceived in cases where an assistant professor is a member of a conceptual organization. For example, an assistant professor's research agenda or vision may be perceived by colleagues as lacking originality or insight because it is linked to the conceptual organization's vision, which would not be credited to the assistant professor. Research funding through a conceptual organization does not have the same requirements or review process as found with national and other funding agencies, and thus may not be as highly valued. Furthermore, a conceptual organization's vision may require expertise from multiple disciplines. When an assistant professor collaborates with others not in the same discipline, it can limit the opportunity for colleagues in her or his discipline to learn about and understand the assistant professor's research contributions. This lack of knowledge or understanding may also contribute to a negative evaluation. Thus, the tenure evaluation process may discourage or even conclude an assistant professors' participation in a conceptual organization, with negative consequences for both the assistant professor and conceptual organization.

Associate and full professors must also be active participants in their local university departments and discipline. Activities encouraged by a conceptual organization, e.g., participation in weekly video-conference meetings providing students at other universities feedback on their research and helping a colleague at another university set up research lab equipment, may not be encouraged or valued by one's local university department and colleagues in the same discipline. Individuals have time constraints and, as a result, a faculty member may find they must make difficult choices between contributing to a local department and their career versus contributing to a conceptual organization.

4. Discussion

Scientific collaboration is complex, yet critical to addressing complex problems that cannot be solved by any one individual, discipline or organization. Collaborative virtual environments can facilitate scientific collaboration but care should be taken when designing both the technology and the organization for which the technology is intended. Traditionally technology design occurs independently of organizational design. Indeed, research in these areas also typically occurs in different disciplines, e.g., computer science and business. However, in practice technology and organizational design are interdependent; each influences and helps shape the other. This chapter is a first effort to consider both technology and organizational design for scientific collaboration.

To address the technical design challenge, we built on previous research in situation awareness as well as interviews and observations of scientists to illuminate the complexity of situation awareness in scientific research and to propose that contextual, task and process, and socio-emotional information is needed to create and maintain situation awareness. Research in virtual reality systems suggests control, sensory, distraction and realism attributes of technology contribute to a sense of presence. We suggest that consideration of these attributes with respect to contextual, task and process, and socio-emotional information provides insights to guide design decisions.

The framework was used to guide decisions regarding technology to support situation awareness for the nM collaboratory system. As a result, the nM collaboratory system includes: consistent shared and private work modes, or spaces; the ability to dynamically switch between those shared and private work modes; the ability to customize an individual view of a shared work space; and multiple pointers that indicate each collaborator's focus of attention, interaction mode and actions simultaneously to all remote sites when in shared mode. The results of a repeated measures, or within-subjects, controlled experimental evaluation of the nM collaboratory system help illustrate the validity and utility of the framework, yet should be interpreted with caution because they are indirect measures of validity and utility.

In addition to considering new frameworks to support technology design, it is important to consider organization design, i.e., how the structure and practices of an organization can be designed to facilitate scientific collaboration. To address the organizational design challenge, we built on previous research as well as an in-depth two-year case study of a successful group of approximately

100 geographically distributed scientists. The resulting framework is the conceptual organization. A conceptual organization should have a long-term vision that addresses large complex and challenging problems of national and global importance. Its goal is to works towards this vision, quickly and effectively contributing to relevant dynamic knowledge bases and meeting diverse stakeholder needs with minimum capitalization and start-up costs. To achieve this, it has an explicit conceptual organizational structure in addition to a physical structure, both of which are interwoven across other external organizational and physical structures. Conceptual organizations engage scientists through the appeal of their vision and management structure and practices that encourage and facilitate collaboration. Challenges for conceptual organizations may arise due to conflicts with traditional norms and practices embedded in university and R&D settings. Social network, co-authorship publication and funding data from the case study setting provide initial evidence of the effectiveness of the conceptual organizational framework.

Additional research, utilizing both the technical and organizational design frameworks in a single setting, would provide increased insight regarding the interplay between the frameworks. However, seldom do researchers get such opportunities; our disciplinary, institutional and funding structures today do not encourage such efforts. Yet as we move towards greater understanding of both technical and social aspects of collaborative virtual environments perhaps such new opportunities may emerge. In the meantime, the technical design framework can help guide the development of CVE technology, in particular its ability to support the creation and maintenance of situation awareness across distances. The organizational design framework can help guide the design of research organizations that are geographically distributed. Both frameworks offer new ways of facilitating distributed scientific collaboration.

Acknowledgements

This chapter is based on research funded by the NIH National Center for Research Resources, NCRR 5-P41-RR02170, and the NIH National Institute of Biomedical Imaging and Bioengineering, P41-EB-002025, and by the STC Program of the National Science Foundation under Agreement No. CHE-9876674. Parts of this chapter were originally published in [4, 9].

Thanks to all study participants; to the team that built the nanoManipulator collaboratory system, including Frederick P. Brooks, Jr., Martin Guthold, Aron Helser, Tom Hudson, Kevin Jeffay, David Marshburn, Don Smith, Richard Superfine, and Russell M. Taylor II; thanks to Kelly Maglaughlin, Ron Bergquist, Bin Li, Atsuko Negishi, Leila Plummer and Eileen Kupsas-Soo who assisted in the nM collaboratory project. I would also like to give special thanks to Mary Whitton for her valuable work during the nM controlled experiment and her assistance in an earlier publication on this work, to Joe DeSimone for his comments on an earlier version of the section on conceptual organizations, and to Ralph Schroeder for his helpful comments on this chapter.

References

1. Vidulich, M., Dominquez, C., Vogel, E., & McMillan, G. (1994). Situation awareness: Papers and annotated bibliography. AL/CF-TR-1994-0085, Air Force Material Command, Wright-Patterson Air Force Base. OH: Armstrong Laboratory.
2. Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M.R. Endsley & D.J. Garland (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum, pp. 3–32.
3. Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*: 225–240.
4. Sonnenwald, D.H., Maglaughlin, K.L., & Whitton, M.C. (2004). Designing to support situation awareness across distances: An example from a scientific collaboratory. *Information Processing and Management, 40*(6): 989–1011.
5. Rogers, E. (1995). *Diffusions of Innovations*. New York: Free Press.
6. Orlikowski, W. (1993). Learning from notes: Organizational issues in groupware implementation. *The Information Society, 9*(3): 237–250.
7. Glaser, B. (1978). *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory.* Mill Valley, CA: Sociology Press.
8. Straus, A. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* Thousand Oaks, CA: Sage.
9. Sonnenwald, D.H. (2003). The conceptual organization: An emergent collaborative R&D organizational form. *Science Public Policy, 30*(4): 261–272.
10. Cool, C. (2001). The concept of situation in information science. In W.E. Williams (Ed.), *Annual Review of Information Science and Technology*, Vol. 35. Medford, NJ: Information Today, pp. 5–42.
11. Klein, G. (2000). Analysis of situation awareness from critical incident reports. In M.R. Endsley & D.J. Garland (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum, pp. 51–72.
12. Bales, R. (1950). *Interaction Protocol Analysis.* Cambridge, MA: Addison-Wesley.
13. Nardi, B., Whittaker, S., & Bradner, E. (2000). Interaction and outeraction: Instant messaging in action. In *ACM 2000 Conference on Computer Supported Cooperative Work*. NY: ACM Press, pp. 79–88.
14. Held, R. & Durlach, H. (1992). Telepresence. *Presence: Teleoperators and Virtual Environments, 1*(1): 109–112.
15. Sheridan, T. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments, 1*(1): 120–125.
16. Witmer, B.G. & Singer, M.J. (1992). *Measuring Presence in Virtual Environments* (DTIC Reference Number AD A286 183 DTIC TR 1014), Ft. Belvoir, VA: DTIC, Defense Technical Information Center.
17. Basdogan, C., Ho, C., Srinivasan, M., & Slater, M. (2000). An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer–Human Interaction, 7*(4): 443–460.

18. Usoh, M., Arthur, K., Whitton, M., Bastos, R., Steed, A., Slater, M., & Brooks, F. (1999). Walking > walking-in-place > flying in virtual environments. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*. NY: ACM Press, pp. 359–364.
19. Finch, M., Chi, V., Taylor II, R.M., Falvo, M., Washburn, S., & Superfine, R. (1995). Surface modification tools in a virtual environment interface to a scanning probe microscope. In *Proceedings of the ACM Symposium on Interactive 3D Graphics—Special Issue of Computer Graphics*. NY: ACM Press, pp. 13–18.
20. Taylor II, R.M. & Superfine, R. (1999). Advanced interfaces to scanning probe microscopes. In H.S. Nalwa (Ed.), *Handbook of Nanostructured Materials, Vol. 10*. Norwood, NJ: Ablex, pp. 217–255.
21. Daly Jones, O., Monk, A., & Watts, L. (1998). Some advantages of video conferencing over high-quality audio conferencing: Fluency and awareness of attentional focus. *International Journal of Human Computer Studies*, *49*: 21–58.
22. Zahorik, P. & Jenison, R.L. (1998). Presence as being-in-the-world. *Presence: Teleoperators and Virtual Environments*, *7*: 78–89.
23. Bellotti, V. & Dourish, P. (1997). Rant and RAVE: Experimental and experiential accounts of a media space. In K. Finn, A. Sellen, & S. Wilbur (Eds.), *Video-Mediated Communication*. Mahwah, NJ: Lawrence Erlbaum, pp. 245–272.
24. Harrison, S., Bly, S., & Anderson, S. (1997). The media space. In K. Finn, A. Sellen, & S. Wilbur (Eds.), *Video-Mediated Communication*. Mahwah, NJ: Lawrence Erlbaum, pp. 273–300.
25. Shneiderman, B. (1998). *Designing the User Interface*. Reading, MA: Addison Wesley.
26. Olson, G.M. & Olson, J.S. (2000). Distance matters. *Human–Computer Interaction, 15*(2–3): 139–178.
27. Berg, B.L. (1989). *Qualitative Research Methods for the Social Sciences*. Boston: Allyn and Bacon.
28. Sonnenwald, D.H., Whitton, M.C., & Maglaughlin, K.L. (2003). Evaluating a scientific collaboratory: Results of a controlled experiment. *ACM Transactions on Computer–Human Interaction, 2*(10): 150–176.
29. Sonnenwald, D.H., Maglaughlin, K.L., & Whitton, M.C. (2001). Using innovation diffusion theory to guide collaboration technology evaluation: Work in progress. In *IEEE 10th International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprise (WET ICE)*. NY: IEEE Press, pp. 114–119.
30. Sonnenwald, D.H. (2003). Managing cognitive and affective trust in the conceptual R&D organization. In M. Iivonen and M. Huotari (Eds.), *Trust in Knowledge Management and Systems in Organizations*. Hershey, PA: Idea Publishing, pp. 82–106.
31. Robson, C. (2002). *Real World Research*. Oxford, UK: Blackwell Publishers.
32. Wasserman, S. & Faust, K. (1994). *Social Network Analysis*. NY: Cambridge University Press.
33. Crane, D. (1972). *Invisible Colleges: The Diffusion of Knowledge in Scientific Communities.* Chicago: University of Chicago Press.
34. Wulf, W.A. (1989). The National Collaboratory: A White Paper. Appendix A in *Toward a National Collaboratory*, unpublished report of a National Science Foundation invitational workshop held at Rockefeller University.
35. Duarte, D. & Snyder, N. (1999). *Mastering Virtual Teams.* San Francisco: Jossey-Bass Publishers.
36. Boulding, K. (1989). *The Three Faces of Power*. Newbury Park, CA: Sage.
37. Kuhn, T. (1970). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

38. Grey, D., Lindblad, M., & Rudolph, J. (2001). Industry-university research centers: A multivariate analysis of member retention. *The Journal of Technology Transfer, 26*: 247–254.
39. Sonnenwald, D.H., Solomon, P., Hara, N., Bolliger, R., & Cox, T. (2002). Collaboration in the large: Using video conferencing to facilitate large group interaction. In A. Gunasekaran and O. Khalil (Eds.), *Knowledge and Information Technology in 21st Century Organizations: Human and Social Perspectives*. Hershey, PA: Idea Publishing, pp. 155–136.
40. Salter, L. & Hearn, A. (1996). *Outside the Lines*. Montreal, Canada: McGill-Queen's University.

Chapter 5

ANALYZING FRAGMENTS OF COLLABORATION IN DISTRIBUTED IMMERSIVE VIRTUAL ENVIRONMENTS

Ilona Heldal, Lars Bråthe, Anthony Steed and Ralph Schroeder

1. Introduction

Working together at a distance has been a long-standing research goal of collaborative virtual environments (CVEs). While desktop-based CVEs have come into widespread use for online games and to some extent for distributed work, immersive projection technology (IPT) systems [1] are still relatively rare and they are typically used in specialized applications such as oil exploration, molecular visualization, and architectural walkthroughs. Using IPT systems in networked mode, so that people can share the environment with life-size avatars representing other people in another location, is even more rare. However, it can be envisaged that this kind of distributed collaboration, with highly immersive and surrounding displays, will become more widespread in the future. This prognosis is partly based on technological developments, which are bringing ever-larger screens, more intuitive interface devices and more powerful network connections. It is also based on the increasing need for co-visualization of complex data and large-scale models. Since it is clear that networked immersive systems have distinctive benefits [2, 3], it is useful to make a start on identifying the advantages and disadvantages of these types of virtual environments for distributed collaboration.

One feature of new and highly specialized technologies is that, unlike widespread commercial products, they have not undergone extensive usability testing. Although existing usability evaluation methods have been applied to single-user virtual environments [4–6], these will not be adequate for multi-user or collaborative environments [7, 8]. It is not just the complexity and novelty of the networked immersive technologies that makes evaluations difficult, but also the fact that social processes within the group add a further layer

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 97–130.

of complexity. It has been shown that both the experience and the effectiveness of collaboration in CVEs are substantially affected by groups' characteristics and the interpersonal dynamics [9, 10].

The aim of this chapter is to take some first steps towards demonstrating a method for analyzing social and technical interactions in immersive CVEs. The method consists of analyzing sequences of interactions which we call "interaction fragments". Interaction fragments usually last only a few seconds and often recur in a similar form several times per session. They either support or disturb the process of collaboration for solving the task. Fragments can be classified as mainly involving interaction via the interface technology, interaction with each other (social interaction), or fragments involving problem solving (or reaching the goals of the tasks). One question that then arises is whether generalizations can be derived from these very brief snippets of interaction. In the conclusion, we will return to the question regarding the lessons that can be learned from these fragments, as well as their broader usefulness.

2. Background

There have been several attempts to classify the experience in virtual environments in order to understand what makes their use more effective and enjoyable [11–13]. Moreover, a number of factors have been identified such as presence, immersion, interaction, etc., that influence the experience of—and performance in—virtual environments [14, 15]. Additional factors that need to be taken into account include the levels of realism of the representation of self (the embodiment), of the objects, and the surrounding space. It has been argued that the aim of virtual reality systems should not necessarily be to reproduce physical artifacts or achieve graphical realism [16], nor to reproduce natural interaction in these environments for all situations [17]. For example, although a more realistic embodiment may support better collaboration, simple embodiments can often be sufficient for interaction [18, 19]. Here, we will leave realism and embodiment mostly to one side, and concentrate instead on how people collaborate with each other.

Presence can influence interaction in different ways. For example, several studies show that presence is correlated in different ways with task performance [20, 21]. Task performance and usability issues also depend on the applications [9] or settings [22, 23]. Nevertheless, we will argue that it is possible to identify some common usability issues if we break interaction down into social interaction, interaction via technology, and interaction in order to reach the goal [24]. Apart from the work of Heldal, the analysis here can draw on several kinds of previous studies: studies of single-user environments [17] and of orientation and navigation [6, 25]. For social interaction, some of the main mechanisms or elements (conversation, coordination and awareness) have been identified by Preece [26] and by Tromp [7] for collaboration and usability, and by Tromp, Steed, & Wilson [10] for social conduct (verbal/phatic communication, turn-taking, etc.). To examine how participants reach their goals, it is also necessary to take into account the strategies of task-focused collaboration, proxemic shifts, as well as observations of changes over the course of time.

This chapter is based on a trial in which the aim was to examine collaboration and interaction in networked IPTs. A number of papers discuss other results related to this trial: one discusses the differences in interaction between strangers and friends, i.e. people who do or do not know each other [27]. Another examines some usability issues and the importance of the awareness of one's partner's intentions during collaboration based on questionnaires and interviews with the participants [9]. A third study examines the successes and failures in collaboration in the immersive networked IPT setting with those that have previously been identified for desktop-based CVE systems for object-focused interaction [28]. In this chapter, we will concentrate on the usefulness of examining fragments of interaction for understanding people's collaboration. (Note that interaction fragments should not be confused with the "fragmented interactions" which Hindmarsh, Fraser, Heath, & Benford [29] identified when they analyzed interruptions in the flow of collaboration in desktop CVEs. In their work, Hindmarsh *et al.* showed that problems arose when the technology got in the way and disturbed the users, who were then unable to collaborate successfully.)

3. Study Design

We examined six pairs of users working together via two IPT systems in different locations. Five pairs spent at least 210 minutes each doing five tasks together in networked immersive virtual environments over the course of a day. For one pair, the trial was stopped approximately halfway through because both partners experienced severe nausea and anxiety. The subjects took a break of between 15 and 20 minutes between tasks, and had a longer lunch break of 60–90 minutes between the first two and the other three tasks. The times that pairs spent for each task session were between a minimum of 25 minutes and a maximum of 70 minutes. The order of the tasks was the same for each group, so that they were exposed to the same experience, and could thus be compared:

1. Puzzle—the task was to do a small-scale version of the popular Rubik's cube puzzle, with eight blocks having different colors on each side so that each side would have a single color (i.e. four squares of the same color on each of the six sides, see figure 5-1).
2. Landscape—the environment in this case was a small townscape with surrounding countryside ringed by mountains (see figure 5-2). Subjects

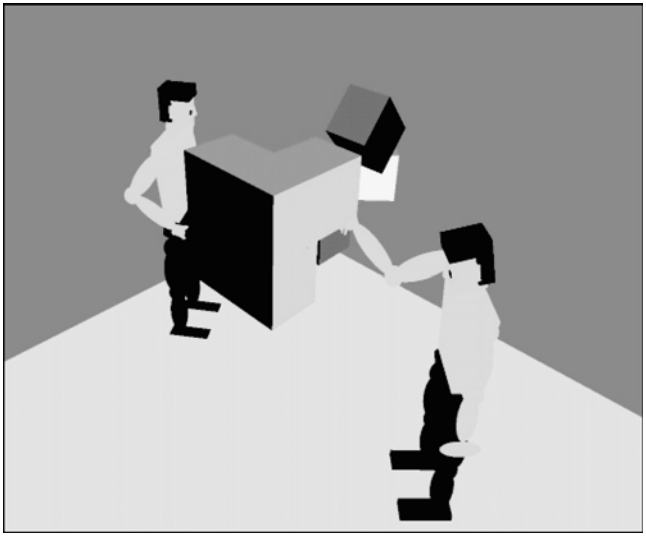


Figure 5-1. Puzzle.

were instructed to familiarize themselves with this landscape and count the number of buildings. They were also told that they would be asked to draw a map of the environment at the end of the task.

3. Whodo—the task was based on a popular game, in this case the murder mystery board game Cluedo. The subjects were asked to find five murder weapons and five suspects in a building with nine rooms (see figure 5-3).

Figure 5-2. Landscape.

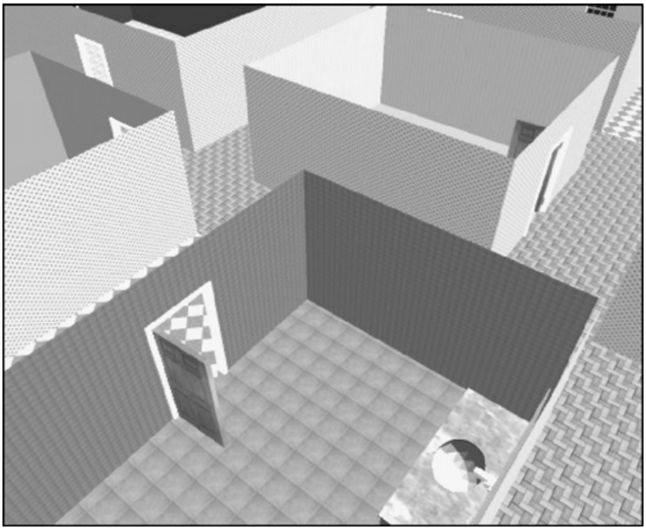


Figure 5-3. Whodo.

They needed to locate the murder victim's body and to find and eliminate weapons and suspects.

4. Poster—this environment consisted of a room with ten posters stuck on the walls (see figure 5-4). The posters each contained a list of six sentence fragments. When all the fragments were put in the right order, they would make a popular saying or proverb.

Figure 5-4. Poster.

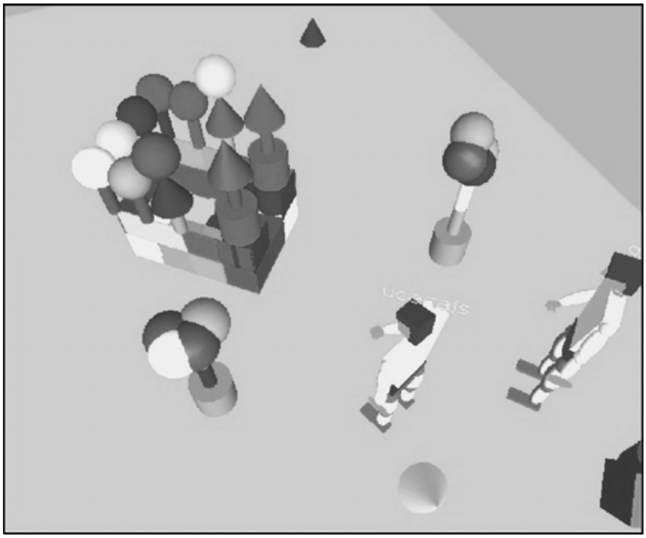


Figure 5-5. Modeling.

5. Modeling—this environment contained 96 shapes (square blocks, cones, etc.) in six different colors. The subjects were told to make a building, or model of a building, to be entered in an architectural competition. They had to use at least three colors and the building had to be a single object. The result was to be their joint "architectural masterpiece" (see figure 5-5).

One of the goals of this study was to cover a wide range of tasks (for a more detailed description, see [27]).

The IPT system at Chalmers University in Gothenburg was a five-sided (no ceiling) 3 m × 3 m × 3 m TAN VR-CUBE. The application was run on a Silicon Graphics Onyx2 Infinity Reality with fourteen 250 MHz R10000 MIPS processors, 2 GB RAM and three graphics pipes. The participants wore CrystalEyes shutter glasses. A Polhemus magnetic tracking device tracked both the glasses and the wand. The rendering performance was at least 30 Hz in the applications discussed here, except for the Landscape world where it was around 10 Hz. The IPT system at University College London was a four-sided Trimension ReaCTor with a floor of 2.8 m × 2.8 m and three 2.8 m × 2.2 m walls. It was powered by a Silicon Graphics Onyx2 with eight 300 MHz R12000 MIPS processors, 8 GB RAM and four Infinite Reality2 graphics pipes. The participants wore CrystalEyes stereo glasses. The head and wand were tracked by an Intersense IS900 system. Rendering performance was at least 45 Hz in the applications discussed here, except for the Landscape world where it was around 10 Hz. Both applications were implemented in a customized version of the Distributed Interactive Virtual Environment (DIVE) system [30, 31].

Each participant was portrayed to the other by the use of a simple avatar with a jointed left or right arm. The participant could not see his or her own avatar, except for a virtual hand drawn in the same position as the physical hand. Although local tracker updates perform at the fastest rate provided by the tracker driver, updates to the remote avatar are only sent at 10 Hz to avoid congestion. The network latency between the two sites was approximately 180 ms. Locomotion was effected by a "move in the direction of gaze" metaphor. The object manipulation metaphor was a ray-casting technique but using a short ray of about 10 cm. The subjects could talk to and hear each other by using a wired headset with microphone as well as earphones. The Robust Audio Toolkit (RAT) was used for audio communication between the participants.

4. Method

For this study, we examined video and audio recordings and referred to some data collected via questionnaires. It is difficult to examine the exact occurrences of certain activities or complete actions in an IPT system because it is not possible to record the activities from all angles [6]. Thus, we recorded the activities from one angle from a camera on top of one corner of the IPT system in Gothenburg and from behind the participant in the IPT system in London. In this way we could watch and listen to details of problematical or supporting sequences over and over again.

We use the following notations in what follows:

LxGx	The *x*th couple ($x = 1, \ldots, 6$), where one partner Lx is working in the IPT system in London and Gx in the IPT system in Gothenburg.
LxGx Task yy:zz	The *x*th couple working on the task indicated has worked for *yy* minutes and *zz* seconds on the task. For example, *L2G2 Puzzle 10:30* means that L2G2 has worked for 10 minutes and 30 seconds (approximately) with the Rubik-cube type puzzle.
Lx:blabla	Lx saying "blabla"
Lx ⇒Wait	Lx interrupts Gx by saying "Wait" or says very quickly "Wait"

Figure 5-6. A subject locomotes straight ahead in the virtual environment. The subject's face and glasses are in the direction of the small line on the circle.

Figure 5-7. A person locomotes backwards.

For the Poster task only we use UPPER-case letters for the sayings on the posters. The lower-case letters are the participants' own words (the notations for how subjects move are shown in figures 5-6 and 5-7).

It is worth mentioning that it is essential to view interaction fragments repeatedly in order to be able to break the interaction down into its significant and detailed parts. For reasons of space we will often present only the transcribed text of particular fragments, including enough description of the actions to provide the context—but ideally the video and audio recordings, and transcriptions of text and movements should be analyzed and presented in conjunction to arrive at a complete understanding of how the collaboration works (some examples are available online at www.mot.chalmers.se/tso/ilona/fragments.html).

5. Interaction Via Technology

5.1. Orientation and Locomotion

The following observations refer to orientation and locomotion fragments, and fragments identified when people responded to and handled the virtual and the real objects around themselves.

For the Puzzle and the Poster tasks, people hardly had to orient themselves or get their bearings since they had a view of the whole virtual space. The orientation in the Landscape, on the other hand, was considered very difficult by all couples. One reason was that the environment contained many similar buildings with a generic background. It was hard to differentiate reference points, which resulted in the pairs using trial-and-error strategies and taking a long time to familiarize themselves with the environment. The conversations show that it took at least 20 minutes for participants to get an approximate feeling of the size of the environment, and to establish facts such that the main road was circular:

L6G6 Landscape 12:30
L6: It feels that we are going in circles.
G6: Yes. The problem is that the horizon everywhere looks the same . . . All the time . . . .

Many pairs said that locomotion in the next task (Whodo) was easier. The corridors and rooms with different appearances helped locomotion and helped

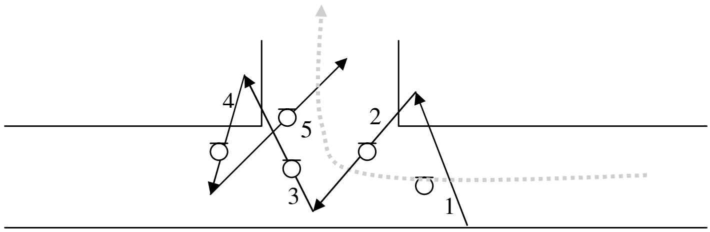


Figure 5-8. L1 tries to follow the gray trajectory, but his movements follow the black arrows. He goes into the wall (1), goes backward (2), tries again and goes into the wall (3), goes backward (4), tries again straight ahead (5) and at this point he is in the corridor going upward (L1G1 Whodo 10:30).

to identify reference points. Some of the people did not follow the corridors and just went through the walls. Those who tried to follow the corridors were constantly colliding with the walls. Figure 5-8 shows the trajectory of one person who tried to go right in a T-junction.

In the Modeling world it was easy to orient oneself even though the space was quite large. This was because the differently colored and shaped objects provided reference points.

For all the tasks we observed that people had difficulties in locomoting straight ahead, and this persisted over long periods of time. They had problems following each other and following the roads (for the Landscape) and the corridors (for the Whodo) tasks. These two tasks were the most locomotion-intensive. Since the task in the Landscape was collaborative exploration, they had to follow each other a lot, but they had a hard time doing this at a constant distance. They often "ran through" each other during a straightforward locomotion or followed the other subject's avatar in a zigzag manner. They commented on this mainly in the beginning of the task, as in the example below, or when they were not preoccupied with the task:

L1G1 Landscape 8:00.
L1: It is fun. It's really like racing.
G1: Yes, it would be funny to [be able to] run so fast!
L1G1 8:40.
L1: Sorry, I'm running through you again.

Just a short while later, between 9:30 and 10:30, L1 ran through G1's avatar four times again, but without mentioning it at all. This is true for other couples too; L6G6 mentions only once that they crossed each other in the beginning of the task, as when L6 apologizes:

L6G6 Landscape 3:00
L6: I think it is hard to go straight ahead . . . Oops! . . . Sorry! [she crosses G6]. I don't think I can go straight ahead.

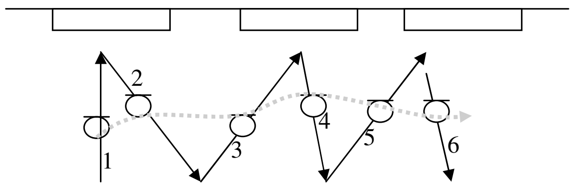


Figure 5-9. Instead of following the gray arrow and locomoting sideways, the subjects usually locomoted back and forth in zigzag in front of the walls with the posters.

In this task the participants could locomote quite fast, which also contributed to the difficulty of adjusting the movements to each other. Among all couples only one person mentioned that he liked it—"it is like racing" (see earlier quote). Several others said that it would be good if they could take a bus or a car to travel with. In the following example, G3 several times asks his partner to slow down.

L3G3 Landscape 6:40
G3 has a hard time following L3.
G3: Can you slow down? . . . So I can adjust . . . .

L3G3 Landscape 9:25
L3 has problems locomoting, but he and G3 ignore this. G3 is only bothered that he cannot follow L3.
G3: Can you slow down? . . . I cannot see you any longer.

Figure 5-9 points to another frequently observed behavior in locomotion. For all tasks, people frequently and easily locomoted backwards. This could be observed in open spaces (Landscape and Modeling), closed spaces (Whodo), as well as smaller places (Puzzle and Poster). In the Poster world, instead of turning somewhat and locomoting sideways, almost all subjects locomoted in a zigzag manner.

In the Modeling world, the subjects often locomoted from the middle of the virtual space, where they built their building, to the border of the environment where the required objects were situated. They did this only by using their joystick; physically they did not move in the IPT system. In other words, they locomoted straight ahead, fetched the object and backed up, rather than turning around physically in the space.

5.2. Using the Interface Technologies

We examine here both the treatment of the virtual objects and the handling of real objects (technical devices and constraints of the surrounding space).

In the subjects' interactions with the technology in this regard, we observed disturbances on three different levels. The first were small interruptions such as usability problems (for example when the subject could not grab an object on the first attempt but could do so on the second), short audio disturbances, inconsistencies in texture updates, and the like. These problems with the system were rarely mentioned at all during the collaboration. An example here is when menu windows were accidentally displayed in the environment by the experimenters (for couple L5G5 Puzzle, 29:39). The subjects did not react at all, probably because the windows disappeared quickly.

A second type of technical disturbances comprised those that were observed and mentioned by one of the subjects but where the partner did not react. For example, the projected hand of one subject got stuck in the bedroom door and the subject complained—even though her partner did not observe the problem at all. In fact this happened because the displayed images become frozen for some seconds when the London tracker lost synchronization for a while:

L6G6 Whodo 4:00
G6 speaks about the task.
L6: My hand got stuck in the door . . . [Some seconds] Now it lets me go.
G6 continues to talk about the task as if nothing had happened.

Another example of this type is when L6 (L6G6 Poster 17:33) mentioned audio problems, and he asked his partner if she heard the "strange noise", but the partner did not answer and no further attention was paid to it.

The third type of technical disturbance affected both partners. For example, if a subject complained for a long time to her or his partner, asked them for help, or informed them explicitly. Here is an example of the latter:

L1G1 Poster 30:00
G5 has problems with his glasses; the operator changes the batteries in his glasses. L5 is solving the problem; he almost does not notice that G5 has problems.
G5: OK. Now it is a little bit clearer. But still blurred a bit. I have to adapt to it more.
L5 [observes that something is wrong with G5]: OK.
G5: This is technology.
L5: Are your glasses a little bit heavy as well? [He speaks about *his own* problem.]
G5: Unfortunately they cut on one of my ears . . . But the problem is that they are blurry.
L5 ⇒ Move on! You are in my way! [He is eager to continue with the task.]
G5: OK.
L5: TO RIGHT stays here.

Sometimes longer breaks occurred in the audio connection. In these cases the subjects usually worked alone and tried not to care too much about the interruption:

L5G5 Poster 5:00–8:45

The couple did not observe that the audio had come back, only when G5 started whistling and L5 heard it did he say:

L5: Hi! Are you back?

G5: Yes, I'm back. But I do not have "buttons" [joystick]. Now I have. Wait, I have to put on the clothes [the glasses].

Even though both subjects were using IPT systems, the images displaying the virtual environment were not completely symmetrical since the two IPT systems were not entirely the same. For the tasks where the subjects had to differentiate colors in order to solve the task properly (Puzzle, Posters), the subjects in London found it easier. This was probably due to the better projection system. For those tasks where the subjects had to manipulate more (Puzzle, Whodo, Modeling), the subjects in the IPT in London again found this easier. In this case it was because of the better tracking system. For some tasks, a more surrounding experience was more valuable, particularly in the Landscape, Whodo, Poster, and Modeling tasks, which were better supported in the IPT in Gothenburg with its five walls. The walls were made of fabric in the IPT in Gothenburg and of acrylic in London. Even this slight difference caused differences in the users' reactions: since the users were supposed to handle the surrounding virtual space separately from the real space, the constraints of the physical surroundings disturbed them, e.g. when one subject collided with the real walls. The subjects in London mentioned these collisions much more often, especially in the beginning of the trials:

L1G1 Posters 11:08

L1: Oh s!&% . . . ! I'm just going straight into the walls of the CAVE again, I think.

G1 [mumbling]: Be careful.

One of the reasons why they mentioned it more often may be that they could hurt themselves more than people in Gothenburg, who collided only with a fabric wall.

For L2G2, it was the person in Gothenburg who often collided with the walls of the IPT system. In the beginning he mentioned this but his partner never answered. We observed that it was generally in the beginning of the tasks that the subjects paid more attention to disturbances caused by colliding with the walls, getting in a muddle with the cables, or complaining about using the 3D glasses. This was not so later on.

Figure 5-10. Working for 20 seconds with the hands outstretched, but not touching the walls to avoid possible collision.

As figure 5-10 shows, one can focus on solving the task, work on fetching objects, and at the same time reach out the other hand to avoid potentially colliding with the wall of the IPT system for a longer period.

During the fourth task, the Posters, only one subject (G5) mentioned the furniture in the room where the Posters were on the walls. This person said that it was a pity that the chair in the middle was not real, since he was tired and he would like to sit down:

L5G5 Poster 41:00
G5: I would like to sit down. We have a chair here. I would like to sit on it. But of course I cannot. Irritating . . . .
L5: Try out! So . . . .
G5: Is this a TV in front of me?
G5: I think that these are mainly computers. But maybe this one is a TV. I'll take it home. Should we jump on the 9 [9th list of words in the Puzzle] now? This is irritating.
L5: [in a higher voice] But this is nearly done! WHERE EVERYTHING IS . . . No, this EVERYTHING cannot be.

The following example shows that the users ignored those objects that were not important for task solving. They would just locomote through them and hardly mentioned them at all if they were not relevant for solving the problems.

Figure 5-11. Ignoring objects (L5G5 Posters 9.00).

In the example above, the couple had already finished three tasks and had worked for approximately 40 minutes on the fourth, indicating that they could have been physically tired and bored in the environment. Otherwise, every user completely ignored the furniture. They just went through the objects or stayed for substantial periods inside them, as the figure above (figure 5-11) shows where someone just stands in the middle of a shelf.

The virtual walls were more often mentioned for the Poster task, both because it was difficult to locomote to and from the posters (see figure 5-9 above) and because the subjects unintentionally and by mistake went through the walls and left the room (too much locomotion straight ahead). Usually, they did not discuss this, but sometimes a subject looked for her or his partner who could be hidden by the walls:

L1G1 Posters 11:22
L1: Are you outside as well?
G1: No, I'm inside.

Even though the objects are important for solving the problems, the couples merely adapt to these virtual objects' non-material structure and pass through them easily. This is true for all tasks. However, they were often aware of the objects. This is shown best in the Whodo application, where the couples were

Figure 5-12. L1G1 Whodo 17:56–18:10. L1 goes to the table (picture 1) through the table (picture 2), turns back, bends down, looks under the chair, and goes down on the knees under the table looking upward (picture 3).

looking for a hidden murder weapon. They went through a bed, sofa, chair, and table in one second, in such a way that the observer had the feeling that they did not consider that there were any objects in the environment. Yet, seconds later one participant just looked under or crawled completely under a virtual object, to see if there was a weapon hidden under or behind it (see figure 5-12).

Despite the fact that the objects had neither mass nor gravity, the subjects handled them almost like real objects. It was possible, as the following example shows, that one person could mark an object, move it to her or his partner, and leave it (the object would stay suspended in the space), while the other "took it" and moved it where she or he wished. This is clearly hard to perform in non-immersive CVEs.

L6G6 Puzzle 16:30
L6: What is happening if I try to take this from you?
G6: I don't know. Oops [surprised as she hands over a cube and L6 takes it] . . . do you have a red one at your place?

A problem that affected collaboration was that many subjects, for all tasks, kept pointing with non-tracked hands. In many cases they failed to observe that their partners could not possibly see what they were pointing at. They did this on several occasions for short or long periods.

Except for the Puzzle task when the cubes snapped together if they were close enough to each other, the objects were quite difficult to adjust into a correct position relative to each other.

L6G6 Modeling 17:30–50
L6 tries for several seconds to fix a cylinder straight in the right place.
L6: How stubborn it is! OK, then. This can be inclined, then [leaves the cylinder at an angle rather than straightened].

5.3. Handling the Partner's Avatar

There was a difference in how subjects handled the partner's avatar during the different stages of collaboration (introductory phase, proper collaboration, end phase). During the introductory phase they often briefly discussed each other's appearance, the colors of their clothes, the sizes of their avatars, the name tags on the top of the other's head and the like. They were also interested in their own appearance, and whether, for example, they looked as "slim" as their partners. They often combined this short information with "real information"—where the other was situated physically, how the weather was in that city, or about their occupations.

During the proper collaboration, when each subject was focused on solving the task, the subjects did not care about their avatars except as points of reference. If they stopped solving the task and talked with each other about general matters, and about the task, they handled the others' avatars almost like real people. For example they would face each other when speaking to each other (figure 5-13).

In the intermediate phase, during proper problem solving, they would locomote through each other more often. There were instances for most of the tasks when one went through the other's avatar several times in the course of one minute, or also worked standing inside the other person's avatar for several

Figure 5-13. L5G5 speaking to each other.

seconds. In these cases they mentioned the other's avatar only if it disturbed problem solving, for example by blocking each other's view. In the following example, L5 tells G5 to move out of his way, and immediately afterwards continues with the puzzle:

L5G5 Poster 32:00
L5 ⇒ Move on! You are in my way.
G5: OK.
L5: TO RIGHT stays here.

A key benefit of using immersive VR technology is that a person can quickly glance at what the partner is doing. These are very small movements that may not have been noticed by the participants. However, they occurred frequently, especially in the large spaces (Landscape, Modeling) where people worked for longer time periods on their own by observing, counting and carrying objects (but also with the Poster task).

Finally, many of the pairs who succeeded with a task "shook" each other's hand (tried to move their virtual hand to the partner's hand), and L6G6 waved goodbye to each other at the end of almost every task, and when they succeeded with the Puzzle, they lifted their hands over their head and tried to "high five" each other.

6. Social Interaction

Here we present fragments that illustrate being-at-ease or having difficulties in the course of social interaction. These are sequences of actions that show how a person in a pair handles her or his partner during conversation and communication, and to what extent one subject is aware of the other.

6.1. Conversation

Small appreciative phrases, e.g. "Great", "Precise", "Cool", "Fine", "Perfect", etc. were effective in supporting collaboration. When one subject in the couple encouraged the other, the collaboration went fine, even though the partners seldom reacted to these phrases. They did not necessarily react to short acknowledging sentences, like "this is nice/hard/right", "I'm here", etc. or questions like "How are you?", "All well?". Nevertheless, these played an important role in keeping the conversation going. By the same token, people who often complained about how difficult the task was, or that they did not think they could manage it, or who were silent—disturbed the collaboration, even though their partner did not necessarily acknowledge these complaints directly:

L3G3 Puzzle 16:20
G3: I still haven't learned how to operate. How to get them. . . .
L3 [silence].
G3: Aha. What happened with that one?
L3 [silence].
G3: Do you think we'll make this in 20 minutes?
L3: Hah [strange laugh]. No. I don't think so!

How quickly a subject answered the partner was very important and directly influenced the success of collaboration. Longer periods of silence confused the partners because they did not know whether the partner was listening at all, intentionally or unintentionally ignoring their comments, or if they were experiencing a technical failure. For L2G2, the subject from Gothenburg constantly made more efforts that kept the conversation going, and the collaboration went much better than for L3G3:

L2G2 Puzzle 6:00
G2: We are going to run out of the cubes soon [says this in a joking way].
L2: I think we may each have a side of different color.
G2: Yes, I think so.
L2 [silence].
G2: OK, black is no color. Maybe you tried to tell me that?
L2 [no answer].
G2: "Ah."

The conversation was problematical for L2G2, but G2 pushed it forward. L2 was passive, but at least he responded with "OK", "Sure", "Right". As a result, they were one of only two pairs who succeeded in putting together all the 11 sayings on the posters (correctly or not). For example, G2, the Swede, often asked his English partner if he had trouble understanding:

L2G2 Poster 34:00
G2: EVERYBODY. MEMOIRS. What does that mean? Remember?
L2: MEMOIRS? . . . Is that . . . like a book of personal experiences?
G2: Ah! Memoirs!

For L3G3, even though G3 had problems with the language, he never asked his partner for help. The following conversation sequences show how L3 gets frustrated because of his passive partner, who often took a long time to respond—and when he did so, it was in an unhelpful way:

L3G3 Whodo 35:00
L3: OK. Ah . . . Last time, when you clicked the rooms, what criteria did you follow to click on the rooms?

G3: Ah, Ah . . . .
L3: You see, you have done it well. So? . . . I don't know . . . what did you do with the rooms?
G3: Last time? Do you mean last game?
L3: Yapp.
G3: [does not respond].
L3: Yes?
G3: Do you mean . . . ? [short pause].
L3: Yes. [upset].
G3: . . . what I did with the rooms last game? [pause].
L3: Yees! [upset].
G3: Jaha. Here I can see you. I'm right behind you.
L3: OK . . . [he gives up].
G3: I can see the dead body again. [pause] I just clicked on the posters.

G3 also had problems with using the devices and seeing objects in 3D, which made him talk in Swedish (to himself, or to the operator), which, in turn, made L3 upset:

L3G3 Whodo 25:00
L3: Hello!
G3: Yes.
L3: So, we have to do it again!
G3 [talks in Swedish to the operator].
L3 [irritated]: He is talking Swedish!

Conversation played different roles during the different phases of problem solving. During the introductory phases for each task, it was especially important to suggest strategies, ask the partner about suggestions, acknowledge the partner's suggestions quickly, and show that one was active and interested in the task. L6G6, a pair who solved the tasks effectively, had intense conversation periods at the beginning of each task, as the Landscape example shows:

L6G6 Landscape 3:30
G6: I think is quite hard to see where a house begins and ends. Should we go nearer to see better?
L6: OK.
G6: Should we go near?
L6: Should we traverse the lawn?
G6 ⇒Yes, I think we should go straight over.

Frequent acknowledgment and turn-taking were important. However, quick comments sometimes resulted in confusion. By not having gestures, facial expressions or eye glances, the conversation was occasionally stopped by unnecessary breaks and starts when partners did not know who should continue speaking:

L6G6 Landscape 30:00
L6 ⇒ This has to be . . . [she stops and waits].
G6 ⇒ Is this . . . [she also stops and waits].
L6: An airfield.
G6: Or something. . . .

Speaking simultaneously was not just distracting but also, on some occasions, supported collaboration. For the Posters, people often read the words at the same time by dividing the posters between themselves, and each subject read the words from the posters located in her or his part of the room. Hearing several words belonging in one way or another to the same saying helped them to put together the proverb more quickly. They often recited the sayings, or possible variants, together. This seemed to be an effective method for all couples in this task:

L6G6 Poster 20:00
L6 and G6 together: DIPLOMACY IS TO SAY THE NASTIEST THINGS IN THE NICEST WAY.
G6: No. TO DO, we have here.
L6 and G6 together: DIPLOMACY IS TO DO THE NASTIEST THINGS IN THE NICEST WAY.

The following example shows an adaptation to the "problematical partner" encountered earlier. L3 in a way got used to G3's complaints and to the fact that from time to time he spoke in Swedish. He filtered the unnecessary words out and tried to put the right sayings together:

L3G3 Poster after 20:00
G3: OK. Seems that I had some problems with my glasses, but now it is OK.
L3: Should we do the number one? A CRITIC IS A MAN WHO KNOWS.
G3 ⇒ THE WAY har jag hittat här [He says in Swedish: "har jag hittat här", which means "I found here"].
L3: Yes! A CRITIC IS A MAN WHO KNOWS THE WAY.
[It seems that he does not care about Swedish being spoken.]

6.2. Communication

It is difficult to separate communication from conversation, awareness, or problem solving that supports collaboration. Treating the other's avatar as discussed earlier can be regarded as (non-verbal) communication. Conversely, people could speak in the environment without communicating with each other. For example G3 sometimes spoke Swedish, or spoke with the operator. G1 sometimes spoke with the objects. Some people described certain movements without being interested in whether the partner answered or not. Finally, one person could be aware of the other without necessarily communicating with her or him. For example, when one subject became frustrated that the other was too quiet. Collaborative problem solving is both an internal and external process [32]. One must communicate with one's partner in order to work together, but thinking about new strategies or understanding strategies that have been suggested does not necessarily entail that one should communicate with one's partner. And engaging with the other's avatar is not necessarily communication—as, for example, when one locomotes through the other's avatar several times without mentioning it.

Besides speaking to each other, as presented in the section on technical interaction, the subjects could easily see their partner's location in the environment during most of the tasks and they could note what the partner was doing. By taking a short glimpse of one's partner, one could ascertain that everything was all right. Immersive technologies also make it possible for participants to easily relate to their partners' size, position in the environment, and direction of movement [28].

Gesturing and pointing was important for collaborating effectively on the tasks, so as to help each other if needed. Seeing what the partner was doing and speaking at the same time enabled a subject to assist the partner in handling the technical devices or to do things instead of her or him:

L3G3 Puzzle 15:30
G3: . . . I don't know how to drop this.
L3: You just release the button.

The real-time communication also created possibilities to discuss inconsistencies in the environment. Even though the subjects could not be sure whether their partner's environment and the devices they used were exactly the same, they appeared to assume this. Here G6, who went through the floor of the simulated environment (due a software inconsistency), tried to help her partner by telling her how she managed to come back after having experienced the same problem:

L6G6 Landscape 27:00–28:00
L6: Now I'm coming up again.
G6: I think it is only to go backward. Anyway, when you are falling down through the texture . . . Are you there somewhere? Or are you stuck?
L6: ⇒ How did you came back before? Did you go backward . . . or so . . . ?
G6: I just went backward. I stayed and went backward.

L6 [to herself]: We should go straight backward. I can't. Failed. Ah, now I'm going backward.
G6: Yes, I see that you are . . . .
L6 ⇒ I'm going out to the road instead . . . I refuse to go further in the quicksand.

For the modeling world, the participants often had to help each other because if one person grabbed an object this person "owned" it and the partner could not grab it. The following is a case when a subject asks her partner for help to move an object. In addition, she instructs the partner to put it in the right place and the correct position.

L6G6 Modeling 46:30
L6: Could you please help me with the red one? This is one that I definitely cannot move. Please put it round this one [pointing with the wrong hand].
L6G6 Modeling 47:30
L6 is giving instructions to G6 where she should put an object.
L6: A little more . . . A little more on my side . . . Now straight . . . Straight . . . a little more straight away . . . A little more . . . Yes!

Here, we have to stress that both partners would use their non-tracked hand for pointing throughout all tasks. Sometimes they noticed their mistake and commented on it:

L5G5 Puzzle around 7:00
G5: Here, here . . . against me . . . Do you see my . . . Of course, you can't see my hand. Damn, I'm waving with my real hand, hmmmm. . . .
L5: Sure, yes . . . If you point with the joystick hand I'll see it.

6.3. *Awareness of the Partner*

The partners were aware of each other since they helped each other directly (when a subject asked for help) or indirectly (when they were looking around and monitoring what the partner was doing or what was happening). For the Puzzle and the Modeling applications, because the environments were small and open, each user could constantly see their partner. The only exceptions were for the participants in the IPT system in London who would occasionally face out towards an open wall. This did not cause any major disturbances in collaboration. For the other applications, there were times when one person could not see the other's avatar. In these cases, the subjects often reported to each other verbally on where they were and what they were doing.

Almost all subjects in the experiment helped their partners occasionally during the tasks. For the Puzzle task, the subjects often asked their partners about the colors of the hidden sides; that is, the sides that they could not see but that they knew their partners could see. Even though manipulating the cubes was easy, the subjects could get a sense of the colors of objects more quickly by asking their partners rather than by manipulating the object to see themselves.

Many of the subjects also verbally described what they were doing in the environment. They followed their own movements and thoughts with a "trail of words". Such behavior was especially prominent for the Whodo task, probably because the subjects were in different virtual rooms and could not see each other but where the task required keeping the partner informed of progress. "Now I'm in the room" was a typical phrase in Whodo. For this task, they often jointly discussed the rooms, weapons and suspects that they had to eliminate.

Saying things together perhaps gave a feeling of being acknowledged by the partner. In the Poster task the couples often said the sayings jointly. In the Landscape, they often jointly counted the houses:

L1G1 Landscape 19:00
L1: How many houses do we have now? Fifteen, or?
G1 and L1 [counting out loud together]: We had 11 before.
L1: Yes, we had 11 so it might be 16. . . .

Two things that seemed to be particularly important were being able to do things together and being able to see what the partner was doing. However, this requires implementation awareness from designers. Designers could, for example, implement objects that support collaboration so that they can be used or viewed simultaneously by both partners. However, when people got to know certain benefits of their partner's technology that they did not have themselves, often in the middle of a task, they got frustrated. Almost all subjects in Gothenburg observed relatively late during the task that they could see the posters on the walls just by turning around. When they mentioned this to the partners from London, who usually knew it already, they felt disadvantaged and frustrated:

L5G5 Poster 17:00
G5: I try a smart strategy here. I just missed a thing at the beginning. This room, the Cube I'm inside in, is almost exactly the same size as this room [the virtual]. Also I do not have to move [locomote] around. I can go forwards and backwards instead.
L5: Yes . . . I do so. . . .
G5: Really? You don't use your "remote control" [joystick] at all?
L5: No, I'm moving instead.
G5 [somewhat irritated]: OK. Well . . . Number 5 [he turns to the task instead].

The following example is another illustration of how a person becomes frustrated when he becomes aware of the advantages that the other person has. This is an example of when the partner is able to manipulate objects much more easily:

L3G3 Whodo 15:00
G3: But you? Don't have any problem, do you?
L3: Well, no. . . .
G3: You understand the task?
L3: Yes. And you?
G3: [sighs] I hope so . . . I don't know how to drop this.
L3: You just release the button.
G3: . . . The one that I was working with is now in front of you [irritated].

7. Problem Solving

To examine how people reached their goals in the environments, we can consider the flow of collaboration in different phases of problem solving. In this trial, the subjects had to solve problems in all five applications. For two tasks (Puzzle, Whodo), they got visual responses and knew whether they solved the problem correctly or not. For the Landscape and Poster task, they did not know, even though many of them were curious about the results. The Modeling was an open-ended task. In examining problem solving, we were interested in how they began to solve the problems, what they did to choose strategies, whether they repeated the same strategies or changed them, and how they acted when they got partial results.

There were very short discussions about choosing strategies before the first task. Almost everyone used trial and error, just attempting to put the cubes together by building one side from four small cubes with the same color. All of the couples ended up redoing their first side. For almost all couples (except one) the strategy they then adopted was to continue in the same vein—to begin with one side of the same color, or sometimes a second simultaneously, and adjust the rest to it. Only two couples (L1G1, L6G6) chose a different strategy at a later stage.

In the Landscape, when people looked around, they often observed that the size of the environment was huge. Many of them suggested, as an initial strategy, to divide the environment between themselves so that each should explore half of the city. However, all couples changed this strategy because it was difficult to communicate the results. Also, they could not easily delimit half of the environment so that they would know what was assigned to them.

For the Whodo task, the couple often discussed how they should proceed since different tasks and investigations took place in different rooms. The importance of continually acknowledging correct strategies, while actively contributing to influence the solution and modify it if necessary, is shown in the following example:

L5G5 Poster around 13:00
L5: WHEN . . . THEY . . . NAME . . . SPELL . . . and . . . YOU ARE. What can this mean?
G5: One more time! Take it again. I did not find any 5 yet.
G5: And I have YOUR here.
L5: OK. YOU ARE FAMOUS WHEN THEY CAN SPELL YOUR NAME. It is done!
G5: OK.
L5: YOU ARE FAMOUS WHEN THEY CAN SPELL YOUR NAME.
G5: YOU ARE FAMOUS WHEN THEY CAN SPELL YOUR NAME.
L5: Perfect. We took this in 10 seconds.
G5 ⇒ No! Wait a minute! I have here...Yes, that is true.
L5 ⇒ No. Then we begin. It was the fifth, or? YOU ARE FAMOUS WHEN THEY CAN SPELL YOUR NAME.
G5: No! It says here: IN KARACHI. Says here. Wait! We did not reach this before. YOU ARE FAMOUS WHEN THEY CAN SPELL YOUR NAME IN KARACHI.
L5G5 [they read together]: YOU ARE FAMOUS WHEN THEY CAN SPELL YOUR NAME IN KARACHI.
L5: OK. YOU ARE FAMOUS IN KARACHI WHEN THEY CAN SPELL YOUR NAME.
G5: No, the opposite! YOU ARE FAMOUS WHEN THEY CAN SPELL YOUR NAME IN KARACHI.

It can be seen here that both partners are constantly repeating to each other what they are saying and thus sharing which part of the problem they are working on. Another example can illustrate a different approach to problem solving: even in cases when a subject mentions a technical problem and the partner reacts, it may happen that the first person does not want to pursue it further but prefers to concentrate on problem solving. This occurred several times for a number of couples, and often followed different technical interruptions. Here, L5G5 contributed to efficient task solving by keeping the partner's attention on the task:

L5G5 Poster 11:20
There are several occasions when L5 works and G5 either works or not because he is commenting on problems with the technology or making observations about the environment, etc.
L5: Could you take your 3's [the words with the number 3 in front of them]?

G5: It is so bad that they are so low.
L5: Low???
G5: Yes, I have to bend down.
L5: Come along [with the task]! Did you find any?
[A few minutes later:].
L5 and G5 "read" together their words: A MAN WHO DOES NOT MAKE MISTAKES DOES NOT USUALLY DO ANYTHING.

This couple chose an efficient strategy in solving the Posters from the start. They took positions in a room such that L5 could easily see half of the room and read the posters well while G5 could do so with the other half. Many of the other couples took similar positions at some point during the problem solving.

The following example also shows that it is important to push the problem solving forward, even though one is not completely sure of the results. For the Landscape task, the couples were rather unsure on several occasions whether they had counted the buildings correctly or not. This was hard without feedback, references, and marking possibilities.

L6: "The 34th was this. Or?"
G6: "OK. But . . . we are going to continue. Straight away. . . ".

Active participants and strong opinions often caused conflicts, as in the following example, where G6 does not agree on the same solution as L6. They stand in front of each other, face each other and argue. Neither of them backed down or wanted to continue with another proverb.

L6G6 Poster 48:30
L6: WHEN EVERYTHING IS BAD IT IS GOOD TO KNOW THE WORST.
G6: No. In the opposite way. WHEN EVERYTHING IS THE WORST THE BAD CAN BE GOOD.
L6: Here you have THE WORST. Because of that I think WHEN EVERYTHING IS BAD IT MUST BE GOOD TO KNOW THE WORST.
G6: Noooo! . . . I think it is one. . . .
L6 ⇒ But I think so!
G6: I don't understand why this should be a proverb in this way! If. . . .
L6 ⇒ [explaining] WHEN EVERYTHING IS BAD IT MUST BE GOOD TO KNOW WHAT IS WORST. . . .
G6: Hah! Hah!
L6: [agree] Yes, a strange proverb, I should say. [She still believes that she is right.] For example: There is sun today. Hmmm. OK. Today . . . Yes, today.

G6: No! Yes, no! TO KNOW IT . . . TO KNOW IT GOOD . . . You cannot say "TO KNOW IT GOOD".
L6: TO KNOW IT GOOD.
G6 ⇒ Can't we say. . . ?
L6: We must say WHEN EVERYTHING IS BAD because the only verb we have is . . .
[L6 and G6 argue with each other. L6 does not give up with her version.] 50:00
L6: I'm positive. I'm sticking to this absolutely!

The next example shows a conflict between L3 and G3. At the beginning, L3 wishes to explore the village alone, but G3 wants to do it together. After approximately half an hour G3 suggests that he should do it alone, which frustrates L3 since it was his idea all along.

L3G3 Landscape 15:00
G3: Do you think this is a circular road?
L3 Yes. This village is not so big.
At this point G3 modifies the strategy again.
L3: Let's count the houses, G3!
[But G3 suggests that they should follow the road in opposite directions and see if they meet again. L3 does not like it but agrees. He makes shortcuts through the lawn and becomes impatient.]
L3: Should we start counting now?
G3: Yes, this is the middle of nowhere. [He is irritated.]
L3: How would you like to proceed now? [Tries to be polite anyway.]
[G3 indicates that he does not care. He suggests again that they should go on their own and each count half the village.]
L3. Yes. Can you count the houses? [Sarcastically.]
G3: Yes, I will count the houses. [At this point G3 follows the strategy that originally was suggested by L3 although he had earlier disapproved of it.]
L3: I leave you to explore this part. I go back to the center of town. OK?
G3: Yes.
L3: Then we meet again.
G3: OK.

Sometimes people changed their methods in problem solving because of technical problems. For example, L6 had problems locomoting over a grassy lawn; she stopped G6 even though they had decided previously to locomote to some part of the town through the lawn.

L6G6 Landscape 28:00
L6 ⇒ I'm going on the way instead. I refuse to go out in the quicksand.

G6: Hah. Hah. Do you see me? I'm staying with the traffic sign on the way.
L6. This lawn seems to be dangerous.
G6: Hah. Then we should keep to the asphalt.

We can now turn to the end stage of the collaboration. L5G5 were generally eager to solve the tasks. The following example shows a short sequence with rapid conversation. The participants are speaking to each other and arguing, not bothering if they make mistakes and they are speeding up the work at the end to solve the puzzle:

L5G5 Puzzle approx 27:00
L5: Only a cube wrong.
G5: And it is right on two sides, only here. . . .
L5 ⇒ Do you see a white one here? Yes, I must. . . .
G5 ⇒ Were there any white ones here? On the wrong one?
L5: No. . . .
G5 ⇒ Yes, it is white on this side!
L5: No!
G5: Yes.
L5: No, but watch this. I didn't see this before. Yes, but then . . . Look at this, boy! This is done now!
G5 ⇒ Wait a little bit now! Wait a little bit now! Is this so?
L5 ⇒ So! This is done!
G5: Yes, this is right in all directions. I see it now!
L5: Me too! Me too!
G5 ⇒ Then we can . . . Then we can shake hands now!
L5: Yes, it is done! Yes, we can!
[They try to shake hands. Then, after several minutes:]
L5: You . . . you are in the middle of the cube . . . you. [for the first time they comment that they can go through the objects] . . . do you feel it?
[Both are laughing.]

People followed the convention of saying goodbye, but also commented more on the simulated avatars and objects at the end too—in other words, they can focus on these elements once the task is finished.

8. Discussion

8.1. Interaction Via Technology

What lessons can be learned from analyzing fragments of interaction in collaborative situations? For interaction via technology we have identified several problems in locomotion, orientation and handling the tasks. Many of the problems for locomotion and orientation showed how people experience clumsy movements, disorientation, problems in following each other, and the like. Other problems in interacting with technology were to due to shortcomings in the operation of the system. Such shortcomings are quite common with a complex new technology that is not yet robust. These included problems with grasping and aligning objects, poor audio, and colliding or passing through objects leading to disorientation. These are immediate usability problems that can be treated in the design of CVEs. One point to highlight is that these problems were often not noticed, or they were mentioned by participants but ignored.

Another type of problem are those arising from the asymmetrical technological settings. Examples here include the person using the Gothenburg IPT having problems manipulating or identifying objects. In these cases the participants often did not know that their partners did not have the same kinds of problems, and this created frustration when they found out that they were handicapped. The misunderstandings about interaction with technology, in turn, had an impact on social interaction and problem solving (see also chapter 7 in this volume for a further discussion of technical inequalities and social effects).

One important set of fragments illustrates how people treated virtual objects: Sometimes participants ignored objects as if they did not exist but seconds later they would treat them as real objects. Similarly, sometimes they treated the partner's avatar as if it was not there, going through it forward and backward, but at other times they apologized when colliding with it. This observation was true for all tasks and all couples (a previous study has also mentioned this, though without explanation, see [10]). This phenomenon raises questions for the study of presence, but it also raises questions about how objects and people should be represented if they can be ignored in one moment but in another treated as if they were real. If participants do not need consistency in this respect, then under what conditions is one or the other mode (treating people and objects as real, or the opposite) or switching between the two problematic? It can be noted here that it is often possible to tell from a participant's focus of attention whether they will treat avatars and objects as real or not.

At the same time, again, participants often easily coped with or ignored difficulties in interacting with technology, including difficulties with devices, with audio, or with interruptions due to technical problems. This includes using the constraints of the real environment in an odd way: figure 5-8 shows a person who focused on the task, using the devices, and automatically using her right hand to avoid collision with the wall. Does this mean that she has less presence since she knew about the walls? Or does it mean that she has higher presence because she had learned to "use the walls" in a similar way as she learned to use the devices? In short, participants learned during the tasks how to use devices, to adjust to the audio, and accept interruptions. At the end of the day, people had also adapted to the technical constraints, cables and working within the walls of the IPT systems.

8.2. Social Interaction

As for social interaction, we have highlighted the role of small conversational fragments. These can be easily overlooked unless recordings are viewed repeatedly, but they also often play an important role. They were more frequently identified as beneficial in that they kept the conversation going. Moreover, encouraging each other and acknowledging each other had a positive effect on problem solving. We have also seen that keeping the conversation going can help problem solving, but that missing communication cues can cause confusion as participants do not know who should speak next. Long periods of silence, especially for the first four tasks, were confusing (perhaps less so in the final task because both could work alone or engage with each other as required). Counting houses together in the Landscape task, on the other hand, and reading words aloud to each other for the Posters and Whodo tasks, helped problem solving.

In this trial there were problems when people did not share the same skills, for example with different language backgrounds, or if they regarded their partner as clumsy or unwilling. However, good intentions and effort promoted the collaboration. For problem solving more specifically, keeping the conversation going (for example, not letting long periods occur without any feedback, especially in the beginning of conversation) increased the effectiveness for problem solving. A frequent pattern of behavior that was observed was that the subjects occasionally glanced at each other in order to maintain awareness. In between conversation, pointing and helping the partner (non-verbally) also contributed to efficient collaboration.

Awareness of their partners was also achieved by asking where the other was if they did not see her or him, discussing their appearances (mainly in the beginning), apologizing if they collided with the other's avatar, and mentioning when they went through each other's avatar. We also observed that the partners could adapt to one another's behavior, for example if the partner always complained because of technical disturbances. This means that the awareness of the other partner, and what one must do to have sufficient contact with the partner, influences both problem solving and technical interaction. To foster better awareness of each other and avoid frustration, a first step would be to make known from the start the differences in technology and the main differences in each other's backgrounds. Simply letting the partner know at the outset that the other needs more help with the language or the technology, for example, would counteract unnecessary frustration during problem solving.

8.3. Problem Solving

We have presented evidence that people interact with each other differently in the different phases of collaboration. Further, we have focused on strategies during problem solving. We highlighted that the main difficulties that directly influence effectiveness and experiences are related to strategies. It is too easy to choose the "trial-and-error" strategy first. Misunderstandings were caused by difficulties in following through on the strategies, changing strategies, and continuing with the same strategy with very little modification.

There were situations when people did not have any clear strategy and hence did not know how to proceed with the task. There were also confusions around collaboration, misunderstanding each other's intentions, and frustration when one could not make the other person aware of one's situation (see also [10]). In a successful situation, people would check a strategy, follow up steps to reach their goal with each other, and have enough feedback to be sure that they performed correctly. All these ways of handling strategies have an impact on supporting or disturbing task-focused collaboration.

9. Conclusions

This chapter has presented a detailed study of collaboration in a CVE using immersive VR systems by examining three processes: interaction via technology, social interaction, and problem solving. Through numerous examples of fragments, we have illustrated the benefits of examining fragments very closely for the design of more usable CVEs. Clearly, the three areas that we have kept analytically distinct are interrelated in practice. We suggest, however, that the distinction between the three can help to address, if not overcome, the problem that analyzing CVEs always seems to depend on the context—including different technologies and different tasks or applications. By separating these three areas, we can perhaps find common properties in different contexts; for example if interaction with a certain type of technology always causes certain problems regardless of task or collaborators.

These observations raise larger issues about how to improve the usability of CVEs. Is it better to improve the systems and features of the environment, or to improve the users' awareness of their activities and settings? Our observations also raise broader questions about the aims, development and uses of CVEs, such as whether the aim should be more intuitive devices and greater realism, and what purposes CVEs are best suited for. One lesson that can be derived from the fragments that we have analyzed is the greater importance of social problems and task-solving problems as opposed to problems associated with technology. We saw, for example, how the subjects often ignored or got used to interacting with the technology. The implication is that potential solutions to these problems should also focus on these two areas rather than on technology issues. However, there may be technological solutions to problems in social interaction and problem solving. For example, it may be possible to develop tools for synchronizing the activities of the participants, or to create computer-aided awareness.

This leads to a more general point, which is that the successes and failures in these processes are something that the outside observers can easily become aware of through the method presented here. These same processes, however, are something that the participants were not necessarily aware of. This can be seen in the examples when the participants successfully coped with the various "unnatural" aspects of this setting and when they failed to overcome problems that they could easily avoid if they became more aware of the situation.

One of the benefits of analyzing fragments over the course of a longer collaboration is to examine whether, in the CVE, people follow conventions of interactions from their everyday experiences in the physical world. Since participants in this case collaborated over a longer period, we can argue that breaks with these conventions are unlikely to have been due to the novelty effect of the technology or the situation. Our examples show that whether or not conventions are maintained might depend on the focus of attention of the participants, what they are preoccupied with, how long they have collaborated and what phase of collaboration they are in.

In future work, it will be useful to gauge the relative weight or importance of these three processes more closely and to systematize the fragments according to the conditions in which they have wider applicability. Many of the fragments that have been analyzed here within each of the three areas will be found in a range of tasks in immersive spaces. By examining these fragments closely, we have highlighted some typical problems and successes. In view of the fact that working together in distributed systems is likely to become ever more widespread and complex, it is all the more urgent to develop systematic approaches for improving the usability of immersive CVEs.

References

1. Cruz-Neira, C., Sandin, D., & DeFanti, T. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In *Proceedings of SIGGRAPH 93*. New York, pp. 135–142.
2. Slater, M., Sadagic, A., Usoh, M., & Schroeder, R. (2000). Small-group behavior in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments, 9*(1): 37–51.
3. Schroeder, R., Steed, A., Axelsson, A.S., Heldal, I., Abelin, A., Wideström, J., Nilsson, A., & Slater, M. (2001). Collaborating in networked immersive spaces: As good as being there together? *Computers & Graphics, 25*(5): 781–788.

4. Gabbard, J.L. & Hix, D. (1997). *A Taxonomy of Usability Characteristics in Virtual Environments*. Report. Blacksburg, Virginia Polytechnic Institute and State University.
5. Kaur, K. (1998). *Designing Virtual Environments for Usability.* Doctoral dissertation. City University, London.
6. Bowman, D., Gabbard, J., & Hix, D. (2002). A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence: Teleoperators and Virtual Environments, 11*(4), 404–424.
7. Tromp, J. (2001). Systematic *Usability Design and Evaluation for Collaborative Virtual Environments.* Doctoral dissertation. University of Nottingham.
8. Heldal, I. (2003). *Usability of Collaborative Virtual Environments. Technical and Social Aspects.* Licentiate thesis. Chalmers University of Technology.
9. Steed, A., Spante, M., Heldal, I., Axelsson, A.S., & Schroeder, R. (2003). Strangers and friends in caves: An exploratory study of collaboration in networked IPT systems for extended periods of time. In *Proceedings of ACM SIGGRAPH 2003 Symposium on Interactive 3D Graphics April 27–30,* Monterey, California, pp. 51–54.
10. Tromp, J., Steed, A., & Wilson, J. (2003). Systematic usability evaluation and design issues for collaborative virtual environments. *Presence: Teleoperators and Virtual Environments, 10*(3): 241–267.
11. Robinett, W. (1992). Synthetic experience: A proposed taxonomy. *Presence: Teleoperators and Virtual Environments, 1*(2): 229–247.
12. Zeltzer, D. (1992). Autonomy, interaction, and presence. *Presence: Teleoperators and Virtual Environments, 1*(1): 127–132.
13. Schroeder, R. (2002). Copresence and interaction in virtual environments: An overview of the range of issues. *Fifth International Workshop on Presence, October 9–11*, Porto, Portugal, pp. 274–295.
14. Draper, J.V., Kaber, D.B., & Usher, J.M. (1998). Telepresence. *Human Factors, 40*(3): 354–375.
15. Steed, A., Slater, M., Sadagic, A., Tromp, J., & Bullock, A. (1999). Leadership and collaboration in virtual environments. In *Proceedings of the IEEE Virtual Reality*, 13–17 March, Houston, Texas; USA, pp. 112–115.
16. Stanney, K.M. (Ed.). (2002). *Handbook of Virtual Environments: Design, Implementation, and Applications*. London: Lawrence Erlbaum Associates.
17. Bowman, D.A. (1999). *Interaction techniques for common tasks in immersive virtual environments. Design, evaluation, and application.* Doctoral dissertation. Georgia Institute of Technology.
18. Bowers, J., Pycock, J., & O'Brien, J. (1996). Talk and embodiment in collaborative virtual environments. In *Proceedings of Conference on Human Factors in Computing Systems*, *CHI'96,*April 13–18, Vancouver, British Columbia, Canada, pp. 58–65.
19. Bente, G., & Krämer, N.C. (2002). Virtual gestures: Analyzing social presence effects of computer-mediated and computer-generated nonverbal behaviour. *Fifth International Workshop on Presence*, *October 9–11,* Porto, Portugal, pp. 233–244.
20. Heldal, I. & Schroeder, R. (2002). Performance and collaboration in virtual environments for visualizing large complex models: Comparing immersive and desktop systems. In *Proceedings of Virtual Systems and Multimedia, VSMM'2002*, 25–27 September, Gyeongju, Korea, pp. 208–220.
21. Sadowski, W. & Stanney, K. (2002). Presence in virtual environments. In K.M. Stanney (Ed.), *Handbook of Virtual Environments: Design, Implementation, and Applications.* London: Lawrence Erlbaum Associates, pp. 791–806.

22. Heldal, I., Schroeder, R., Steed, A., Axelsson, A.S., Spante, M., & Wideström, J. (2005). Immersiveness and symmetry in copresent situations. In *Proceedings of IEEE Virtual Reality*, 2005, Bonn, Germany, pp. 171–178.
23. Yang, H. & Olson, G.M. (2002). Exploring collaborative navigation: The effect of perspectives on group performance. In *Proceedings of Collaborative Virtual Environments, CVE 2002,* 30 September- 2 October, Bonn, Germany, pp. 137–142.
24. Heldal, I. (2004). Usability development for collaborative virtual environments. *VIRART workshop: Designing and Evaluating Virtual Reality Systems*, 22–23 January, Nottingham, England.
25. Stanney, K.M., Mollaghasemi, M., & Reeves, L. (2000). Development of MAUVE, the multi-criteria assessment of usability for virtual environments system (Final Report, Contract No. N61339-994-C-0098). Orlando, FL: Naval Air Warfare Center-Training Systems Division.
26. Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction Design: Beyond Human-Computer Interaction*. New York: John Wiley & Sons.
27. Spante, M., Heldal, I., Steed, A., Axelsson, A.S., & Schroeder, R. (2003). Strangers and friends in networked immersive environments: Virtual spaces for future living. In *Proceedings of HOIT 2003, Home Oriented Informatics and Telematics*, April 6–8, Irvine, CA, USA.
28. Heldal, I., Steed, A., Spante, M., Schroeder, R., Bengtsson, S., & Partanen, M. (2005). Successes and failures in copresent situations. Forthcoming in *Presence: Teleoperators and Virtual Environments*.
29. Hindmarsh, J., Fraser, M., Heath, C., & Benford, S.J. (1998). Fragmented interaction: Establishing mutual orientation in virtual environments. In *Proceedings of ACM conference on computer supported cooperative work, CSCW'98*, 14–18 November, Seattle, WA, USA, pp. 217–226.
30. Frecon, E., Smith, G., Steed, A., Stenius, M., & Stahl, O. (2001). An overview of the COVEN platform. *Presence: Teleoperators and Virtual Environments, 10*(1): 109–127.
31. Steed, A., Mortensen, J., & Frecon E. (2001). Spelunking: Experiences using the DIVE System on CAVE-like Platforms. In B. Frohlich, J. Deisinger, H-J. Bullinger (Eds.), *Immersive Projection Technologies and Virtual Environments*. Wien: Springer-Verlag, pp. 153–164.
32. Fjeld, M., Lauche, K., Bichsel, M., Voorhorst, F., Krueger, H., & Rauterberg, M. (2002). Physical and virtual tools: Activity theory applied to the design of groupware. *Computer Supported Cooperative Work, 11*(1–2): 153–180.

Chapter 6

THE IMPACT OF DISPLAY SYSTEM AND EMBODIMENT ON CLOSELY COUPLED COLLABORATION BETWEEN REMOTE USERS

David Roberts, Robin Wolff and Oliver Otto

1. Introduction

Trends towards greater collaboration between organisations increase the need for effective, efficient and safe ways to collaborate within distance teams. Technology has already greatly reduced the need for face-to-face meetings. Telephones, text messaging, email, web, and classical video conferencing have all thrived in supporting specific aspects of tele-working. There is, however, still a need for face-to-face meetings, even though the cost to business individuals and the environment can be significant. The holy grail of tele-collaboration is to support the full range of communication used within a co-located group.

Psychologists categorise social communication between humans as verbal, non-verbal, the role of objects and that of the environment [1]. Immersive displays surround the senses within an information world, which, compared to desktop systems, is believed by some to increase the feeling of presence and by others to increase task performance. Immersive Collaborative Virtual Environments (ICVE) allow a number of people to share an interactive synthetic experience from a true first person perspective. The use of these technologies in tele-immersion allows geographically separated people to interact using a variety of verbal and non-verbal communication within a shared meaningful environment and through shared information objects. We believe that this is the first technology to support these four primary categories of social human communication in a natural and intuitive way.

Linking walk-in displays, such as Caves, supports unprecedented naturalness of communication between physically remote people by placing them together in a shared scene in which they can naturally move around, talk, gesture and manipulate shared objects. We have investigated the impact of display

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 131–149.

device and embodiment on the perception and performance of collaboration during a shared task requiring various forms of social human communication and ways of sharing objects. The task, building a garden gazebo, has been routinely evaluated in sustained trials between as many as four linked interactants, distributed between four sites in the UK and Austria.

1.1. Placing People in an Information and Social Context

Inhabited Information Systems (IIS) encapsulate the principle of placing people within an information context [2]. Immersive virtual reality technology is well suited to this, as it physically places a person within a 3D scene where one can naturally look around and move. Geographically separated people are able to interact and communicate with each other when linking such immersive display interfaces through a CVE, which usually represent a remote user by a human-like articulated character, the avatar. A tracking system placed on the head allows natural control of gaze, while another on a hand can allow an object to be pointed at or reached for. Within the physical confines of the display and tracking system, participants are able to naturally walk around, walk up to and face each other and objects, thus supporting natural social behaviour such as proxemics, communicational gaze and gross gesturing. In the natural world, gaze is extremely important for controlling conversational turn taking, representing attention and emotion. Most immersive systems track gaze to control the viewpoint and this may be communicated to remote participants. However, most present systems do not distinguish between head and body rotation, thus communicating horizontal gaze movement by moving the avatar's torso.

The predominant method for immersion was for many years the Head Mounted Display (HMD). Large screen display systems, such as cubic walk-in (Cave-like) displays, Workbenches and Panoramas, have gradually gained acceptance and are now predominant in industrial applications. Like HMDs, these place people within the scene and offer natural control of gaze and reach, but Panorama and walk-in displays offer far greater field of view than typical HMDs, additionally reducing the risk of motion sickness and nausea. Another important distinction between HMD and walk-in displays is that the user can observe his physical body within the scene, which is thought by some to increase the feeling of presence [3].

1.2. Level of Cooperation

Levels of cooperation within CVEs have been categorised by a number of research groups in similar ways: Ruddle *et al.* described the different levels of cooperation as level 1—co-existence and shared-perception; level 2—individual modification of the scene; and level 3—simultaneous interactions with an object [4]. A similar taxonomy was presented for haptic collaboration that describes the respective levels as Static, Collaborative and Cooperative [5]. Our studies provide a more detailed taxonomy of level 3, which will be described later.

1.3. Impact of Display Configurations

Closeness of collaboration depends upon the supported level of communication. We now look at the impact of display configuration on characteristics associated with the feelings of presence and co-presence across the forms of social human communication described above. The majority of CVEs have been experienced through desktop interfaces. The limited field of view and unnatural viewpoint–control characteristic of desktop interfaces introduced problems in navigation and observation. A technique for increasing the field of view by introducing an avatar to represent the local participant and attaching the viewpoint behind the head has become well established through the computer games industry. However, this approach does not utilize a human's peripheral vision and is difficult to control from natural head gaze. The former may be addressed by providing a desktop surround display, e.g. from three monitors joined in an arc [6]. The latter can be addressed through an immersive display with head tracking. Our studies attempt to address both former and latter by using a walk-in display with head tracking.

Significant problems in communicating the referencing of objects of interest within a crowded environment were experienced through desktop interfaces [7]. When participants were given the task of relocating furniture in a shared virtual room, more talking was dedicated to identifying objects than to deciding where to put them. This problem was reduced but not removed by again placing the viewpoint behind the local avatar.

Several studies have investigated the effect of linking various combinations of display systems for collaboration. It was found that immersed users naturally adopted dominant roles versus desktop users [8]. A study by Schroeder *et al.* investigated the effect of display type on collaboration of a distributed team [9]. Their work extended the concept of a Rubik's cube by splitting the composite cube such that two people could concurrently interact with individual component cubes while observing each other's actions. The study compared three conditions based on display combinations: two linked walk-in displays; face-to-face; and a walk- in display linked to a desktop. The primary finding was that the asymmetry between users of the different systems affects their collaboration and that the co-presence of one's partner increases the experience of the CVE as a place.

1.4. Naturalness of Interaction

Immersive displays are typically imbalanced in terms of input and output bandwidth to the senses. We have observed that most people seem very impressed on entering a walk-in display and observing the immersive rendering but become quickly disappointed when they attempt to interact with the scene. There is considerable scope to improve the naturalness and intuitiveness of interaction in today's systems. Two important factors are multi-sensory alignment and responsiveness.

A walk-in display system gives a perfect alignment of visual and kinematical senses as the user can see his or her own body moving within the space. It is well accepted that the feeling of touch through the haptic sense improves the realism of many simulated tasks, for example in medical training. Multi-user haptic systems are now gaining maturity with control systems used to stabilise interaction in the face of network delays. However, haptic rendering requires considerably higher and less jittery frame rates than vision and combining the two senses in multi-user worlds has proved problematic [10].

Responsiveness is of prime importance. Latency in viewpoint changes following head movements appears to increase feelings of motion sickness and disorientation. Latencies in interacting with objects cause frustration. Responsiveness to interactions with the scene can be improved through localisation and replication. That is, each display system has a dedicated computer that contains a replication of at least part of the object database. The effect of a user's actions on the scene, including their viewpoint, can be calculated and rendered locally. Changes to the scene are sent across the network to update other replications. Consistency control is required to ensure that replications do not diverge to an unstable degree.

2. Experimental Set-Up

In order to investigate the effect of the display system and user embodiment on closely coupled collaboration in virtual environments as well as its impact on human communication and the CVE system, we designed a benchmark application that involves various forms of coordinated shared object manipulation.

2.1. Benchmark

Our benchmark allows the analysis of each form of social human communication across a variety of collaborative scenarios, including various forms of shared object manipulation. The task, the construction of a garden gazebo,

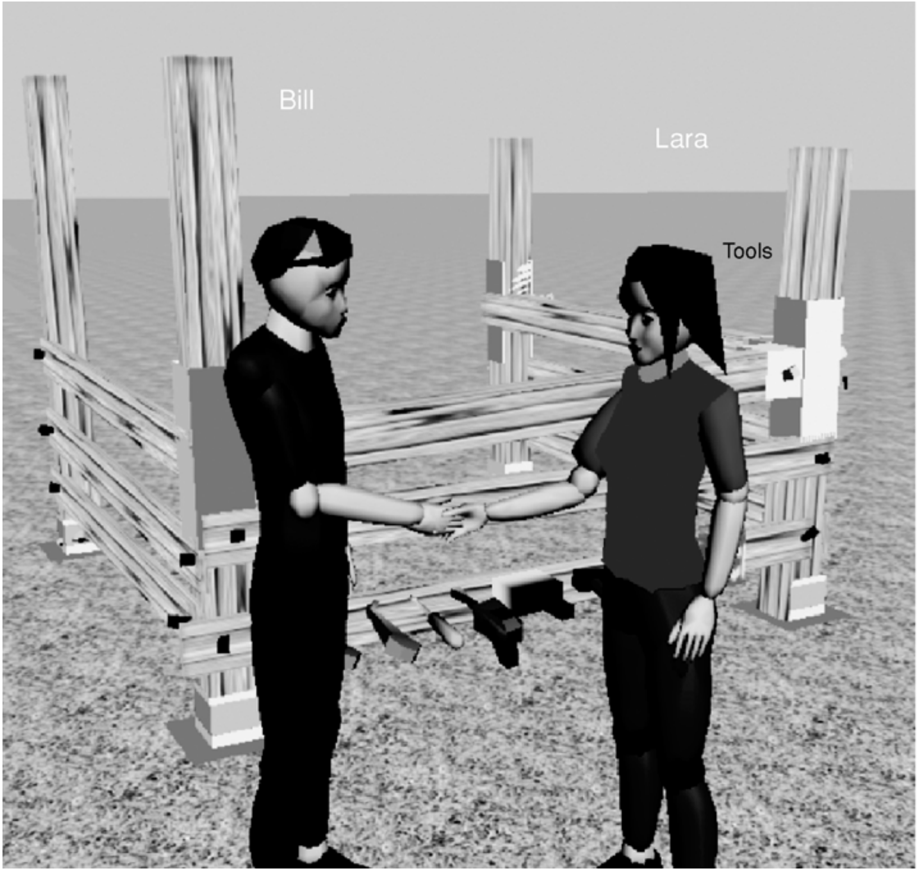


Figure 6-1. The Virtual Gazebo application.

cannot be done alone. The time taken to complete each scenario is a measure of the success of collaboration. Scenarios include planning, passing, carrying and assembly, each requiring a distinct mix of non-verbal communication and method of object sharing. Figure 6-1 shows a snapshot of the Virtual Gazebo and figure 6-2 shows a user accessing the application through a walk-in display.

The experiment starts by situating the user in a virtual garden setting, along with all the necessary building materials and tools scattered on the ground. Avatars that represent remote users appear as the rest of the team enters the garden. A user can pick up material and start to build the gazebo. However, constructing a gazebo on your own is not possible. The simulation of gravity intentionally prohibits leaving materials in thin air. Beams have to be held by one person while another fixes them with screws and a screwdriver tool. Moving, positioning and joining of beams require teamwork. Other tasks, such

Figure 6-2. A local user (left) interacting with a remote user (right) within a walk-in display.

as gathering materials, can be done sequentially, but still require coordination between the sub-tasks. The limited set of tools provokes competition for shared resources.

2.2. *Collaborative Scenarios*

We now examine the four scenarios of planning, passing, moving and fixing, as summarised in table 6-1. In the planning phase, users reference and discuss shared objects. For example, deciding which material to start with, where to get it from and where to carry it, will make use of both spoken word and

Table 6-1. Scenarios of object sharing.

Scenario	**Figure**	**Description**	**Method of sharing**
Planning	6-3a	Discussing how to proceed.	Referencing objects and environment.
Passing	6-3b	A tool or material is passed from one user to another.	Sequential sharing and manipulation of the same object attribute.
Moving	6-3c	A wooden beam is too heavy to lift alone requiring one user to lift each end.	Concurrent sharing of an object through the same attribute.
Assembling	6-3d	A wooden beam must be held in place by one user, while another fixes it by drilling a hole and inserting a screw.	Concurrent sharing of an object through distinct attributes.

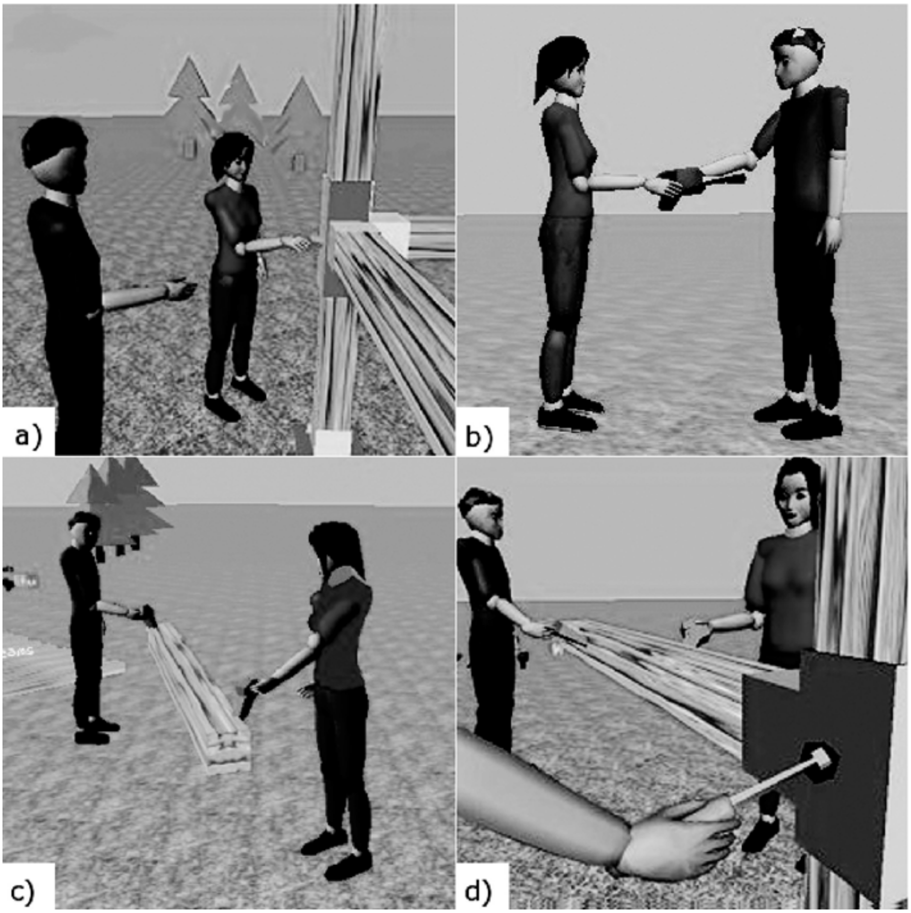


Figure 6-3. (a) Planning, (b) passing a tool, (c) moving a beam, (d) assembling an object.

referential gesturing. Passing an object, for example a screwdriver, between people requires sequential manipulation of the same object attribute, movement in this case. Moving an artificially heavy object, such as a beam, requires two people to share the manipulation of the movement attribute. Assembly often requires one person to hold a material while another drills a hole or inserts a screw. The existence of a hole or screw is defined as an attribute and so the object must be shared through distinct attributes if held while fixed. The effectiveness of communication of intention, attention and emotion, all impact on the performance of each scenario (see figure 6-3).

2.3. Referencing Objects

The performance of collaboration focussed on shared objects is dependent on the efficiency of communicating objects of attention. Other work has shown that this is problematic when using desktop interfaces [7]. We have earlier proposed the hypothesis that problems in referencing objects using desktop displays arise from restricted field of view and indirect control of gaze and pointing. We further proposed that the use of walk-in displays should reduce this. Although not designed for the purpose, the gazebo application is well suited to test our hypotheses. The construction site environment is similarly cluttered as the room of furniture in Hindmarsh *et al.* [7], but has a greater diversity in size of objects and scale of environment. The task of building the gazebo routinely requires communication of the referencing of objects as well as the place within the environment that they are to be taken. Communication of referencing must reflect nuances of speech and gesture and the interface must not restrict the recipient from capturing these. When using a walk-in display, control of gaze and pointing are driven through a tracking system and the user is surrounded by the display surface to the front, both sides and the floor. The complexity of the task requires the collaborative planning of a number of steps, which may involve several collaborators and objects. A wide field of view and direct control and communication of gaze and pointing should allow efficient referencing, location and identification of each.

A typical conversation between two collaborators—Bob and Lara in this case—engaged with moving a heavy beam using carry tools, is reproduced below.

Bob: Hey, let's move this beam over there. [Points with his hand to the beam in front of him and to the left.]
Lara: [Rotating to see Bob and then follows his hand movements.]
Lara: Ok, I will take this carry tool here. [Points to the tool and moves to pick it up.]
Bob: I'll take the other tool then. Where is that?
Lara: Just next to you.
Bob: Ah, ok. [Rotates and picks up the tool.]
Lara: I have my end of the beam now.
Bob: Yup, I am right with you. Ok now let's move it over there.
Lara: I am following your direction.

2.4. Implementation

The Virtual Gazebo has been implemented on top of the established CVE platform DIVE [11] in version 3.5. Each remote user is embodied through a human like avatar. A non-jointed avatar represents the non-immersed desktop user, whereas the immersed user's avatar shows dynamic head and arm-articulation controlled from head and hand tracking data. Our application included various interactive objects that imitated the behaviour of building materials and tools.

An object's behaviour has been implemented in the form of DIVE/Tcl scripts that describe a set of procedures to change an object's state in a specific way. For example, when a screwdriver intersects with a screw, all objects intersected by the screw are fixed together. All behaviour scripts are reactive and triggered by specific events. These are update messages generated by the CVE system to update replicated versions of the distributed virtual environment. DIVE supports several event types. These include object transformation events, such as movement or rotation; object interaction events, such as grasp, release or select events; object collisions; and changes to object-specific properties and flags. Most functionality of the Virtual Gazebo is triggered by collisions of material and tool objects. For example, when a drill tool is held closely to a material object so that they collide, the resulting collision event would trigger a procedure in the material object's behaviour script to increment the property that counts the number of holes in the object.

Certain object behaviour scripts additionally provide a level of consistency control within the application level. For example, during most interactions a user-defined flag is set to signal a definite object state as an acknowledgment of the successful action of a tool. Hence, a level of causal ordering and discarding of events is realised by constraining the order of manipulations through such flags and properties that must be set in a certain order. For example, a hole must be drilled between two materials and then a screw inserted through both, before they can be fixed together with a screwdriver. The material conditions of possessing a hole and being fixed are each achieved in the code through setting a respective flag. If such a condition was not fulfilled the action would not be successful and the current event discarded, forcing the user to repeat the step.

Early research claimed that generic concurrent manipulation of shared objects would be not possible unless the CVE system fulfils hard real time and reliability constraints [12]. We relaxed these constraints and, again, provided some level of concurrency control through the application level. Allowing script procedures to set several distinct properties concurrently has enabled the concurrent object manipulation of distinct attributes, avoiding exclusive object ownership. For instance, a counter attribute of an object could be set while its position was continuously updated. In contrast, concurrent manipulation of the same attribute has been realised through intermediate procedures that gather and process events before updating the actual attributes of the manipulated object. For example, carrying a beam together with a remote user would be performed with help of specialised carry tools, used at each end of the beam to attract its transformation. These tool objects send their current position to the carried object, which in turn attempts to find an average transformation between them and finally communicates the result to the remote peers that update the scene accordingly. Instead of manipulating the carried object directly, the users interact through intermediate objects that, due to their relatively little overhead in behaviour scripts, show responsive feedback. Hence, carry tools are used also to hide the effect of network latency while concurrently sharing the manipulation of objects.

2.5. Effects of Remoteness

All four scenarios of collaboration have distinct requirements on responsiveness and consistency of the system. For example, the representation of gestures must be sufficiently complete to be recognised and their subtleties understood. Nuances, relating to referencing or manipulating objects, must be sufficiently communicated. This requires timely capture and synchronisation. Communication of events across a switched network, such as the Internet, introduces the possibility of delay, loss and disorder of events. Although this can be addressed by reliable protocols, these introduce dependencies between sender and receiver that can increase jitter and, thus, affect synchronisation. The vast majority of events in most CVE applications represent movement. As movement can be communicated atomically, synchronisation can be increased by discarding all but the most current movement event for each object.

Excessive delay in communication between users can confuse concurrent manipulation of objects, turn taking in conversations, as well as sequential object manipulation. Although reductions in consistency in avatar representation may result in confusion and lower the performance of collaboration, they are unlikely to produce unwanted outcomes in the overall task, unless understood. The loss of events, such as the drilling of a hole, insertion of a screw and tightening the screw with a screwdriver, can cause significant confusion, as can the loss of ordering between them. Furthermore, it becomes confusing if the movement events for an object are delayed and displayed after it has been fixed. Ideally, such events should be discarded before they affect the model. If not, the effect must be overridden quickly.

3. User Trials

Our benchmark has been used to support routine user trials over a two year period between four sites in Europe. Typical configurations of the involved display types can be seen in table 6-2. Initial trials, linking three walk-in displays at Reading and London in UK and Linz in Austria, demonstrated that supporting both verbal and natural non-verbal communication during shared object manipulation was achievable with today's CVEs, but not without significant network-induced inconsistencies [13]. For example, our initial prototype induced around two seconds of delay which was clearly noticeable, as actions appeared to lag behind their words. This problem was due to the CVE system becoming overwhelmed by the scale of tracking events and was fixed in a later

Table 6-2. Display configurations.

Display type	View	Input	Audio	Avatar*
Walk-in (CAVE-like)	Stereo	Tracked head & hand	Yes	Medium realism, dynamic body
Workbench	Stereo	Tracked head & hand	Yes	Medium realism, dynamic body
Desktop	Mono	Mouse & keyboard	Yes	Medium realism, static body

* Avatar: The remote representation of the local user.

prototype with a simple filter mechanism based on the magnitude of movement allowed to generate an event. We found that a good compromise between sufficient detail to support understandable non-verbal communication and sufficient synchronisation to achieve shared object manipulation was to only send movements of above 1cm. Many events below this seemed to be caused by tracking jitter rather than real human movement.

A second and more substantial problem was the loss of event messages vital to causal state changes of the shared objects during manipulation. Although intermediate movement events may often be lost without causing undue confusion, vital events, such as the fixing of a beam to a foot, cause considerable confusion when lost on a remote site. Although we found that the collaborators quickly learned to detect and rectify the problem by carrying out that part of the task again, this did cause frustration, a reported reduction in the feeling or realism and in productivity.

3.1. *Experimentation*

The subjects were all higher degree students from the three participating universities. Most were aged between twenty and thirty and were studying computer science related subjects. Teams of two or three people, in distinct locations, were given a short tutorial in using the application, told the goal of the task and left to plan and undertake it together. An average session took between a half and one hour, during which, both quantitative and qualitative measurements were taken for each scenario or sub-task. Quantitative measurements included the time to complete each sub-task along with the resultant event traffic. Qualitative measurements of the impression of collaboration were obtained through a post-questionnaire. Most questions were based on those of Usoh *et al.* [14]. Answers could be given on a scale of 1–7, where 7 represents total agreement and 1 total disagreement. Asking sets of related questions reduced errors arising from a user's misinterpretation of a question. Overall, more than one hundred students helped us to gather data and to study different aspects of such close collaboration.

We adapted our application to reduce the dependency on vital events and reduced the goal of the task to building one corner of the gazebo, thus eliminating redundancy in the scenarios. The first of these revised trials were undertaken between a walk-in display and two desktop systems at the University of Reading [15]. The purpose of the trial was to test the impact of asymmetric displays on supporting non-verbal communication including the shared manipulation of objects in comparison to a study by Schroeder *et al.* [9]. We found that novice users adapt quickly to the interface and remoteness of peers. Typically, after three sessions of approximately 15 minutes their performance doubles, approaching that of expert users (5 minutes to perform the task). Immersive users undertook most parts of the task far more efficiently than their desktop counterparts.

3.2. User Performance

Our benchmark requires close collaboration at numerous points. This means that a faster worker must often wait for the slower before beginning the next step. Schroeder *et al.* found that the perception of collaboration is affected by asymmetry between users of the different systems [9]. Our results showed that the time taken to complete a collaborative task is also affected. When roles in the collaborative task are ill-defined, the performance of the team equals that of the weakest link. However, the performance is greatly increased when the immersed user undertakes the more difficult part of every sub-task. The results of our questionnaire confirmed that the perception of contribution is affected by asymmetry of linked displays when carrying a beam [15]. However, this is clearly not the case when fixing a beam. This suggests that the interface plays a major role during the sharing of an object's attribute and a minor role when sharing an object through distinct attributes.

Surprisingly, neither the interface, nor the form of object sharing, is perceived to affect the level to which the remote user has hindered the task. This appears to contradict the results of the performance analysis above. From the perspective of immersed users, collaboration is considerably easier with a symmetric user. Desktop users, however, found the type of remote display to play little part in the degree of collaboration. Another finding of this trial was that walk-in displays are much more efficient than desktop systems in terms of team performance, especially when positioning objects. In some cases, however, a basic desktop interface may be easier to use, such as when holding an object in the air, as this is just a mouse click. The extent of usability and natural interaction, though, depends on application and interface.

Figure 6-4. A desktop interface with headset for verbal communication.

One would have expected verbal communication between remote users to become more natural when the technology is transparent, that is when the microphone and speakers are hidden. However, we observed a significant increase in verbal communication when the user is constantly aware of a familiar communication device, such as a headset with microphone and earphones [1]. When this was introduced in the trials (see figure 6-4 for a desktop set-up), the team worked together more successfully and each participant made greater use of the non-verbal communication influences. Verbal communication, compared with non-verbal communication, was perceived to be of the greatest importance. Little difference was perceived in the importance of the other influences.

The current state of technology is still some way from providing natural social human communication between remote participants. Novice users commonly commented that the lack of a feeling of touch made interaction with objects unintuitive. We addressed this in an extended study by placing a haptic system within the walk-in display [16]. In addition to possibly being the first group to do this, the exercise resulted in two important findings: Firstly, decoupling graphics and haptics rendering onto separate machines can maintain suitable frame rate, latency and jitter characteristics for both visual and haptic senses, while maintaining sufficient consistency between them. Secondly, the frame rate of visual representation affects the usability of the haptic interface.

3.3. Event Traffic

In order to design a consistency management scheme it is vital to have a detailed understanding of the requirements of interaction and collaboration and the characteristics of event traffic that is likely to be generated while supporting these between particular interfaces. We also used the Virtual Gazebo to study the event traffic across all three levels of collaboration, and across various symmetric and asymmetric display configurations, paying particular attention to the form of object sharing in levels two and three [17]. We looked at the impact of the four categories of collaboration on event traffic within a local area network at University of Reading through both a desktop system and a walk-in display system. The frequency of specific event types to update avatar movement, object manipulations and consistency control, were measured, analysed and compared.

We found that the interface provided by a particular display device, as well as the form of object sharing, have significant impact on the frequency of generated events. Event bursts, arising from greater magnitudes of human movement, occurred during shared object manipulation that often resulted in event queuing. This was often manifested through the jumping around of the shared object, with the magnitude of discontinuity depending on a combination of interface and type of object sharing. The bursts were exacerbated by events generated to ensure consistency of shared objects by bringing them to an objective state. We found that concurrent object manipulation can result in more traffic than sequential manipulation, whereas concurrent manipulation of the same attribute impacts more than distinct attributes. Again, erroneous results arose from the delay or loss of vital events, such as those that change the hierarchy of the scene graph. Vital events were rare but tended to coincide with or bound bursts of non-vital events. We concluded that a CVE does not yet exist which is capable of supporting applications like the Virtual Gazebo across walk-in displays, without unnaturally constraining the application and laboriously tuning event passing.

3.4. Relating Event Characteristics to Semantics

Another set of trials investigated the relationships between event characteristics and the semantics of human movement. These trials were undertaken between linked walk-in displays at the University of Reading and Johannes

Kepler University in Linz [18]. Again, the event traffic has been measured in terms of event frequency. But this time we focused on the motion of the tracked head and hand during collaboration and the resulting effect on the event traffic. We found that supporting natural interaction and non-verbal human communication implies increased event traffic for detailed movement representation of avatar and shared objects. The magnitude and importance of movements was shown to depend on the kind of collaborative scenario. Within our trials, the frequency of avatar movement events is fairly continuous and constitutes 64% of event throughput. Object movement events only occur during or shortly following interaction. The highest peaks in frequency of events come from shared object attribute manipulations and are caused by the added effect of the consistency mechanism. Vital events, which synchronise or trigger actions and are essential for steering the application, often bound the manipulation of a shared object attribute and sometimes coincide with it. Tracing movement across the various collaborative scenarios (figure 6-5) has revealed interesting results that

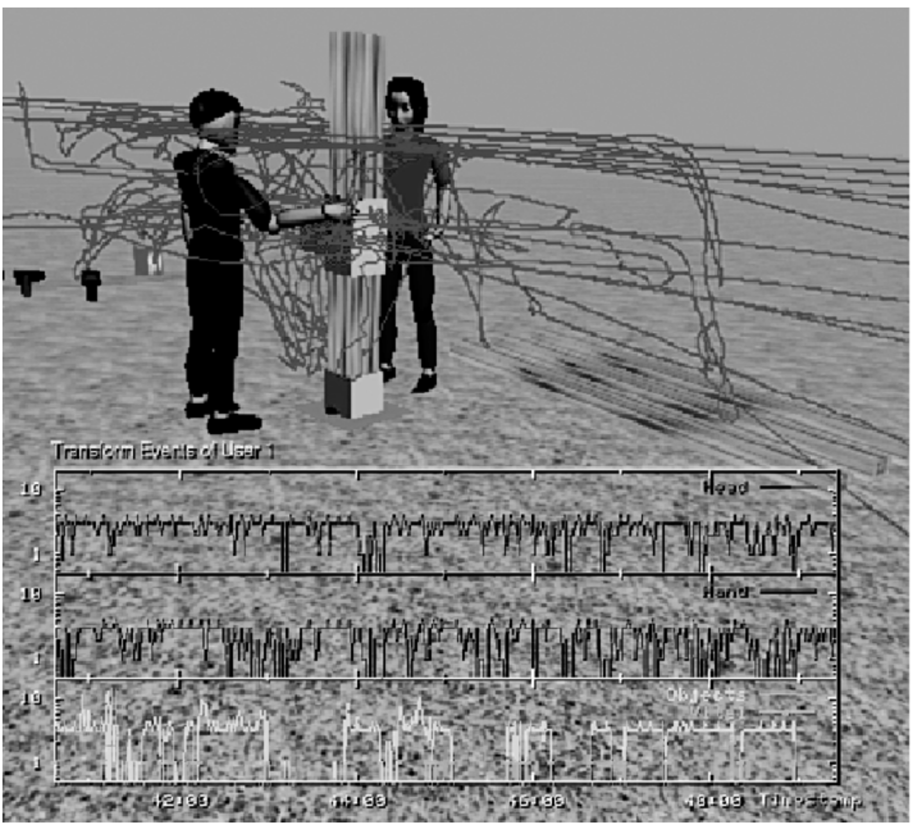


Figure 6-5. Traced head and hand motion of avatars.

offer potential for optimisation to reduce overall event traffic. Although the form of collaboration seems to play little part in the frequency of head and hand movement events, it plays a considerable role in the magnitude of change described by the events. Furthermore, only one member of the team seems to be highly animated at any one time.

The exclusive use of walk-in displays significantly improved the observable naturalness and performance of a distributed team, compared to the exclusive use of desktop, or of an asymmetric pair of both display types, as used in previous trials. As with the above furniture moving application [7], we observed considerable verbal communication related to the referencing of objects. However, we did not observe this leading to excessive or unnatural delays. However, when linking walk-in and desktop displays, we have observed such problems encountered by desktop users. Furthermore, we have found that users in walk-in displays perform much better than those using desktop systems in object placement tasks. We suspect there to be a number of factors behind this. Firstly, the walk-in displays have a much wider field of view than a desktop display and secondly, objects and others can be located through natural head glance which is much quicker and easier to control than moving the view of an avatar; thirdly, placing people within the environment, as opposed to allowing them to "look into" it, is likely to encourage natural proxemic behaviour including a sense of the relative position of people temporarily out of one's field of view.

4. Conclusion

Although advances in technologies, such as tele-conferencing, threaten the justification for the full-size avatars of immersive systems, they still offer considerable advantages in supporting social human communication, especially through and around shared objects. A variety of organisations have expressed a strong need for such tools to meet the growing demands of collaboration fuelled by globalisation. One example is in Aerospace, where engineers, often in different countries, need to discuss the routing of pipe work around an engine. Here, it is necessary to efficiently communicate which pipe is under discussion and what is thought about the route it follows. Different solutions may need to be explored by manipulating the pipe work in real time and a reliable collective opinion must be reached.

There is currently no commercial system capable of communicating sufficient levels of naturalness in attention, understanding and emotion to remove the need for a face-to-face meeting. Technologies such as tele-conferencing and pure tele-immersion can communicate gestures, posture and facial expression, however, neither objects of interest nor the environment are truly shared. Extensions, such as Access Grid, allow data to be presented in an adjacent window, but then the whole team is physically removed from it. Furthermore, participants are usually constrained to face into one direction. Collaborative Computer Aided Design and digital mock-up systems allow a model to be shared and some allow it to be collaboratively updated and annotated, but again they separate the people from the data. Although CVEs allow a fairer and more powerful method of sharing objects, previous trials have indicated problems referencing these objects between a team.

We hypothesised that the combination of walk-in display technology and CVEs would better overcome the separation of people and data by placing people within it such that they could use and see their own bodies interacting with data and others. We further hypothesised that the characteristics of these displays and their tracking interfaces would reduce problems in referencing objects. We classified various forms of closely coupled collaboration and related this to a set of distinct collaborative scenarios and built a benchmark application.

Our studies have shown that combining the characteristics of walk-in displays and collaborative virtual environments allow people to interact through and around shared objects with unprecedented levels of naturalness. We have demonstrated that even a particularly demanding application of this genre can be supported with current technology. Major failings reported of CVE technology were found to be mostly overcome using walk-in displays. Both qualitative and quantitative measures showed walk-in displays to perform better than desktop displays. Developing this application was hard work using current technology. A detailed investigation of the relationships between interaction scenarios and event traffic demonstrated requirements for consistency control and suggested optimisations for increasing scalability. In summary, this technology appears to take us considerably closer to removing the need to travel to support teamwork involving shared artefacts.

Future work that follows on from what has been presented here and that deserves exploring more closely includes communicational eye gaze, closely coupled visual-haptic collaboration and applications within health, culture and engineering.

Acknowledgements

The authors would like to particularly thank Anthony Steed, both for porting the DIVE CVE to immersive devices and for his high level input on the drawbacks of current technology. Thanks also goes to David Swapp at University College London, Dieter Kranzlmueller, Christoph Anthes, Paul Heinzlreiter, Jens Volkert and Johann Messner at Johannes Kepler University in Linz, Vassil Alexandrov, Ali Al-Khalifah and Detlef Krischker at the University of

Reading, all students, lecturers and everyone else involved in the user trials and the creators and developers of DIVE at The Swedish Institute of Computer Science.

References

1. Otto, O. & Roberts, D.J. (2003). Importance of communication influences on a highly collaborative task. In *Proceedings of The Seventh IEEE International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*. Delft, The Netherlands, pp. 195–201.
2. Snowdon, D.N., Churchill, E.F., & Frécon, E. (Eds.) (2004). *Inhabited Information Spaces: Living with Your Data*. London: Springer, 2004.
3. Schuemie, M.J., Straaten, P.v.d., Krijn, M., & Mast, C.A.P.G.v.d. (2001). Research on presence in VR: A survey. *CyberPsychology & Behavior, 4*(2): 183–202.
4. Ruddle, R.A., Savage, J.C., & Jones, D.M. (2002). Verbal communication during cooperative object manipulation. In *Proceedings of the ACM Conference on Collaborative Virtual Environments (CVE)*, pp. 120–127.
5. Buttolo, P., Oboe, R., & Hannaford, B. (1997). Architectures for shared haptic virtual environments. *Computers & Graphics, 21*(4): 421–429.
6. Lessels, S. & Ruddle, R.A. (2004). Changes in navigational behaviour produced by a wide field of view and a high fidelity visual scene. In *Proceedings of the 10th Eurographics Symposium on Virtual Environments (EGVE'04)*, Aire-la-Ville, Switzerland, pp. 71–78.
7. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (2000). Object-focused interaction in collaborative virtual environments. *ACM Transactions on Computer–Human Interaction (ToCHI), 7*(4): 477–509.
8. Slater, M., Sadagic, A., Usoh, M., & Schroeder, R. (2000). Small group behaviour in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments, 9*(1): 37–51.
9. Schroeder, R., Steed, A., Axelsson, A.S., Heldal, I., Abelin, Å., Wideström, J., Nilsson, A., & Slater, M. (2001). Collaborating in networked immersive spaces: As good as being there together? *Computers & Graphics, 25*(5): 781–788.
10. Park, K.S. & Kenyon, R. (1999). Effects of network characteristics on human performance in a collaborative virtual environment. In *Proceedings of IEEE Virtual Reality*, Houston TX, March 13–17, pp. 104–111.
11. Carlsson, C. & Hagsand, O. (1993). DIVE—A platform for multi-user virtual environments. *Computers & Graphics, 17*(6): 663–669.
12. Broll, W. (1995). Interacting in distributed collaborative virtual environments. In *Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS)*, Los Alamitos, CA, pp. 148–155.
13. Roberts, D.J., Wolff, R., & Otto, O. (2004). Pushmepullyou: The reality of interaction with shared objects in networked walk-in displays. In *Proceedings of the 17th International Conference on Parallel and Distributed Computing Systems (PDCS)*, San Francisco, CA, USA, no pp.
14. Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments, 9*(5): 497–503.
15. Roberts, D.J., Wolff, R., & Otto, O. (2003). Constructing a Gazebo: Supporting team work in a tightly coupled, distributed task in virtual reality. *Presence: Teleoperators and Virtual Environments, 12*(6): 644–657.

16. Seelig, M., Harwin, W., Roberts, D., Otto, O., & Wolff, R. (2004). A haptic interface for linked immersive and desktop displays: Maintaining sufficient frame rate for haptic rendering. In *Proceedings of 17th International Conference on Parallel and Distributed Computing Systems (PDCS)*, San Francisco, CA, USA, 2004, no pp.
17. Wolff, R., Roberts, D.J., & Otto, O. (2004). A study of event traffic during the shared manipulation of objects within collaborative virtual environments. *Presence: Teleoperators and Virtual Environments, 13*(3): 251–262.
18. Roberts, D.J., Wolff, R., & Otto, O. (2004). Supporting social human communication between distributed walk-in displays. Forthcoming in *Proceedings of ACM Symposium on Virtual Reality Software and Technology (VSRT)*, Hong Kong, November 10–12.

Chapter 7

THE GOOD INEQUALITY: SUPPORTING GROUP-WORK IN SHARED VIRTUAL ENVIRONMENTS

Maria Spante, Ann-Sofie Axelsson and Ralph Schroeder

1. Introduction

Most of our daily social interaction takes place face-to-face—this applies to work meetings as well as to personal encounters. Yet, in recent years, mediated interaction of various kinds and for various purposes has become more common, including the use of email, videoconferencing, tele-meetings, instant messaging, and online computer games—to mention just some of the main ones. In recent years virtual reality technologies (VR) have been increasingly used for distributed work and play, making it possible for people physically distributed to interact with each other and with a spatial graphical 3D interface in real time. In order to be able to connect to and interact with the Shared Virtual Environment (SVE) generated by the VR technology, the user must use some kind of VR system; for example a high-end immersive projection technology system (IPT, also known as Cave-type system) or, more commonly, a desktop system. All VR systems provide the user with some kind of output from the SVE, most commonly visual and audio output, and the possibilities to make input into the SVE by using some kind of tool (e.g., a mouse, a pointer, or a joystick) for interaction with the graphical environment and with other users.

When people interact with each other in a SVE, they do not necessarily use the same kind of VR system—for example, desktop computers or Head-Mounted Displays (HMDs). Or again, it could be that members of a group use the same kind of output technology (for example, two networked Caves) but different types of input technology (for example, a 3D joystick at one end and a cyberglove at the other).

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 151–166.

As has been pointed out in previous writings in relation to system differences in distributed group-work [1], users are seldom aware of these differences. The reasons for this are mainly two: (1) since the users only see each other as graphical representations (avatars) in the SVE, they are not aware of the physical surroundings—including technical system setup—of the other group members, and (2) since people have a tendency to overestimate other peoples' similarity to themselves on attitudes, behaviors, and personality traits (the so called "false consensus effect" [2]), they also see themselves as the norm technically; that is, when working together in a distributed group, they believe that the other group members have the same technical capabilities as they do themselves and do not bother to ask about what kind of technology the other group members are using. As has been reported elsewhere [3], the use of different VR systems (different input and output features) in distributed group-work gives the group members unequal possibilities to interact with the SVE and with each other, something which often causes confusion, misunderstandings, and difficulties in collaboration. However, due to the unawareness of the situation that causes these problems of inequality, collaborators seldom discuss this problem or solutions to it. This means that, instead of jointly tackling the collaboration problems that occur, they try, not always successfully from a group-work perspective, to handle the problems individually.

It is highly unlikely that distributed collaborators will be using exactly the same VR systems for group-work in SVEs in the future, for example for reasons of costs (highly immersive systems can be provided for a limited number of participants). Thus, there will continue to be a serious need for managing the collaborative problems that an unequal technical setup causes. One way to manage this is simply to make the collaborators familiar with each others' systems in order to reduce the misunderstandings about what the different systems can and cannot do, and perhaps even to take advantage of the system differences and distributing the labor in a way that makes better use of the two or more VR systems.

This chapter reports on a trial where individuals worked together in pairs on a collaborative problem-solving task in a SVE using very different VR technologies; a desktop VR system on one side and a high-end immersive VR system on the other. Half-way through the task (after approximately 10 minutes) they were asked to switch system with their partner. The hypothesis was that the change of perspective would lead to better possibilities of dealing with issues that are related to distributed group-work and thereby improving the group-work process. It was found that there are several advantages with experiencing different and unequal systems when dealing with a collaborative task of this kind. Partners learn not only about the strengths and limitations of the different systems, but also about collaborating with others and about the implications of using different technologies. The chapter concludes with the implications of this "good inequality" for the design and use of SVEs.

2. Background and Previous Research

Even though many researchers have shown the benefit of face-to-face collaboration in comparison with mediated interaction [4], distributed collaboration is becoming increasingly common. Various technologies have contributed to this development, including tools for text-based computer-supported collaborative work, videoconferencing as well as SVEs [4, 5]. However, mediated collaboration faces both technical and social challenges. Access to technology varies within and between countries [6] as well as within and between groups [7]. Even within organizations it is often the case that different sites have access to different technologies, and practices in using technologies for mediated collaboration can vary between sites (see, for example, [8] for videoconferencing). So when people work together at a distance via computers, videoconferencing systems or in SVEs, they often use systems with different capabilities.

The consequences of this asymmetry can be problematic or they can go unnoticed. Partners may "divide the labor" between themselves, taking on different tasks without being aware of this (for SVEs, see [9]). Or the consequences may be problematical insofar as the differences between systems can lead to inequalities in leadership or in status and thus in how people interact and work together—again, without the collaborators being aware that this inequality has been introduced or shaped by the technology and its different capabilities.

One reason for this effect is the absence of social cues in computer-mediated communication. This effect has been studied at least since Short, Williams and Christie's studies [10], which compared different communication technologies and face-to-face communication in terms of "social presence". A SVE can be considered as a "rich" medium in terms of social presence, compared to for example text chat, since the SVE generates a strong sense of "being there together" (for definitions of presence and co-presence, see [11–13]). How people take advantage of the lack of social cues and of the social cues they have available in SVEs in comparison with face-to-face situations is critical for this kind of mediated collaboration.

Previous research on collaboration in SVEs involving different systems has shown a variety of effects. A study by Slater and colleagues [14] of small group collaboration involving three people where one used a HMD and the other two used ordinary desktop computers, showed that the person who used the HMD was considered the leader, without knowing what kind of system the others were using. Schroeder and colleagues made similar findings for pairs working together, one using an IPT system and the other person using a desktop system [15]. Again, the person using the IPT was considered the leader and as contributing the greater share to the task. When doing the same task face-to-face, desktop-to-desktop, or IPT system networked with another IPT system, no such leadership or unequal contribution tendencies were found. An additional finding was that the subjects "naturally" divided the labor between themselves, with the immersed subject taking a more active role in the spatial aspects of the task and in manipulating objects, and the non-immersed subject taking a more "supervisory" role—again, without being aware of the differences between the two systems [9]. Finally, Axelsson [3] has analyzed the findings from a number of studies of interaction between people using immersive and non-immersive systems, and discussed the different problems that occur when people use asymmetrical technologies.

Findings by Hindmarsh, Fraser, Heath, & Benford [16] showed that problems of working together using networked desktop systems for a spatial task occur because of the restricted field of view and because the collaborators are not aware of what their partners can and cannot see. Heldal, Steed, Spante, Schroeder, Bengtsson, & Partanen [17] by contrast found that this problem on the whole does not apply to collaboration using networked IPT systems. The benefit of using symmetrical technological set-ups to facilitate equal collaboration has thus been shown by a number of studies.

3. Method

In contrast to the studies mentioned in the previous paragraph which used a between subject design, the current study used a within subject design, giving users an experience of both types of technologies. Eighteen subjects arranged into nine pairs participated in the trial. Each pair met in a SVE to solve a Rubik's-cube type puzzle (see figure 7-1) using an immersive and a non-immersive VR system. The trial was limited to 20 minutes and the subjects changed systems half-way through the trial. The subjects were 17 postgraduate students taking a pedagogical course at a technical university and their teacher. There were 4 females and 14 males from various disciplines at the university. The subjects had all met during the course, but they had no previous experience of working together.

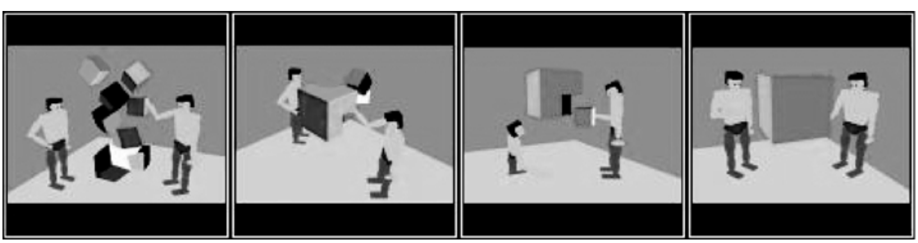


Figure 7-1. Two subjects (visible as human-like avatars) involved in a puzzle-solving task in a SVE using an IPT and a desktop system.

3.1. Technology and Task

The immersive system used was a 3 × 3 × 3 meter TAN VR-CUBE with stereo projection on five walls (no ceiling). The application was run on a Silicon Graphics Onyx2 Infinity Reality with 14 250 MHz R10000 MIPS processors, 2 GB RAM and 3 Infinite Reality2 graphics pipes. The subjects wore CrystalEyes shutter glasses and used a 3D wand for object manipulation. A Polhemus magnetic tracking device tracked the head via the glasses and the hand via the wand. The non-immersive desktop system consisted of a Silicon Graphics O2 with one MIPS R10000 processor and 256 MB RAM and a 19-inch screen display. The dVise 6.0 software was used.

The task was to solve a puzzle involving eight blocks with different colors on different sides and to rearrange the blocks such that each side of the finished cube would display a single color. The colors on the sides of the 8 blocks were red, blue, green, orange, yellow, white, and black.

Using the IPT system, the subjects could move the virtual blocks by putting their virtual hand into a block, pressing the button of the 3D wand and moving the wand in the desired direction. Navigation was purely by moving around physically and pointing with the 3D wand. Using the desktop system, the subjects navigated by pressing the middle mouse button. In order to move a block, the subjects had to first select a block by clicking on it with the left mouse button, then keep the right mouse button pressed and move the mouse in the desired direction. By pressing the right mouse button combined with the shift key, the subject could rotate the virtual block. When selecting a virtual block as described above, the outlines of the block appeared as dotted lines (also visible to the partner). Whereas the avatar in the immersive system (representing the immersed subject) was dynamic and represented the subject's actual movements, the avatar in the desktop system was static (e.g. no gestures) since the subject's physical body was not tracked. We used the Robust Audio Toolkit (RAT) for audio communication between the subjects via headsets with microphones and earphones.

3.2. Procedure and Experimental Design

Before each trial session, the two subjects were given verbal instructions concerning the aim of the task and the various functions of the input- and output devices. The subjects were deliberately not informed about their partner's system. They had 5 minutes to familiarize themselves with the system but were not allowed to start communicating with the partner. The total time for the task was 20 minutes and halfway through the subjects changed system. Post-trial interviews with the subjects took between 5 and 15 minutes and focus

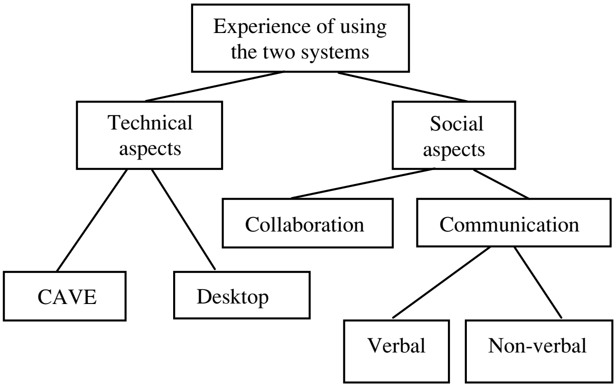


Figure 7-2. The different technical and social aspects that the subjects reflected upon during the post-trial interviews.

group discussions involving 4–6 subjects were held which took between 45 and 60 minutes.

3.3. Data Collection and Analysis

The trial sessions were audio- and videotaped and the post-trial interviews and discussions were audiotaped. The analysis presented here is based only on the transcribed post-trial interviews. In order to interpret the interviews we used content analysis as described by Altheide [18]. We were particularly interested in how the subjects experienced the change of VR systems; how that influenced their sense of collaboration and their experience of the two VR systems. During the post-trial interview we asked the subjects to reflect on the task, the collaboration and the change of VR systems with specific reference to the technical aspects on the one hand and the social aspects on the other. These aspects were operationalized as described in the above scheme (figure 7-2).

4. Results

In the following section we present, first, the data where the subjects reflect on their experiences in relation to technical aspects, which will be followed by reflections in relation to social aspects. Quotes from subjects are coded as follows: "I1" informs us that the quotation comes from the first subject who started the trial using the IPT system, and "D1" informs us that the quotation comes from the first subject who started the trial using the desktop system—and so forth.

4.1. Experiences of Using Different Systems: Technical Aspects

One question concerned how the subjects experienced the use of the two different systems during the trial. Typically, the subjects remarked on the different technical functionalities. In general, they experienced the IPT system as more intuitive to use and the manipulation of the virtual blocks as easy. The desktop system was experienced as more difficult to use because of the need for button pressing to manipulate the objects. Typically subjects commented on similarities and differences:

> I started at the workstation [referring to the desktop system] . . . When I came down here [to the room where the IPT system was located] it was more intuitive in a way what I should do. I saw where I was and I had only one type of control apart from my own movements. The only thing I needed to do was to grasp and release. There I could twist and do the turns. At the workstation I had to do it with control –alt-shift or control-shift and a mouse click so it was harder for me to do the task there. I thought I noticed that for [name of the partner] as well. When she came up there [to use the desktop system] she didn't really know how to move either. (D1)

> One was more handicapped there [on the desktop system]. One cannot do everything. The function one could have is to stand aside and look since one does not have the same functionalities. Here [IPT system] one is much smoother. (D7)

> I think you get a better view from here, from the computer [desktop system], but you cannot handle the things from here very well. I don't know if there is a problem with a cable or with the special joystick. But I think when you are there on the stage [IPT system] it is easier to move and manipulate objects. (I11)

Regardless of which system the subjects started with, most subjects shared the same view, that the IPT system was more intuitive and easy to use. However, in contrast to the majority opinion, one subject said:

> I think that the computer [desktop system] was easier, but one is used to computers. At the same time, here [IPT system], some things were easier to do here such as walking around faster and seeing the colors. That one could do faster down here, but the mere use of it, the computer was easier, but that is surely a habit issue, one is used to control [with the keyboard]. Here, it was more of trying to see the hand and, well, click on the right spot. I experienced that as harder. (D6)

The majority of the subjects could intuitively and easily use the IPT system even though they had no previous experience of the technology. A few subjects expressed minor difficulties grasping objects with their virtual hand using the wand, but the general experience was that it was easy to manipulate the virtual objects using the IPT system in comparison with the desktop system.

Another experience in relation to the technology was the difference in immersion. Even though the majority of the subjects said that the IPT system was more intuitive to use in comparison with the desktop system for actions such as rotating the virtual blocks, some subjects experienced the IPT system as being too immersive:

> My feeling is that you can manage better the system through the computer [desktop system]. That cube [the IPT system] is . . . it . . . causes a lot of difficulties. You feel surrounded by cubes. And you sort of . . . you can grab one of them. But for me from this monitor [desktop system] I can see everything and probably I can manage my tools. I think so. (I13)

> There they [the virtual blocks] are all around you so it is hard to get a real overview. (I16)

Even subjects that did not mention the feeling of being "surrounded by cubes" said that the desktop provided a better overview of the puzzle. The following two quotes illustrate the general view of the subjects concerning the advantage of using the desktop in comparison with the IPT system:

> I think it was easier in front of the desktop using the mouse and keyboard to have an overview and perhaps help out a bit and check it out and think a bit. (D9)

> I think you get a better view from here from the computer but you cannot handle the things from here [desktop system] very well. (I11)

It was not the case that the subjects were completely in favor of one system compared to the other. The subjects appreciated the two different systems in different ways. The different technical functionalities of the IPT and desktop systems were useful for different purposes in solving the puzzle together. The IPT system was seen as useful for manipulating objects and the desktop system for obtaining a clearer overview of the puzzle. At the same time, the different technical functionalities also caused different types of difficulties: the desktop system was associated with problems in moving and manipulating the virtual blocks due to the need for button pressing; using the IPT system, on the other hand, was experienced as easy to use in relation to navigation and object manipulation but difficult in relation to getting an overview of the puzzle. This yields the following picture (table 7-1).

Table 7-1. Difficulties associated with the two different VR systems.

	IPT	**Desktop**
Manipulation	Easy	Difficult
Overview	Difficult	Easy

4.2. Experiences of Using Different Systems: Social Aspects

The subjects were also asked a number of questions concerning their experience of using the two systems in relation to different social aspects. We will present, first, their views on collaboration and then their views on communication.

The majority of the subjects experienced the trial as a highly collaborative situation and expressed themselves in a positive way about working together. Regardless of whether they solved the puzzle or not, working with a partner on a task like this was seen as a good thing. However, although 11 out of 18 subjects had a positive experience of the collaboration, three subjects felt that they could have very well solved the task without a partner. Only two subjects reported that their collaboration was not working.

Most subjects thought that their collaboration improved after they had changed systems, and thought that they could use this knowledge about how the different systems worked to improve their collaboration:

> I thought it [the collaboration] worked well. I thought it worked very well when one knew, when one had tried out each other's tools. In the first instance one did not know what kind of capabilities the other had. I noticed that he could move around much easier but I did not know if that was because of him being better to manage the terminal [the desktop system] or what it was. I didn't know that he was down here [IPT system], that he had a tool like this. It became much easier after, when one knew, then we could divide the work better between us. (D7)

> I think the collaboration with my partner was really fruitful and especially because we had two different views. From the computer I can see above better, and he can handle better the cubes, so I think the collaboration is necessary to solve the task faster. (I11)

> You know, we started with no strategy at all. That was actually bad because we didn't see what next. But during this final stage we understood each other better and that was a relief. (I13)

Even subjects who thought that they did not really make active use of their knowledge about the different technologies believed that changing systems had improved their collaboration.

> I think it is better than working on my own because obviously there are some tasks that are more difficult to do from the computer [the desktop system] but it is easier for me to move around and he can turn around more easily different sides of the block. So I think that it is good to compensate but all we need is to have a better plan if we know the task earlier or in the middle. We should have some time in the beginning to just [talk about]—how we should do the task . . . It's easier to control when you sit here in front of your computer [desktop system] of course. So maybe it is good to have a strategy and then do some work from the computer first and then go downstairs [IPT system] to make the detail. (I10)

In relation to the experience of collaboration, most subjects thought that collaboration was useful and that it was improved by the fact that they had changed systems. The change of system led to an increased understanding of each other's perspectives and capabilities. This understanding enabled them to divide the labor between themselves based on the capabilities of the technology—such that the IPT person took a more active role in manipulating and moving the virtual blocks, and the desktop person had an overview and took a more "supervisory" role.

As for communication, subjects regarded verbal communication via the audio channel as crucial, but they also considered it to be as important to see their partner's avatar movements and actions. Typically they commented on the way their partner moved around in the environment. In particular, those who started off using the desktop system found it remarkable that their partner moved around so easily and smoothly in comparison with themselves:

> In some way I realized that he had a different tool. One understands that at once when one can see how smoothly he can move. One understood that quite quickly. Then it took a while before we talked about what kind of tool the other one had, but that became obvious when we changed. (D7)

This quote also illustrates how some subjects attributed the differences in movements to the technology without any knowledge about the differences between the systems. Some subjects, however, associated this to their partners' skills:

> I thought it was a superman I had met that could do exactly as he pleased with his keyboard. (D3)

The ability to refer to objects by pointing at or moving them back and forth facilitated communication concerning which block was being handled at the moment. However, as one subject noticed, the ability to refer to objects was different in the two systems:

> He was there in a way. It was really hard to express when one was upstairs [desktop system]. Then one had to grab a cube and say—"I am over here". But for him [in the IPT system] he could say "here I am", or, "I am going here". In some way he was there but I was not. (D8)

This quote demonstrates the subtle mix of verbal and non-verbal communication in a SVE which is further complicated by an asymmetrical setup. The possibility of referring to objects depended on the system: when using the desktop system the subject had to select a virtual block to indicate to the partner what object s/he was talking about. When using the IPT system the subject could refer to a block by simply pointing at it. Movement could be conveyed by means of the dynamic avatar, which showed the physical movements of the user, something which was not possible on the desktop system. Not only was action more intuitive in the IPT system, but language use was also more intuitive in the sense that "here" and "there" could be conveyed through the interface in the same sense as in the physical world. Subjects realized that knowledge of the two different systems also improved the way they communicated:

> [Changing systems halfway through] was fun. One could see these different possibilities. But that also meant, given that one had tried both systems, that one could more easily communicate with the systems and [also] communicate better with each other. (D6)

Changing systems was thus important for a better understanding of each other's possibilities and constraints, which helped subjects to agree on who should do what.

Finally, the experience of collaboration and communication is also reflected in the subjects' comments related to "being there together" or co-presence:

> Without voice communication it would have been difficult, so it was crucial. (D8)

However, this same subject also felt that he sometimes forgot his partner while busy handling the objects using the IPT system:

> But also, since I did not see him, or rather he was over there so to speak, he was not close to the cubes. Then it was very easy to forget [him] . . . not until I was working alone I thought: oops, now I'm doing too much! (D8)

This quote also illustrates that it is the avatars that subjects respond to (not to the "real" partner), and in this case, when the subject was busy with the virtual blocks, he felt that he lost awareness of his partner.

5. Discussion

Previous studies have highlighted the disadvantages of asymmetrical setups. But as shown in this chapter, if users can be made aware of the differences entailed by different systems, users can obtain a better understanding of the possibilities and constraints of the systems, and collaboration can in this way be enhanced. The subjects in this study clearly thought that changing systems was positive and that they could collaborate and communicate better after the change. They also thought that they made use of the different capabilities of the technology and that this improved their strategy for solving the task. They recognized that the IPT system provided them with better possibilities for object manipulation and the desktop system provided them with a better overview of the puzzle.

Interestingly, subjects recognized the benefits as well as the drawbacks of each system. These were insights that they in some cases were able to implement immediately, during the ongoing trial, after having changed systems, and in other cases these new insights occurred to them only after they had completed the session, insights that they should be able to make use of in similar future collaborative situations.

Research that points to the importance of face-to-face meetings for distributed work typically highlights the importance of shared knowledge about the situational context in which the collaboration takes place—such as knowledge regarding what it looks like at the other site and "getting a feel for" the work culture that may vary from site to site [19, 20]. As in many previous studies, the present study underscores that shared knowledge enhances collaboration. However, in addition to the shared knowledge about the context in which the collaboration takes place, it is also, as the current study shows, important to facilitate knowledge-sharing regarding the capabilities of the technical systems, since this knowledge is hidden to the users.

Before we discuss the implications of these findings further, it will be useful to recall some of the disadvantages of unequal systems. One of the main disadvantages is that the collaborators may not be aware of the unequal technical capabilities that they and their partners have, which may lead to misunderstandings since they will communicate and collaborate with each other on the basis of their own technical system. Although users will experience co-presence in the SVE, they will still not be able to see or do things in the same way. Hence, the users may adopt a poor strategy for solving the task since they believe that they can contribute fairly and equally, which is not the case. In addition, people who collaborate using unequal technical systems can obtain incorrect impressions of their partner since the partner's way of handling the technology is interpreted as an indication of personal skills and character rather than of the technology the partner is using.

However, overcoming the disadvantages of different technologies by "trading places" will often not be possible. The point of distributed work is usually that partners should work together at a distance without meeting face-to-face, in order, for example, to save time or travel costs. However, when the technical set-up is so different as when an IPT system is linked to a desktop system, there may be a trade-off between this disadvantage and effective collaboration. A further insight from this study is that one should consider whether it is worthwhile to invest in and use a costly immersive system for an object manipulation task when the partner on the opposite site has a non-immersive system. The drawbacks with this set-up might actually be greater than the benefits. However, although previous research has shown the benefit of symmetrical set-ups for object manipulation [14, 15] collaboration *can* and *must* work with asymmetrical set-ups of VR systems. Some of the disadvantages of asymmetrical set-ups can be overcome with enhanced knowledge about the possibilities and constraints that the systems provide. If people know about the system differences, they can make use of them in their collaboration.

It is also worth mentioning that there may be advantages for two or more participants to have different technologies and actually take on different roles—for example, when people need to perform different complementary tasks. One result of using unequal systems is that, as mentioned earlier, even when collaborators are unaware of the type of system that their partner is using, they may be able to divide the labor between themselves. In the present study, when the participants found out about the reason for their unequal participation, they said that they could make use of this knowledge to figure out a better strategy for carrying out the task. In other words, creating a "common ground" in a situation of missing social cues allowed them to collaborate better [21]. It can be seen that VR technology is not only a tool *for* social interaction, but also an important feature *in* social interaction [22].

6. Conclusion and Future Work

The finding that distributed group-work via VR technology can be enhanced by increased awareness of the technological capabilities will have obvious relevance for the design of VR systems and their uses. How then can VR systems provide knowledge about different capabilities? Can knowledge be built into the systems, or should task sessions be structured so as to allow for "putting on the other person's virtual shoes"? One suggestion that one is tempted to make in response to this study is that the differences between the VR systems—the technical capabilities—should be made obvious to users, and that this can easily be done with VR technology. For example, avatars could have labels that specify what type of system and input/output devices the users possess.

Note, however, that this solution would also have drawbacks: for example, such labels might create a cognitive overload on the users' part (how much information can the user "take in"?). Moreover, the whole point of VEs is that they are supposed to be natural interfaces, without the need to bring extraneous information into the environment—information that the users then need to maintain awareness of. It may also be that by focusing on figuring out what capabilities their partners have, collaborators lose the ability to "naturally" divide the labor between themselves and thus add to rather than reduce the time they need in order to carry out the task. As in everyday work life, we collaborate with people with various capabilities that are not made explicit, and this seldom causes collaboration difficulties. On the contrary, different capabilities often increase collaboration between people [23].

In the case of SVEs, however, the present study showed the benefit of being made aware of the various capabilities of technology by means of trading places.

Another suggestion for improving collaboration in SVEs might therefore be that users develop a verbal protocol whereby they talk about what tools they have access to instead of displaying this information. Developing a verbal protocol could be that after: "Hi how are you?" it might become a convention to ask: "What does the technology look like at your end and how does it work?". Mark [24] has demonstrated the benefits of explicit shared conventions in computer-supported collaborative work whereby users save both time and effort by using these.

One limitation of our trial was that it was short and that the subjects had only one opportunity to solve the task together. It may be that collaborators could easily adapt to the different capabilities or to the absence of social cues which could make them aware of these differences, over the course of time [25] (for time effects in text-based communication and collaboration, see also [26]). It would be interesting in future research to test whether such adaptation takes place, as well as whether longer sessions with different systems could mitigate the need for "trading places"—or if doing so could be even more valuable during longer sessions. It would also be interesting to examine whether simply communicating the different capabilities verbally or by means of the partners demonstrating them to each other remotely could be just as effective as experiencing the different systems.

To sum up: putting yourself into the other person's virtual shoes can enhance the interaction and the strategy in a collaborative task, as well as providing people with valuable insights into the use of VR systems. In other words, knowing about the different capabilities of technologies can enhance collaboration, thus creating "the good inequality". As the discussion has shown, however, such a setup for "trading places" may not always be possible or desirable to implement. Future research will show under what circumstances it can be useful to change VR systems—for designers of VR systems, and to participants involved in distributed group-work beyond the trial setting of this study.

Acknowledgements

This research was part of the project "Living in Virtual Worlds: Implications of the Uses of Virtual Reality in Long-Term Settings" supported by VINNOVA, and the Chalmers C-SELT project "Engendering Good Learning".

References

1. Axelsson, A.S. (2004). *Framing Social Interaction in Shared Virtual Environments: The Influence of Technical and Social Factors*. Doctoral Dissertation. Chalmers University of Technology.

2. Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attributional processes. *Journal of Experimental Social Psychology, 13*: 279–301.
3. Axelsson, A.S. (2002). The digital divide—Status differences in virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*, London: Springer, pp. 188–204.
4. Hinds, P. & Kiesler, S. (Eds.) (2002). *Distributed Work*. Cambridge Massachusetts: The MIT Press.
5. Churchill, E., Snowdon, D., & Munro, A. (Eds.) (2001). *Collaborative Virtual Environments: Digital Spaces and Places for Interaction*. London: Springer.
6. DiMaggio, P., Hargiatti, E., Neuman, R., & Robinson, J.P. (2001). Social implications of the Internet. *Annual Review of Sociology*, (27): 307–336.
7. Haythornthwaite, C. & Wellman, B. (Eds.) (2002). *The Internet in Everyday Life*. Oxford: Blackwell Publishing.
8. Sonnenwald, D.H., Solomon, P., Hara, N., Bollinger, R., & Cox, T. (2002). Collaboration in the large: Using video conferencing to facilitate large group interaction. In A. Gunasekaran & O. Khalil (Eds.), *Knowledge and Information Technology in 21st Century Organizations: Humans and Social Perspectives*. Hershey, PA: Idea Publishing Co, pp. 155–136.
9. Axelsson, A.S., Abelin, Å., Heldal, I., Nilsson, A., Schroeder, R., & Wideström, J. (2001). Cubes in the cube: A comparison of a puzzle-solving task in a virtual and real environment. *CyberPsychology & Behaviour, 4*(2): 279–286.
10. Short, J., Williams, E., & Christie, B. (1976). *The Social Psychology of Telecommunications*. London: John Wiley & Sons.
11. Scheumie, M.J., van der Straaten, P., Krijn, M, & van der Mast, C. (2001). Research on presence in virtual reality: A survey. *CyberPsychology & Behaviour, 4*(2): 183–201.
12. Schroeder, R. (Ed.) (2002). *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer.
13. Biocca, F., Harms, C., & Burgoon, J.K. (2003). Towards a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and Virtual Environments*, *12*(5): 456–480.
14. Slater, M., Sadagic, A., Usoh, M., & Schroeder, R. (2000). Small group behaviour in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments, 9*(1): 37–51.
15. Schroeder, R., Steed, A., Axelsson, A.S., Heldal, I., Abelin, Å., Wideström, J., Nilsson, A., & Slater, M. (2001). Collaborating in networked immersive spaces: As good as being there together? *Computers & Graphics, 25*(5): 781–788.
16. Hindmarsh, J., Fraser, M., Heath, C., & Benford, S. (1998). Fragmented interaction: Establishing mutual orientation in virtual environments. *Proceedings of CSCW'98*, pp. 217–226.
17. Heldal, I., Steed, A., Spante, M., Schroeder, R., Bengtsson, S., & Partanen, M. (in press). Successes and failures in co-present situations. *Presence: Teleoperators and Virtual Environments.*
18. Altheide, D. (1996). *Qualitative Media Analysis*. Thousand Oaks: Sage Publications.
19. Kraut, R., Fussell, S., Brennan, S., & Siegel, J. (2002). Understanding effects on proximity on collaboration: Implications for technologies to support remote collaboration. In P. Hinds, & S. Kiesler (Eds.), *Distributed Work*. Cambridge Massachusetts: The MIT Press, pp. 137–162.
20. Nardi, B. & Whittaker, S. (2002). The place of face-to-face communication in distributed work. In P. Hinds & S. Kiesler (Eds.) *Distributed Work*. Cambridge Massachusetts: The MIT Press, pp. 83–110.

21. Axelsson, A.S., Abelin, Å., & Schroeder, R. (2003). Communication in virtual environments: Establishing common ground for a collaborative spatial task. *The 6th International Workshop on Presence*. Aalborg, Denmark, October 6–8.
22. Spante, M. (2004). *Shared Virtual Environments: Technology, Social Interaction, Adaptation and Time*. Licentiate Thesis. Chalmers University of Technology.
23. Albrecht, M.H. (2001). *Managing Diversity in the Workplace*. Oxford: Blackwell Publishers.
24. Mark, G. (2002). Conventions for coordinating electronic distributed work: A longitudinal study of groupware use. In P. Hinds, & S. Kiesler (Eds.), *Distributed Work*. Cambridge Massachusetts: The MIT Press, pp. 259–282.
25. Nilsson, A., Heldal, I., Schroeder, R., & Axelsson, A.S. (2002). The long-term uses of shared virtual environments. An exploratory study. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 112–126.
26. Walther, J.B. (2002). Time effects in computer-mediated groups: Past, present, and future. In P. Hinds, & S. Kiesler (Eds.), *Distributed Work*. Cambridge Massachusetts: The MIT Press, pp. 235–257.

Chapter 8

CONSEQUENCES OF PLAYING VIOLENT VIDEO GAMES IN IMMERSIVE VIRTUAL ENVIRONMENTS

Susan Persky and Jim Blascovich

1. Immersive Virtual Environments versus Desktop Platforms

One way to gauge the implications that immersive virtual environments (IVEs) hold for research, work, and play is to compare participant experiences using IVE platforms with participant experiences using non-immersive platforms. Hence, virtual environment (VE) researchers have increasingly investigated platform type as an important factor influencing media impact. Experimental researchers can and have evaluated platform effects on participants' interactions and experiences experimentally by varying platform type or configurations, such as IVE systems and non-immersive systems (e.g., desktop computers), while holding content constant across platforms. The outcomes of such experimental studies can help investigators determine the added value of IVE technologies in various domains. Such outcomes can be positive, contributing an important research tool for behavioral, biomedical, and social scientists; increasing productivity and efficiency in the workplace, and providing desirable leisure experiences. However, such outcomes might also be negative, causing researchers to alert users to potential hazards of various platform types and identifying application-platform combinations in which caution should be exercised and care taken in development and distribution.

Immersive virtual environment technology is increasingly used to simulate relevant environments for government and military training, for performance assessment of various activities such as driving an automobile or piloting an airplane, and for research in the behavioral and computer sciences. Research supports the value of IVE use in such arenas. Empirical studies have shown that IVEs are more effective than non-immersive platforms for various types of training and learning purposes. For example, participants who were trained

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 167–186.

on a simple search and navigation task in a simulated environment using an IVE platform that incorporated a head mounted display (HMD) and tracking system later performed the same task in a non-simulated environment more quickly and consistently than participants who were trained using either of two non-immersive platforms [1]. In a more complex search task, participants who were trained using a digital IVE training simulation later performed better in an analogous physical simulation than did participants who were initially trained using a desktop computer version of the digital training simulation or who received no prior simulation training [2]. In addition to transfer of training skills, research shows improved performance via IVEs compared to desktop platforms as a medium for spatial and navigation training and tasks. For example, participants asked to learn about configuration and movement of 3D chess pieces demonstrated better performance if the learning phase took place via an HMD-based IVE rather than via a desktop monitor screen [3].

Use of IVE systems has also been shown to produce benefits for collaborative task performance in terms of general efficiency, productivity and leadership. In a small study in which one of three participants in a collaborative group was immersed in a shared IVE using an HMD and in which the other two participants shared the environment via desktop monitors, the immersed participant consistently emerged as a leader during the criterion task [4]. Other researchers have also reported a positive correlation between use of IVE technology and leadership in a molecular visualization task [5]. In terms of task performance, in a study in which participants collaborated in a virtual building task in either a Cave-type IVE or desktop platform, researchers reported a performance increase for users of the Cave-type interface, particularly for the more difficult tasks [6]. In addition, performance has been shown to suffer when one of two cooperating participants uses a desktop rather than an IVE. Furthermore, studies comparing collaborative performance in an IVE to performance in a physical environment have found little or no differences in performance between the two [7] suggesting that IVE collaborative groups are able to function at a high level.

Immersive virtual environment use may not always result in the highest levels of task performance, however. In one study, participants were assigned to perform tasks in a single-user map-based battlefield simulation using four types of platforms including a desktop computer and a Cave-type IVE system. In this study, researchers reported the best task performance for users of the desktop computer system [8]. However, the battlefield simulation used in this type of simulation provides little advantage or immersion for users in IVEs compared to desktop platforms. Indeed, other work has suggested that presence in an IVE is a major factor in the advantages conferred by IVE use. Tasks involving environments that require unrealistic viewpoints and actions might be more appropriate for and therefore benefit more from a non-immersive interface. Clearly, then, IVE use will not confer an advantage in all situations, however, for the tasks that are enhanced by IVE use, it is important to consider why this particular type of platform confers such an advantage.

Simulations in IVEs can be quite compelling and can be experienced as more realistic than on other platforms. When IVEs include advanced tracking and orientation systems, users can interact seamlessly with media content, ignore the physical environment in which they are located, and interact within the environment in a naturalistic way. In many contexts in which IVEs are used, enhanced realism and opportunity for naturalistic interaction within these environments can prove beneficial to users. In other contexts, however, enhanced realism and naturalistic interaction can prove problematic.

One of the most obvious arenas in which IVE technology may prove problematic is in situations in which it involves violent content. The increased immersion produced via IVE platforms is particularly relevant to the study of violent video games in which users act aggressively and perform the violent actions necessary for game play. Before discussing the particulars of how immersion, presence, and violent content can interact to produce heightened anti-social outcomes, a review of the video game violence literature will be helpful.

2. Violent Video Game Effects Research

Compared with the study of other types of media violence, research on violent video game effects is relatively young with a small but growing research literature. A spate of narrative review papers was published between 1998 and 2001, advancing varying conclusions as to the richness of the video game violence literature [9–14]. At the time of their publication, most suggested that it was too soon to draw broad conclusions about the potentially problematic effects of violent video games, and cited mixed results in the available body of literature. Also at issue were common methodology and demand problems, and a general dearth of experimental publications.

Two meta-analyses were published during the same period, each reporting a small-to medium effect size of violent video game play on various aggression-related outcomes [15, 16]. One of these analyses found that year in which a study was performed was the strongest positive indicator of overall effect size and suggested that the technological advancement of games over time was responsible in large part for mixed experimental findings [16]. As game technologies became more advanced over time, they led to more realistic, more intense, and ultimately more violent games that, in turn, appear to have led to increases in aggressive outcomes.

More recent experimental studies of violent video games have resolved many of the previous methodological issues. These recent studies also tended to use contemporary games replete with high levels of realistic violence. With few exceptions [17], these studies contributed to increasing evidence of detrimental outcomes resulting from violent video game play including aggressive behavior [18–20], aggressive cognitions [18], increased hostility [21, 22], and implicit associations between the self and aggressive traits [23]. The consistency of these latest studies support the proposition that the mixed results in the early video game violence literature may stem, at least in part, from the inconsistent levels of realism in the violent games used in experiments.

Clearly, if technological advancement and increasing realism in violent video games underlies an increase in aggression resulting from game play, these technological developments are an important focus for research. Immersive virtual environment technology is one such development that promises to contribute to increasing realism in violent video games and indeed is already beginning to do so.

3. Video Game Playing in IVEs

Immersive virtual video games are not new. Such games have been popular since the 1990s when they first appeared at arcade gaming centers in shopping malls, amusement parks, and entertainment complexes. Immersive virtual video games generally took one of two forms, a vehicle simulation pod in which a user would sit and watch a screen that acted as his or her window to the virtual world, or a head mounted display system with various forms of tracking. IVE games varied in content, but shooting or battle-themed games were quite prevalent. Many of the original VE game centers have since closed, but newer centers, such as Disney Quest in the United States, have opened to take their place.

Still, IVE games are not as widespread as might be expected given the levels of enjoyment and repeated use reported by players [24]. Perhaps this is due in part to a disconnect between what inexperienced users would expect of virtual reality given its representation in popular media and the actual sophistication of the technology in its current state. Technology for IVE simulations continues to advance, however, due not only to the work of researchers and game manufacturers, but due also to the development of training simulators for military purposes. Furthermore, arcade game devices have begun to include aspects of immersive technology such as motion tracking that does not create full immersion but comes close. These developments promise not only more advanced IVE games in public spaces and centers, but also a move toward more immersive gaming experiences in the home [25] where video games are played most frequently.

As we move closer to widespread diffusion of IVE games, it is essential to consider the ways in which this IVE technology will change gaming experiences for the better and also, importantly, for the worse. Examination of how platforms, violent video games, and presence might interact will help us understand what a move toward immersive technologies for gaming might mean for post-play outcomes.

4. Immersive Virtual Environments, Violent Video Games, Presence, and Aggression

4.1. Implications of Virtual Violent Video Games for Presence

We define presence as a psychological state in which one perceives one's self as existing in an environment. Presence, then, exists as a phenomenological state independent of any particular hardware system. Researchers have identified many antecedents to, and consequences of, the experience of presence within IVEs both theoretically and empirically [26]. By investigating the ways in which immersive virtual violent video games influence antecedents to presence, we can arrive at a clearer understanding of the mutually reinforcing nature of this relationship.

Media characteristics in a given context are influenced by their content, in this case violence. Visual display characteristics such as image quality, motion, color, and dimensionality can all enhance presence. These characteristics are often quite rich and realistic in modern video games, especially those with violent content. Such games are typically designed to visually represent high levels of violence realistically with high image quality. Auditory display characteristics operate similarly. Sound quality in games continues to improve, as home stereo systems continue to grow and designers create more graphic sounds designed to accompany and amplify the violent visual images that users see. Enhancement of the stimuli (e.g., haptic, olfactory) sent to other senses can also increase presence though their implementation is not widespread. That many immersive virtual video games allow user movements to directly control character movements engages the kinesthetic and orienting systems leading to an increase in both breadth and depth of stimulus quality. Thus, in the realm of media characteristics, the way that contemporary violent video games are designed serves to increase users' experiences of presence.

User characteristics that can increase presence are particularly relevant in the context of immersive violent video games. The experience of presence should increase when the user experiences meaningful content and motivation to engage in that content. So, if a user is prone to violent or aggressive cognitions and is well suited to process such information, as users who voluntarily and recreationally choose to play violent video games may well be, then the user should be more likely to experience a heightened sense of presence when playing such games. Finally, as users gain practice with IVEs, feelings of presence may increase [27]. Video games are popular among adolescents and young adults; perhaps even addictive in certain cases [28, 29]. Video game users are likely to engage in high frequency play because of the enjoyable and/or addictive nature of the games. As these games are increasingly played using an IVE platform, heightened experiences of presence should result.

Because of media characteristics such as visual and auditory display, and user characteristics such as dispositional tendencies for aggressive cognitions and behavior, using IVEs to play violent video games should increase experiences of presence over other types of content. While violent video game content is expected to lead to heightened aggression in IVEs, a user's level of presence in violent games is also expected to affect the relationship between immersive violent video games and resulting aggression.

4.2. Implications of Presence for Playing Virtual Violent Video Games

Presence is central to the prediction of differences in the consequences of playing violent video games using IVEs versus traditional platforms (e.g. desktop computers, arcade units). On a conceptual level, the better a situation is simulated, the less likely individuals will activate a media schema to interpret the situation [30]. Media schemata signal that what one is experiencing is artificial and that responses that would be appropriate in natural situations may not be appropriate. Hence, more realistic simulations should lead to greater presence increasing the likelihood that individuals will experience a simulation situation naturally and thereby be less likely to activate media schemata to control their behaviors.

Other factors have previously been identified as affecting aggression and presence in the violent game playing literature. Early writings and research proposed that video games would lead to increased aggression over other less participatory or interactive media in part because video game players generally take an aggressor's viewpoint [10]. Media violence research supports the claim that identification with the aggressor in a media context increases later aggression against another individual [31]. Increase in the potential for individuals to identify with the aggressor in video games is apparent in two trends. First, modern video games often adapt a viewpoint in which the player's view is identical to an aggressors' view. The player sees through the eyes of the aggressor, experiencing virtually what the aggressor experiences when committing violent, hostile acts. A second development adding to increased likelihood of identification with the aggressor is the ability of players to choose their own (i.e., the aggressor's) representation from a pool of characters who differ in race, sex, build, and other individualizing characteristics. Players therefore have the opportunity to construct a character and an identity.

Virtual environments can further increase identification with the aggressor vis-à-vis presence. When using a virtual platform, not only is the player's view an aggressor's view, the player's body motions are the aggressor's body motions.

Identification with the aggressor and presence are clearly connected as the more an individual experiences presence in the game environment, the more he or she should feel connected to the aggressor's actions.

Other work suggests that aggression should increase as the realism of violent acts increases. The most important type of realism is behavioral realism capturing violent movements, sounds, and even facial expressions. Though, as evidenced by studies of early video games, photographic realism is not a necessary condition for linking violent game play and subsequent aggression, links between increased realism and increased aggression have been found (see below). With new technological advances not only has behavioral realism increased but so has photographic realism likely adding to players' experiences of presence. Game development in the 1990s (e.g. Mortal Kombat) brought more photographically realistic depictions of decapitation and other gruesome actions. Contemporary games bring more photographically realistic blood, screams, wounds, and other such features of violent encounters.

There is some evidence that increases in realism engender increases in violent behavior [32]. Also as mentioned above, more recent studies of violent video games and aggression demonstrate larger effect sizes than earlier studies [16]. Increases in realism, both behavioral and photographic, are likely responsible in large part for increased findings of post-game aggression among video game players. In the case of IVEs, again it follows that heightened realism will play an important part in any increased aggression effects. Environments, opponents, and objects as seen in IVEs are more realistic behaving and looking. In IVEs, players experience objects and characters as three-dimensional. Because presence in a violent game will in turn make experiences in the game environment more realistic, those experiences should lead to increases in aggression.

Tamborini et al. [33] proposed that the strength of the impact of violent content on aggression-related cognition and behavior should be determined by the level of presence experienced during game play. More specifically, Tamborini et al. suggested that a game's impact rides on its ability to enhance involvement and immersion. Involvement and immersion are two critical features of presence [27] where involvement stems from focus on and attention to the environment and immersion results from an environment's ability to isolate and interactively engage the user.

Tamborini et al. suggested that because video games with high-resolution graphics can engage multiple sensory systems, they have increased vividness over other media forms. Because users can almost instantaneously influence the form and content of the game environment, video games are also more interactive than other media. These qualities, vividness and interactivity, share a positive relationship with involvement and immersion such that increases in the former lead to increases in the latter. Therefore, video games in general are a medium that should lead to heightened involvement and immersion, and the addition of IVE technology should increase that involvement and immersion to an even higher level. Tamborini et al. suggested that one of the mechanisms by which immersion in a multi sensory environment would lead to increased aggression is that it should facilitate the learning of more complete aggressive scripts by increasing available and salient cues. The cues in response to which actions are performed should also seem more realistic and should be more easily recognized in the future. According to a cognitive neoassociationist theory of media violence effects, recognizing cues to violence that were present in violent media can underlie violent behavior outside of the game context by priming related cognitions and increasing their accessibility.

All of the routes from immersive video games to increased presence and, in turn, to increased aggression provide strong reason to believe that playing violent video games in IVEs will lead to heightened aggressive responses. Though one study was previously conducted by Tamborini and his colleagues to investigate this possibility [33], the results were mixed, failing to replicate some of the most well established findings in both the presence and the media violence literature. Reasons for the inconsistencies were likely due to characteristics of the VE platform used because the interface was found to be non-intuitive and difficult to master. Convinced of our own and Tamborini et al.'s reasoning, we sought to eliminate this methodological shortcoming. We therefore undertook a series of studies aimed at answering the basic empirical question, does playing violent video games in IVEs increase aggression more than playing these games on traditional platforms?

5. Effects of Immersive Virtual Violent Video Games

5.1. Effects of IVE Platforms versus Traditional Platforms

We conducted two studies [reported in detail in 34] both of which shared the same basic procedure. Participants were asked to play a specific video game either on a desktop computer platform, or using an IVE platform. On the desktop platform, participants viewed the game on a 17″ monitor; they played in a standing position using a hand-held arcade gun controller both to move in the game and to aim and shoot at opponents. On the IVE platform, participants viewed the game in a Virtual Research V8 stereoscopic HMD; they moved in the game using natural body movements while their location and head orientation were tracked. These participants also used the hand-held arcade gun controller, but here they only used it to aim and shoot at opponents.

We developed a gun fight-themed violent video game that could be played on either an IVE or desktop platform. The players' experience was designed to be as identical as possible between the desktop and IVE platforms with

Figure 8-1. Screenshot of the violent gun fight-themed video game created for our experimental studies.

the obvious exception of the differences inherent in the two platforms (e.g., wearing an HMD or not). The game was designed to be simple and easy to play to eliminate issues of differences in playing interface. In the game, players simply shot at two opponents located at the opposite end of the virtual room and were able to hide behind any of three virtual walls to dodge opponents' bullets (see figure 8-1).

To attempt to answer the question of whether playing violent video games in IVEs would lead to increases in aggression compared to the desktop platform, we recorded several types of dependent measures. We included a self-report measure of aggressive feelings composed of a multi-item scale as well as a multi-item presence scale and items assessing playing experience. In addition to our self-report measures, we included two measures to assess behavioral aggression. The first measure was the proportion of head hits to opponents because shooting at the head is arguably more violent than shooting elsewhere on opponents' bodies.

We also, however, wanted to use a more established behavioral measure, one that measured aggression outside of the game context. For this purpose we included a computer-based competitive game. This measure was essentially a reaction time contest with a partner. Participants were told that it was a competition to see who could respond faster to a stimulus. The loser on any given trial was punished by a blast of white noise, the level of which was set by his or her opponent. Participants played against the computer, there actually was no opponent, and noise levels and win/loss trials were randomized. The overall measure here was how high participants set the aversive noise, from no noise at all to the highest setting of about 105 dB, to be delivered to their partner.

In addition to the traditional self-report and behavioral aggression measures, we believed it was important to include a measurement paradigm that was invulnerable to demand and participant manipulation. For this reason we included physiological measures in our experiment, specifically cardiovascular indices of challenge and threat based on the work of Blascovich and his colleagues [35]. Although challenge and threat do not directly measure aggression, challenge and threat do point to benign versus malignant ways of coping with a situation, and furthermore, threat and aggression can be closely tied conceptually, as reactive aggression can be a response to threat. In the challenge and threat paradigm, we looked for three cardiovascular indices to differentiate the motivational states of challenge and threat. Challenge is defined as the condition that exists when resources to perform a particular task outweigh the demands, and threat is defined as the condition when demand outweighs resources.

To index challenge and threat, we examined cardiac performance (indexed by heart rate and ventricular contractility), cardiac output, and total peripheral resistance. The challenge pattern is indexed by increased cardiac performance, large cardiac output increases, and decrease in total peripheral resistance. The threat pattern is indicated by increased cardiac performance, small increase or no change in cardiac output, and small increase or no change in total peripheral resistance.

5.1.1. Findings and Implications

First, enjoyment and satisfaction with the game playing experience as measured by self-report questionnaire items was found to be higher for participants who played the game using an IVE platform versus a desktop platform [detailed results can be found in 34]. This result lends support to the proposition that IVE technology is more enjoyable and a preferred way to play video games and is therefore likely to expand in popularity and use in the coming years.

Our self-reported aggressive feelings measure, reported after the game play period, indicated that participants reported significantly more aggressive feelings after playing the exact same game in an IVE versus on a desktop computer. Our measures of behavioral aggression produced the same results, participants in the IVE condition had a significantly higher proportion of head hits than participants in the desktop condition. In terms of the competitive reaction time measure, participants who played the violent video game in an IVE administered significantly higher levels of aversive noise to a partner than participants who had played on a desktop computer platform.

Our physiological results were slightly complicated by the fact that we had considered the possibility that being attached to physiological measuring equipment, ambulatory though it was, might substantially change the experience of immersion in an IVE, possibly by creating breaks in presence. To examine this possibility and, hopefully, demonstrate that use of physiological measures was appropriate for our paradigm, we varied their inclusion as an independent variable, analyzing the responses of participants who had and had not been hooked up to physiological equipment. These analyses revealed no differences in self-report or behavioral measures by physiological measure inclusion, satisfying us that inclusion of the measures did not alter participants' experiences of the game environment. Analysis of the physiological data from those participants who had been hooked up to the measuring equipment revealed that participants who played the violent video game using an IVE platform displayed the threat pattern of response. Participants who had played the violent game on a desktop platform, however, were not significantly physiologically influenced by the experience.

Taken together, our four measures of aggression and aggression-related outcomes clearly suggest that playing violent video games using an IVE platform results in more detrimental aggression-related outcomes than does playing an identical game using a traditional platform. As previously suggested, we had reason to believe that this difference might be caused by differences in presence in the game environment. This hypothesis was partially supported by our data. We found full statistical mediation by self-reported presence of the relationship between playing platform and our self-report measure of aggression. Though presence fully explained the relationship between platform and our self-report measure, statistically, it did not explain the relationship for either of our behavioral measures. We suspect that this disconnect in mediating construct between our self-report and behavioral measures may be due to the explicit, self-report nature of our particular presence measure in these experiments, though this has not yet been experimentally confirmed.

These experiments support our hypothesis that playing violent video games using IVEs lead to increases in aggression over play using traditional desktop platforms, as well as lending support to our proposal that presence is responsible for this effect. Still at issue though, is whether the content of the game matters, as the previous experiments assessed only play in a violent game. Conventional wisdom would suggest that IVE use should intensify only content-related outcomes, in this case aggression from violent games. If this is the case, playing a nonviolent game in an IVE should not lead to increases in aggression. This prediction, however, had not yet been tested. Also yet unexplored was whether nonviolent themes might similarly be intensified through experiences in an IVE game with related content.

Figure 8-2. Screenshot of the nonviolent art-themed video game created for our experimental studies.

5.2. Effects of Violent versus Non-Violent Games

To address questions of game content-specific intensifying effects, we designed a second, nonviolent game with art-themed content [fully reported in 36]. The non-violent game was designed to remain as similar as possible to the violent gun fight-themed game, including a similar game environment and identical body movements. In this game, players shot paint at a canvas located across the virtual room to create an abstract painting. Players selected paint colors by moving to different locations in the virtual room. By randomly assigning participants to play either the violent or new nonviolent game using either a desktop computer or an IVE platform, we were able to investigate content effects (see figure 8-2).

In this particular experiment, we included two types of dependent measures, self-report and physiological. Our self-report measures included a multi-item scale of aggressive feelings, as well as a multi-item scale of artistic and creative feelings which were content-specific to the art-themed game. Physiological measurement for challenge and threat states was included for all participants.

5.2.1. Findings and Implications

As expected, our self-report aggressive feelings measure revealed a significant interaction between game violence content and game playing platform [detailed results can be found in 36]. Participants who played the violent game in an IVE reported the highest level of aggressive feelings whereas participants who played a nonviolent game in an IVE reported the lowest level. This result confirmed that IVEs, at least in this case, only intensify aggressive outcomes for games with violent content. Additionally, analysis of the cardiovascular psychophysiological data revealed a pattern consistent with our self-report aggressive feelings measure. Specifically, only participants who had played the violent game on an IVE platform exhibited increases in physiological responsivity indicative of threat influenced during game play.

On our self-report measure of artistic and creative feelings, however, we found no differences by condition. Neither game content type nor game playing platform influenced how creative and artistic game participants reported feeling after the game play period. This suggested that artistic and creative feelings did not seem to be transferred when playing an art-themed game, nor were these feelings intensified by IVE use. Because of the nature of these results, it is difficult to draw a firm conclusion as to their implication. It may be that our particular art game did not engender strong artistic feelings to begin with, or that these particular nonviolent feelings do not tend to transfer from video game play. These results might also suggest, however, that there is something special about aggression and violence that transfers and is intensified particularly easily by IVE game play.

5.3. Research Conclusions

Overall our program of research on video games played in IVEs has allowed us to clarify various issues related to advances in video game technology. We have shown that IVE use for violent video game play does lead to increased aggression and physiological threat over play on a traditional desktop computer platform. Further, we have shown that presence is apparently the critical factor by which IVE use leads to increases in self-reported aggressive feelings as well as congruent physiological responses. Furthermore, this intensifying effect of IVE game play is limited at present to violent and aggressive game content.

6. Virtual Violent Video Game Effects in Multiplayer Games

Though findings with respect to the effects of IVE video game play in single-player environments are important and useful in and of themselves, networked gaming is becoming quite popular. Multiplayer networked games appear to be a major direction in which mainstream video gaming is headed. How multiplayer IVE violent video games might affect players' interactions with one another, then, is not simply an academic issue. To date, however, very little research has been performed in this area, and in particular there has been no research specifically addressing any distinct effects of multiplayer games and potential unique patterns of resultant anti-social responses.

Though there may not be much in the way of research focusing on unique effects of multiplayer games, we do know that the outcomes following multiplayer games are similar to the effects of single-player games such that violent games played in a multiplayer context also lead to more aggressive outcomes than nonviolent multiplayer games [37]. In terms of multiplayer IVE violent games, then, it would follow that playing against an avatar (i.e., a representation of an actual human other) rather than against an agent (as in the study reported here) would tend to evoke the same sorts of aggressive responses.

A single study performed by Buckley and her colleagues examined effects of playing a violent video game in a multiplayer IVE [38]. In this study, participants played either a two-player shooting-themed game or a two-player game of virtual tennis. Findings from this experiment were mixed. Results from a post-game behavioral reaction time competitive task, similar to the one we used, revealed that participants administered more noise to opponents if they had played the violent game using an IVE (i.e. Cave-type) platform rather than a desktop platform. Researchers also reported an interaction such that levels of noise administered after playing the nonviolent game on a desktop system were by far the lowest. Results from written story completion task responses rated for aggressive cognitions revealed an interaction such that participants completed the highest proportion of stories aggressively after playing the desktop violent game and the IVE nonviolent game. Story completion results were inconsistent both with the behavioral results in the same study and with previous work. However, in general, results of the behavioral measure do show a similar pattern to results found for single-player IVE games.

More interesting than general aggression effects from the standpoint of a multi-user environment, however, is the question of whether there might be special consequences of playing against an avatar, particularly for a post-game shared IVE experience. One way we can look at this question is through the lens of Blascovich's [39, 40] threshold model of social influence in IVEs.

7. Social Influence in Virtual Environments

The threshold model of social influence posits that social influence will increase as social presence increases. According to this model, social presence itself varies as a function of four factors, two of which are internal and two external to the target of influence. According to the model, behavioral realism and agency (whether interaction partners are believed to be agents or avatars) are the two external factors influencing social presence. The internal factors are self-relevance of the situation and whether the social influence response in

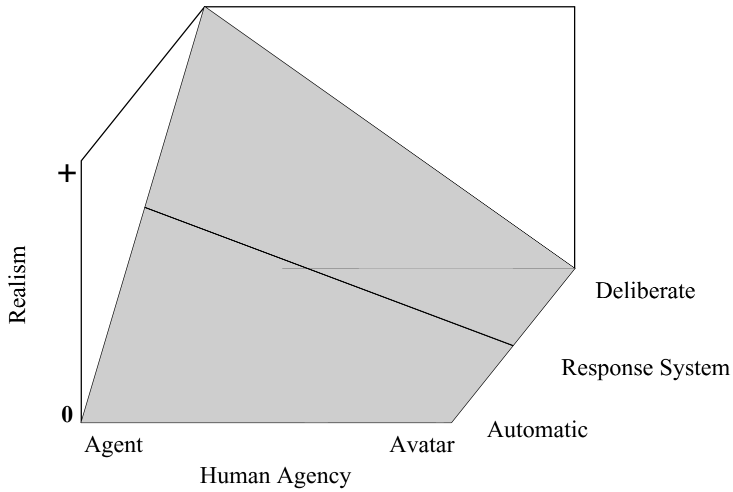


Figure 8-3. The three-dimensional threshold surface of social influence as a function of behavioral realism, human agency, and response system level.

question is a high versus low-level response. The threshold of social influence will be lowest when the target is interacting with what he or she believes is an avatar in a non-self-relevant situation where the response in question is low-level or automatic (see figure 8-3).

Previous research has used the model to determine both deliberate, or high-level social influence effects, and non-deliberate, or low-level social influence effects. The model has been used to identify the level of certain social responses based on the conditions under which a given response was elicited. For example, because a conformity response was elicited by both an agent and an avatar in a particular simulation, conformity was determined to be a lower-level response than social comparison which was elicited only by an avatar (and not elicited by an agent) in an identical situation [41]. According to the model, because conformity in this situation can be elicited by social interaction with an agent whereas social comparison cannot, we are able to conclude that conformity is the more automatic response. When the level of the social influence response behavior in question is held constant, however, the threshold model can also be used to examine the effect of variations in one of the other three factors affecting the social influence threshold.

In terms of our investigation of immersive virtual violent video games, aggression is the relevant social influence outcome. Aggression, particularly behavioral aggression, can be considered a low-level behavior when it is activated without awareness, as is often the case [42]. In the model, low-level processes such as aggression result in a lower threshold for social influence. Examination of the other three contributing factors can help us understand how likely such social influence aggression outcomes are, given the typical characteristics of the violent video game play situation.

Self-relevance, the other internal factor, involves the extent to which an interaction involves a central aspect of the self. There are some aspects of a virtual violent video game environment that are likely to be somewhat high in self-relevance as one might be motivated to play well so as to be favorably evaluated by one's opponent. In addition, individuals for whom video games are a large part of their life (i.e. gamers) may have a part of their self-esteem wrapped up in video game performance. In general, however, playing games is likely to be less self-relevant than many other activities, such as work-related collaboration and interpersonal interaction, which can occur in IVEs. We therefore would consider violent video game play situations, while not irrelevant, to be relatively low in self-relevance, which again is predicted by the model to lead to a lower threshold for social influence.

Examining the third factor, behavioral realism, it is important to note that while not all violent video games are the same, it is likely that there are more commonalities than differences in terms of behavioral realism. Inherent in video games is both a limited context and a limited range of player actions and responses. Though behavioral realism of a game opponent is partially tied up in the complexity of artificial intelligence programming, given context and input limitations, it does not take much for an agent or avatar in most games to appear to be enacting realistic behaviors. In most video games, then, particularly those of a violent nature where the context is often extremely limited, behavioral realism is likely to be relatively high. As behavioral realism increases, the threshold of social influence becomes lower.

The final factor, agency, is where we focus our examination on how the model can inform us as to the likely effects of multiple versus single-player IVE violent video games. Agency speaks directly to this issue as in single-player games one's opponent is by definition an agent whereas in multiplayer games, one's opponent is almost always an avatar. Because characteristics of virtual violent video game environments and of virtual violent video game play, overall, should result in low-to-moderate self-relevance and high behavioral realism, and because aggression, our social influence outcome, is relatively low-level, we would expect a low, shallow threshold of social influence. Given the configuration of factors we have outlined, therefore, we would not expect agency to matter in the majority of cases. In other words, whether a game opponent is or is believed to be an agent or an avatar, immersive virtual violent video games are expected to result in an aggression response. This interpretation is supported by the studies presented earlier where violent video game play against agent opponents resulted in aggressive social responses within the IVE (an increased proportion of head hits). The experiment also found increased behavioral aggression in a supposedly unrelated computer-mediated competitive game (the competitive reaction time/noise-blast game) which though in this case occurred outside of the IVE, may be somewhat analogous to a post-game interaction within the IVE.

We would also expect, however, that in the case where any of the other three variables, level of response, self-relevance, or behavioral realism should change such that social presence was decreased, the agency of the gaming opponent might begin to matter. For example, if we were to use a very high-level measure of anti-social feelings as our social influence response or if we were to limit the opponents' range of actions so as to reduce their behavioral realism, we might increase the threshold such that agency would become a crucial variable. Therefore, we would expect that playing a violent VE video game against an avatar would be more likely to result in aggression than playing against an agent, and would also be more likely to result in higher level, more intentional, forms of aggressive and hostile behavior.

8. Conclusion

The threshold model of social influence predicts that playing violent games in a networked multiplayer game environment versus a single-player environment should make a difference for aggressive outcomes within IVEs. Though not specified in the model, it would be logical to expect unique or intensified responses to one's game opponent over other unrelated individuals. Whether playing against avatars versus agents will prove to lead to heightened aggression and whether that aggression will be more acutely experienced by a gaming opponent versus an unrelated other, however, are still open empirical questions. Questions such as these highlight important themes for future research both in the area of video game effects and in the area of IVE research.

In a practical sense, however, it is unclear how often and in what situations game playing opponents currently have contact with one another either within IVEs outside of the game context or outside of IVEs following the game period. At present, VE video games are generally played in busy public spaces in situations where one might not even be aware of one's gaming opponents' identities. Furthermore, as technology advances, it brings closer the promise of combining immersive virtual technology and networked gaming where opponents will be separated by physical space, as in the networked non-immersive games of today. In these situations, too, it at first seems unlikely that gaming opponents will have much opportunity to interact outside of a gaming context. However, the same technology that brings distributed gaming promises to bring non-physically located social interactions, so perhaps more and more we may be meeting our gaming opponents in social VEs. Based on our investigations reported here, it seems likely that when these social interactions follow multiplayer violent video games, they will be marked by aggressive social influence.

For now though, we know that in both single and multi-player contexts aggression resulting from virtual violent video game play can be diffuse in nature and directed toward non-opponent targets. This alone seems to be reason enough to begin to expand our awareness to include the content of material experienced in IVEs and effects of the immersive nature of IVEs when considering their social or potentially anti-social effects.

References

1. Boyd, C. (1997). Does immersion make a virtual environment more usable? *CHI'97 Electronic Publications*. Available at http://www.acm.org/sigchi/chi97/proceedings/short-talk/cb.htm
2. Youngblut, C. & Huie, O. (2003). The relationship between presence and performance in virtual environments: Results of a VERTS study. In *Proceedings of IEEE Virtual Reality Conference 2003*, pp. 277–278.
3. Slater, M., Linakis, V., Usoh, M., & Kooper, R. (1996). Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess. In *Proceedings of the Virtual Reality Software and Technology*, pp. 163–172.
4. Tromp, J., Bullock, A., Steed, A., Sadagic, A., Slater, M., & Frecon, E. (1998). Small group behavior experiments in the COVEN project. *IEEE Computer Graphics and Applications, 18*: 53–63.
5. Axelsson, A.S., Abelin, Å., Heldal, I., Nilsson, A., Schroeder, R., & Wideström, J. (1999). Collaboration and communication in multi-user virtual environments: A comparison of desktop and immersive virtual reality systems for molecular visualization. In *Proceedings of the Sixth UKVRSIG Conference*, pp. 107–117.
6. Roberts, D.J., Wolff, R., & Otto, O. (2003). Constructing a gazebo: Supporting team work in a tightly coupled, distributed task in virtual reality. *Presence: Teleoperators & Virtual Environments, 12*: 644–657.
7. Schroeder, R., Steed, A., Axelsson, A.S., Heldal, I., Abelin, Å., Wideström, J., Nilsson, A., & Slater, M. (2001). Collaborating in networked immersive spaces: As good as being there together? *Computers & Graphics, 25:* 781–788.
8. Gabbard, J.L., Hix, D., Swan, J.E. II., Livingston, M.A., Höllerer, T.H., Julier, S.J., Brown, D., & Baillot, Y. (2003). Usability engineering for complex interactive systems development. Paper presented at *Engineering for Usability, Human Systems Integration Symposium*, June, Vienna, VA.
9. Bensley, L. & Van Eenwyk, J. (2001). Video games and real-life aggression: Review of the literature. *Journal of Adolescent Health, 29*: 244–257.
10. Dill, K.E., & Dill, J.C. (1998). Video game violence: A review of the empirical literature. *Aggression and Violent Behavior, 3*: 407–428.
11. Freedman, J.L. (2001). Evaluating the research on violent video games. Paper presented at *Playing by the Rules: The Cultural Policy Challenges of Video Games*, University of Chicago, Chicago, IL.
12. Griffiths, M. (1999). Violent video games and aggression: A review of the literature. *Aggression and Violent Behavior, 4*: 203–212.
13. Ivory, J.D. (2001). Video Games and the Elusive Search for their Effects on Children: An Assessment of Twenty Years of Research. Paper presented at the *National AEJMC Conference*, August. Available at http://asuwlink.uwyo.edu/~poison/video_games_and_children_.htm

14. Unsworth, G. & Ward, T. (2001). Video games and aggressive behaviour. *Australian Psychologist, 36*: 184–192.
15. Anderson, C.A. & Bushman, B.J. (2001). 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, 12*: 353–359.
16. Sherry, J.L. (2001). The effects of violent video games on aggression: A meta-analysis. *Human Communication Research, 27*: 409–431.
17. Funk, J.B., Buchman, D.D., Jenks, J., & Bechtoldt, H. (2003). Playing violent video games, desensitization, and moral evaluation in children. *Journal of Applied Developmental Psychology, 24*: 413–436.
18. Anderson, C.A., Carnagey, N.L., Flanagan, M., Benjamin, A.J., Eubanks, J., & Valentine, J.C. (In press). Violent video games: Specific effects of violent content on aggressive thoughts and behavior. *Advances in Experimental Social Psychology.*
19. Anderson, C.A. & Murphy, C.R. (2003). Violent video games and aggressive behavior in young women. *Aggressive Behavior, 29*: 423–429.
20. Bartholow, B.D. & Anderson, C.A. (2002). Examining the effects of violent video games on aggressive behavior: Potential sex differences. *Journal of Experimental Social Psychology, 38*: 283–290.
21. Bushman, B.J. & Anderson, C.A. (2002). Violent video games and hostile expectations: A test of the general aggression model. *Personality and Social Psychology Bulletin, 28*: 1679–1686.
22. Panee, C.D. & Ballard, M.E. (2002). High versus low aggressive priming during video-game training: Effects on violent action during game play, hostility, heart rate, and blood pressure. *Journal of Applied Social Psychology, 32*: 2458–2474.
23. Uhlmann, E. & Swanson, J. (2004). Exposure to violent video games increases automatic aggressiveness. *Journal of Adolescence, 27*: 41–52.
24. Heeter, C. (1995). Communication research on consumer VR. In F. Biocca & M.R. Levy (Eds.), *Communication in the Age of Virtual Reality*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., pp. 191–219.
25. Pham, A. (2004). Eye toy springs from one man's vision. *Los Angeles Times*, January 18. Available at http://www.latimes.com/business/la-fi-eyetoy18jan18,1,4576015.story?coll=la-home-business
26. Lombard, M. & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication, 3*, Available at http://jcmc.huji.ac.il/vol3/ issue2/lombard.html
27. Witmer, B. & Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*: 225–240.
28. Funk, J.B. (1993). Reevaluating the impact of video games. *Clinical Pediatrics*, *32*: 86–90.
29. Griffiths, M.D. & Hunt, N. (1995). Computer game playing in adolescence: Prevalence and demographic indicators. *Journal of Community and Applied Social Psychology, 5*: 189–193.
30. Ijsselsteijn, W. (2002). Understanding presence. Paper presented at the *Fifth Annual International Workshop PRESENCE,* October, Porto, Portugal.
31. Leyens, J.P. & Picus, S. (1973). Identification with the winner of a fight and name mediation: Their differential effects upon subsequent aggressive behavior. *British Journal of Social and Clinical Psychology, 12*: 374–377.
32. Geen, R.G. (1990). *Human Aggression*. Belmont, CA: Brooks/Cole Publishing Co.
33. Tamborini, R., Eastin, M., Lachlan, K., Fediuk, T., Brady, R., & Skalski, P. (2000). The effects of violent virtual video games on aggressive thoughts and behaviors. Paper presented at the

86th *Annual Convention of the National Communication Association*, November, Seattle, WA.
34. Persky S. & Blascovich J. (Under review). Violent video games and aggression: Effects of immersion and presence.
35. Blascovich, J. & Tomaka, J. (1996). The Biopsychosocial model of arousal regulation. *Advances in Experimental Social Psychology, 28*: 1–51.
36. Persky S. & Blascovich J. (Under review). Platform and gender effects on violent and non-violent video game play.
37. Ballard, M.E. & Wiest, J.R. (1996). Mortal Kombat™: The effects of violent videogame play on males' hostility and cardiovascular responding. *Journal of Applied Social Psychology, 26*: 717–730.
38. Buckley, K.E., Tapscott, R.L., Sidhartha, R., Rypma, C.A., Gentile, D.A., Anderson, C.A., Nacin, C., Sannier, A., Oliver, J.H., & Bushamn, B.J. (2004). The effects of violent virtual reality games on aggressive behavior and cognitions. Paper presented at the *American Psychological Society Annual Meeting*, Chicago, IL.
39. Blascovich, J. (2002). Social influence within immersive virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments.* London: Springer, pp. 127–145.
40. Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments. *CVE 2002: Proceedings of the 4th International Conference on Collaborative Virtual Environments*, pp. 25–30.
41. Swinth, K.R. & Blascovich, J. (2001). Conformity to group norms within an immersive virtual environment. Paper presented at the *American Psychological Society Annual Meeting*, Toronto, Ontario.
42. Todorov, A. & Bargh, J. (2001). Automatic sources of aggression. *Aggression and Violent Behavior, 7*: 53–68.

Chapter 9

THE PSYCHOLOGY OF MASSIVELY MULTI-USER ONLINE ROLE-PLAYING GAMES: MOTIVATIONS, EMOTIONAL INVESTMENT, RELATIONSHIPS AND PROBLEMATIC USAGE

Nick Yee

1. Introduction

Every day, millions of users [1, 2] interact, collaborate, and form relationships with each other through avatars in online environments known as Massively Multi-User Online Role-Playing Games (MMORPGs). For example, in a planetary system known as Corbantis, geological surveyors are busy inspecting their chemical harvesting installations to maintain their daily quota for a cartel of pharmaceutical manufacturers. These manufacturers, allied with a rebel faction, are struggling to research and supply key medical supplies to the front lines of the conflict. Corbantis is an incredibly sophisticated online environment capable of supporting thousands of users at a time. Users log on to the server from remote locations independent of each other, and interact with each other through graphical avatars to accomplish complex goals. But Corbantis is merely one planetary system out of many other equally complex worlds. These online environments offer tantalizing glimpses of how millions of avatars interact on a daily basis outside a laboratory setting and what users derive from that experience.

The study of MMORPGs is highly relevant to research on social interaction in Shared Virtual Environments (SVEs) and avatars at work and play in these environments. Although many of the theoretical implications of social interaction in virtual environments have been explored in the artificial confines of Virtual Reality (VR) research laboratories [3–8], MMORPGs are the only existing naturalistic setting where millions of users voluntarily immerse themselves in a graphical virtual environment and interact with each other through avatars on a daily basis. The opportunity to study what

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 187–207.

people actually do when they choose to be in a virtual environment with thousands of other people cannot be overstated, and the results and implications of a survey study of 30,000 MMORPG players will be presented in this chapter.

In the following section, the history and structure of MMORPGs will be presented, followed by an overview of the methodology used in the survey study. The demographics, usage patterns and motivations of users will then be presented. A combination of quantitative and qualitative data will guide the discussion of relationship formation, role exploration, skill transfer, and problematic usage in these environments. Finally, potential uses of these environments for social science research will be discussed.

2. History of MMORPGs

MMORPGs are a new class of Multi-User Domains (MUDs) – online environments where multiple users can interact with each other and achieve structured goals. The first MUD—an adventure game in a persistent world that allowed multiple users to log on at the same time—was created in 1979 by Roy Trubshaw and Richard Bartle [9]. While it is commonly thought that MUDs descended from table-top role-playing games (RPGs) such as Dungeons and Dragons, the two genres emerged around the same time and co-evolved beginning in the early 1970s [10] and became popular during the 1980s. Both games allow users to create characters based on numerical attributes (i.e. Strength, Dexterity, Intelligence) and templated roles (i.e. Warrior, Cleric, Druid) with different strengths and weaknesses. Game-play typically revolved around a combination of interactive story-telling and logistical optimizations under the guise of slaying monsters and attaining higher levels and skills. In RPGs, a designated Game Master controlled the outcome of events based on dice-rolls and references to charts and tables. In MUDs, this is controlled by the server.

As the graphical and processing capabilities of the modern personal computer increased, and as accessibility to the Internet became widely available, it became possible in the early 1990s to build MUDs with graphical front-ends. *Ultima Online*, launched in 1997, is recognized to be the first MMORPG—a persistent, graphical, online environment that allowed thousands of users to be logged on at the same time. The number of active users that Ultima Online could support was what distinguished MMORPGs from existing graphical MUDs. The second MMORPG, *EverQuest*, launched in 1999, quickly achieved a sustained user base of 400,000 and remains the most popular MMORPG in North America as of 2004 [2] even though at least 10 competing MMORPGs have emerged since then.

3. Details of MMORPGs

Users must purchase or download the specific MMORPG client and pay a monthly subscription fee (around 10–15 USD) to access the central servers. Users view the world in real-time 3D graphics, and use an avatar (a humanoid graphical representation of the user/player in the game) to interact with the environment and each other through a combination of a mouse-driven user interface and keyboard commands. Communication between users occurs through typed chat and animated gestures and expressions. The worlds of MMORPGs are vast and presented in rich, graphical detail. It typically takes several hours to traverse the entire world of an MMORPG, although different types of transportation are available to teleport users to different locations in the world. Users are given a large degree of control over the appearance of their avatars. For example, in the recent MMORPG *Star Wars Galaxies*, users are able to manipulate their avatar's gender, race (Wookie, Human, Rodian, etc.), skin tone, age, height, weight, musculature, cheek prominence, jaw prominence, brow prominence, nose shape, eye shape, eye color, hair style, hair color, mouth shape, lip fullness, and the presence of body markings or freckles.

Users choose from a set of professions or roles that the MMORPG provides. The permanence or fluidity of roles varies depending on the design of the environment. Each role has varying strengths and weaknesses and most MMORPGs are designed such that users must often collaborate to achieve goals within the environment. While early MMORPGs were based on fantasy medieval worlds made popular by RPGs and contained only combat-oriented roles (i.e. Warrior, Archer, Healer, etc.), recent MMORPGs have offered more diverse profession alternatives. For example, in *Star Wars Galaxies*, one can become a skilled musician, chef, hair stylist, pharmaceutical manufacturer, or politician.

Goals and rewards in MMORPGs typically use a random-ratio reinforcement schedule based on operant conditioning. Early achievements are quick, almost instantaneous, and gradually take more and more time and effort until progression becomes almost imperceptible. Most forms of advancement in MMORPGs require increasing cooperation or dependency on other users, oftentimes mutually beneficial. In *Star Wars Galaxies*, surveyors locate deposits of chemicals and minerals across different planets. To harvest these resources in bulk, surveyors must purchase mining installations from architects. Alternatively, surveyors can choose to sell the locations of rich deposits (i.e. the information itself) to miners rather than harvesting the resources themselves. Surveyors who choose to harvest resources may then become resource brokers who market those resources to artisans and manufacturers who need those resources to produce goods. Combat, medical or fashion goods then are sold on the open market and bought by mercenaries, doctors and other members of the general public.

Ultimately, each user decides which form of advancement they will pursue, and the richness and complexity of the environment eliminates the need for super-ordinate goals or storylines. Every user is motivated by a different combination of the possible rewards. The result is that adventures, stories, and most importantly, meaningful interactions and relationships between users emerge. Functional constructs within the environment facilitate these social networks—combat groups (temporary collaboration between a few users), guilds (persistent user-created membership organizations), and ideological alliances (agreements between guilds or "racial" groups).

4. Collaboration in MMORPGs

Combat-oriented collaborations in MMORPGs become incredibly complex once users have advanced beyond beginner levels. In typical battle-oriented scenarios, groups of 4–8 users are confronted by multiple enemy agents based on fairly sophisticated AI. These groups of users are typically composed of a balanced combination of roles and must communicate and perform effectively as individuals using a predetermined group strategy. Consider a fairly typical crisis situation. Certain enemy agents will run away and elicit help from allied agents when they are badly wounded. In a dungeon setting, these enemy agents typically run towards deeper, more dangerous locations. If the agent succeeds, he will return with several stronger agents. But if one user chases the agent, while the others decide not to, then that jeopardizes the group as well. This situation typically occurs while the group is still engaged with other half-wounded agents. Also remember that different users have different personalities (risk-taking propensities, assertiveness, and so on) and different stakes at this point of their adventure, and differ in their loyalty to the group and each other. In the span of 5–10 seconds, the risk-analysis, opinions and decisions of the group communicated over typed chat, or the solitary actions of a particular user, will determine the life or death of all members of the group. This particular type of crisis is also embedded into the larger context of existing tensions such as emergent leadership, group polarization, and personality differences.

More recent MMORPGs such as *Star Wars Galaxies* have also created collaboration scenarios of an entirely entrepreneurial nature. All non-basic goods in the environment (clothing, housing, pharmaceuticals, etc.) are produced by users. Unlike earlier MMORPGs, users cannot sell goods back to the server itself. All transactions, and the resulting supply, demand, and pricing of specific goods, are user-driven. The environment has mechanisms allowing users to survey for resources, harvest those resources, research schematics for assembling resources into sub-components, construct factories to mass-produce finished goods, and market those goods to the public. The process is so complex, time-consuming, and distributed over several skill sets that users typically specialize as resource brokers, manufacturers, or retailers, and typically form quasi-business entities with other users to facilitate that process. These entities have to communicate effectively, develop a coherent product strategy, assess market competition, and ensure that the production chain is running smoothly. Many users comment that being part of such entities feels like having a second job.

5. Existing Literature on MMORPGs

More than a decade has passed since Dibbel [11] pondered the significance of a virtual rape in *LambdaMOO* as it was embroiled in a political reform. The academic interest in MUDs it sparked was almost entirely driven by qualitative scholars. Turkle [12] articulated how these environments revealed the fluid and decentralized nature of identities. Others [13] have challenged the utopian visions of cyberspace, arguing that online communities do not foster racial equality but merely make racial minorities easier to suppress. The behavioral sciences have kept their distance from these online environments. With regards to video gaming in general, the field of psychology seems fixated on whether video games cause real-life aggression [14–23]. Considering that new forms of social identity and social interaction are emerging from these environments, is aggression the only thing worth our attention?

Academic attention in MMORPGs has largely been driven by economic and legal scholars. Castronova [24] has calculated the Gross National Product of the world of *EverQuest* by aggregating e-Bay sales of virtual items and currency, and has also shown that male avatars sell for more than female avatars of exactly the same capabilities [25]. Legal scholars [26] have examined the ownership of virtual property and whether avatars have enforceable legal rights. Griffiths [27] has also aggregated online poll data at websites catering to *EverQuest* players to provide the basic demographics and preferences of *EverQuest* players.

In essence, there has been very little research available that has explored the social interactions, relationship formation and derived experiences of the users of MMORPGs. Since the spring of 2000, the author has carried out an extensive survey study of over 30,000 MMORPG users that has examined who uses MMORPGs, what motivates their use, and the salience and impact of the experiences that emerge in these environments. While a previous paper [28] focused on presenting statistical findings from the data set, the following chapter summarizes and elaborates on those findings with qualitative data to provide a richer perspective of these online environments.

6. Methodology of Survey Study

The survey study consisted of a series of online surveys that were publicized in web portals that catered to MMORPG users from the years 2000 to 2003. The approximate number of active subscribers to each existing MMORPG was publicly available [1], and it was usually clear which MMORPGs comprised the bulk of all MMORPG players. Therefore, users of the four most popular environments were targeted for the study—*EverQuest*, *Dark Age of Camelot*, *Ultima Online*, and *Star Wars Galaxies*. A survey with new content was usually publicized every two to three months. Each survey took about 5–10 minutes to complete, and typically 2000–4000 respondents would complete each survey. In each survey, respondents were asked to provide their email if they were interested in participating in future surveys. At the beginning of each survey phase, in addition to the recruitment at websites, respondents already in the database were contacted via email to notify them of the new survey in which they could participate. Over the course of four years, 30,000 unique users participated in the survey study.

Lack of motivation and integrity in web-based surveys are two potential concerns, but studies have shown that web-based respondents are typically highly-motivated due to self-selection and anonymity does not have an adverse affect on data integrity (for review, see [29]). Sampling bias is also a concern. In particular, a skewed representation of dedicated and heavy users is possible. Because of market competition, demographic information about users of these environments is not publicly available; however, informal communication with representatives of some of these companies has corroborated the basic demographic representativeness (average age and average hours per week) of the sample. Also, the sampling bias in using a large, non-random sample of MMORPG users to generalize to other MMORPG users is probably not any riskier than the standard practice in experimental psychology of using small, non-random samples of mostly Caucasian students between the ages of 18–22 who are enrolled in introductory psychology courses to generalize to all of humanity (for example, see [30]).

7. Only For Teenagers?

The stereotype that only teenagers partake in these environments discourages broader interest in studying these environments. Indeed, the Journal of Adolescence recently dedicated an entire special issue to the topic of video game violence (February, 2004), fostering the stereotype that adults do not engage in these kinds of activities, or that somehow adolescents interact with video games in an entirely different way from how adults interact with them. Data from Griffiths' study [27] as well my survey study [28] challenge that stereotype. The average age of MMORPG respondents was 26.57 ($n = 5509$, SD = 9.19); the median was 25, with a range from 11 to 68. The lower and upper quartile boundaries were 19 and 32, respectively. Thus, in fact only 25% of MMORPG users are teenagers.

Many MMORPG users have stable careers and families of their own [28]. 50% of respondents ($n = 2846$) worked full-time, 36% were married, and 22% had children. The data showed that teenagers, college students, early adult professionals, middle-aged homemakers, as well as retirees were part of these environments. Indeed, these seemingly disparate demographic groups would oftentimes be collaborating and working together to achieve the same goals similar to the ones mentioned earlier. This finding is particularly striking given that these disparate demographic groups seldom collaborate in any real life situation.

8. Time Investment

The demographic reality of these environments is important to establish to frame the significance of the following data on usage patterns. Users spend on average 22.72 hours ($n = 5471$, SD = 14.98) each week in their chosen MMORPG. The lower quartile and upper quartile boundaries were 11 and 30 respectively. The distribution showed that about 8% of users spend 40 hours per week or more in these environments—the equivalent of a normal work week. The significant amount of time that users are willing to invest in these environments is further highlighted by the finding that 60.9% of respondents ($n = 3445$) had spent at least 10 hours continuously in an MMORPG. The correlation between age and hours spent per week was not significant, implying that the appeal of these environments is comparable for high-school students, middle-aged professionals and retirees.

9. Emotional Investment

The appeal and salience of these environments is further demonstrated by the degree that users are emotionally invested in their avatars and the environment. When respondents were asked whether the most positive experience they had experienced over the period of the past 7 days or the past 30 days occurred in an MMORPG or in real-life, 27% of respondents ($n = 2170$) indicated that the most satisfying experience over the past 7 days occurred in the game, and 18% of respondents indicated the same when the wording was changed to "over the past 30 days". With regards to the most negative event, 33% of respondents indicated that the most negative experience over the past 7 days occurred in the game, and 23% of respondents indicated the same when the wording was changed to "over the past 30 days".

Open-ended questions asking users to elaborate on examples of these experiences drew responses that revolved around interactions with other users [31]. Typical positive experiences involved an unexpected altruistic or courageous action by another user.

> He showed rare courage by staying until everyone was clear, including me, knowing that he would probably not make it out alive. That was the most selfless thing I had seen done before or since. He stayed, knowing the corpse retrieval that awaited him, the experience he would lose, and the wasted time he was about to experience because of it. He could have run and lived, but he did n't for our sakes. When you make sacrifices for people, they will remember, and the best groups are those built on loyalty, self-sacrifice, and courage. [male, 32]

On the other hand, typical negative experiences involved the selfish actions of other users, or actions or behaviors that constituted an attack on the competence or self-worth of a user.

> I was playing my enchanter at the time, and his partner turned out to be an enchanter, a level higher than I was. I was medding up after buffing the group and switching my spells back to hunting/guarding spells, when the new enchanter started casting everything I had just cast, overwriting everything I had done, telling the group what to do and commenting on how they obviously hadn't had a chanter with them who knew how to take care of their group and they were lucky he was there, he'd make sure they didn't get into TOO much trouble. I disbanded and headed for the zone, in tears of frustration. To be overwritten, pushed aside, and belittled was unbearable. [female, 36]

What is clear is that these environments encourage both time and emotional investment from the users, and that users derive salient emotional experiences from these environments.

10. Motivations

The usage patterns of users force us to examine what makes these environments so appealing. What motivates users to become so invested in these environments? User responses expose the varied and multi-faceted reasons for why users engage in these environments.

> After many weeks of watching I found myself interested in the interactions between people in the game, it was totally absorbing!!!! The fact that I was able to immerse myself in the game and relate to other people or just listen in to the "chatter" was appealing. [female, 34]

> I play MMORPGs with my husband as a source of entertainment. Overall it can be a cheaper form of entertainment where you can spend quite a bit of time with a significant other. To play well you end up developing more ways of communicating. [female, 31]

> I like the whole progression, advancement thing . . . gradually getting better and better as a player, being able to handle situations that previously I wouldn't have been able to. [male, 48]
>
> No one complains about jobs or other meaningless things. It's a great stress reducer. I like that I can be someone else for a couple hours. [female, 28]
>
> Currently, I am trying to establish a working corporation within the economic boundaries of the virtual world. Primarily, to learn more about how real world social theories play out in a virtual economy. [male, 30]

Having an empirical framework of articulating motivational differences between users is the foundation to understanding the emergence of more complex behaviors and interactions in these environments. This framework provides the foundation to explore whether different sections of the demographic are motivated differently, and whether certain motivations are more highly correlated with usage patterns or in-game preferences or behaviors.

In an attempt to create an empirical framework for articulating motivations for MMORPG usage, a series of 40 statements covering a broad range of motivations were generated based on open-ended responses as well as Bartle's [32] theoretical framework of "Player Types" based on his experiences in MUDs. Examples of the resulting items include: "I like to feel powerful in the game," and "I like to be immersed in a fantasy world." These statements were presented using a 5-point Likert-type scale and then analyzed using an exploratory factor analysis to arrive at a parsimonious representation of the associations among the 40 items [28].

The analysis produced five factors. The "Relationship" factor measures the desire of users to interact with other users, and their willingness to form meaningful relationships that are supportive in nature, and which include a certain degree of disclosure of real-life problems and issues. The "Manipulation" factor measures how inclined a user is to objectify other users and manipulate them for his personal gains and satisfaction. Users who score high on the "Manipulation" factor enjoy deceiving, scamming, taunting and dominating other users. Users who score high on the "Immersion" factor enjoy being in a fantasy world as well as being "someone else". They enjoy the story-telling aspect of these worlds and enjoy creating avatars with histories that extend and tie in with the stories and lore of the world. The "Escapism" factor measures how much a user is using the virtual world to temporarily avoid, forget about and escape from real-life stress and problems. And finally, the "Achievement" factor measures the desire to become powerful in the context of the virtual environment through the achievement of goals and accumulation of items that confer power.

It was found that male users score higher than female users on Achievement and Manipulation, whereas female users scored significantly higher on the Relationship, Immersion and Escapism factors. In other words, male users are more likely to engage in these environments to achieve objective goals, whereas female users are more likely to engage in MMORPGs to form relationships and become immersed in a fantasy environment. These gender differences resonate with findings by Cassell and Jenkins [33] and suggest that MMORPGs do not have one set of factors that appeals to everyone equally well, but instead, have a host of appealing factors each of which draws in users with different motivations. With regard to how these motivations related to usage patterns, among male users, age was inversely correlated with the Manipulation ($r = -0.33$, $p < 0.001$) and Achievement ($r = -0.27$, $p < 0.001$) factors, implying that younger male users tend to objectify both the environment and other users for their own personal gains. Among female users, age was inversely correlated with the Manipulation ($r = -0.15$, $p < 0.001$) and Immersion ($r = -0.13$, $p < 0.001$) factors.

The articulation of the different reasons why users engage in these environments allow researchers to explore usage preferences and behaviors in relation to the motivations of the user in addition to gender and age differences. It is simply not the case that all users engage in these environments for the same reason.

11. It's All Pretend?

Because these environments are labeled "role-playing games", it is easy to assume that users treat it as a simplistic game of pretend-play. The emotional investment that these environments derive from users is one way of countering that assumption. Users in fact take these environments very seriously. Other survey data also show that the majority of users indicate that the way they behave and interact with others in these environments is very close to how they behave and interact with others in the material world [31]. In other words, most users are simply being themselves rather than experimenting with new identities or personalities. It is also easy to assume that nothing serious or meaningful happens in or can be derived from these environments because they are merely semi-sophisticated forms of play. The following sections provide multiple lines of evidence to argue that many different kinds of serious social phenomena occur in these environments.

12. Relationships in MMORPGs

When asked to compare the quality of their MMORPG friendships with their material world relationships, 39.4% of male respondents ($n = 2971$) and 53.3% of female respondents ($n = 420$) felt that their MMORPG friends were comparable or better than their material world friends [28]. Furthermore, 15.7% of male respondents ($n = 2991$) and 5.1% ($n = 420$) of female respondents had physically dated someone who they first met in an MMORPG [28]. Thus, both platonic and romantic relationships seem to occur with significant frequency in MMORPG environments. This finding resonates with Walther's [34] notion of the hyperpersonal effects of computer-mediated communication (CMC).

Indeed, the ingredients that Walther proposed for hyperpersonal interactions—interactions that are more intimate, more intense, more salient because of the communication channel—all exist in MMORPGs. First, the communication channel allows the sender to optimize their self-presentation because interactants do not have to respond in real-time. Second, the receiver forms an impression of the sender by inflating the few pieces of information that the sender has optimized. Third, participants can reallocate cognitive resources typically used to maintain socially acceptable non-verbal gestures in face-to-face interactions and focus on the structure and content of the message itself, which comes across as more personal and articulate. Finally, as interactants respond to personal messages with equally personal and intimate messages, the idealized impressions and more personal interactions intensify through reciprocity. The cumulative effect is that the interaction becomes more intimate and positive.

It has also been suggested that there are factors unique to MMORPGs that facilitate relationship formation [35]. The kind of high-stress crisis scenario outlined earlier in the chapter occur with great frequency in these environments under different guises. When paired with the degree of emotional investment users place in these environments, many relationships are in fact triggered by these trust-building scenarios, analogous to boot camps and fraternity initiations in the material world.

> To succeed in EQ you need to form relationships with people you can trust. The game does a wonderful job of forcing people in this situation. Real life rarely offers this opportunity as technological advances mean we have little reliance on others and individuals are rarely thrown into life-or-death situations. [male, 29]

While it may appear that meeting other users with compatible personalities and interests seems like finding a needle in a haystack in these environments, users are in fact pre-selected for compatibility. 36% of employed respondents ($n = 1099$) work in the IT industry, and 68% of respondents ($n = 3415$) have experience with table-top role-playing games. IT workers are typically analytical and rational; RPG players are typically imaginative and idiosyncratic. Both tend to be non-conformist. MMORPG environments are a very specific form of entertainment—gradual advancement via avatars in a fantasy medieval or futuristic world with other users. Thus, in fact, MMORPG users are probably similar in more ways than not.

And finally, the fantastical metaphors employed in these environments encourage idealizations that parallel cultural myths of chivalric romance—knights in shining armor, clerics with glowing auras. Thus, these metaphors also encourage idealization in addition to the underlying inflated sense of compatibility due to hyperpersonal interactions. MMORPGs are environments where users are in fact falling in love with knights in shining armor:

> The MMORPG relationship is inexplicably more romantic, more epic, more dramatic . . . [female, 16]

MMORPG environments allow us to think about how the mechanics and functional constraints of a constructed world could be used to engineer the relationships that form [36]. User dependencies, the mechanics of death, and other structures all play a role in encouraging or discouraging relationships to form in these environments. MMORPGs allow us to ask questions about how the mechanics of a world influence the communities that form instead of focusing on individual interactions.

13. Romantic Partners and Family Members

There is a very different kind of relationship "formation" that can be explored in MMORPGs. Many MMORPG users participate in the environment with a romantic partner or family members. 15.8% of male respondents ($n = 1589$) and 59.8% of female respondents ($n = 311$) participate in the environment with a romantic partner, while 25.5% of male respondents and 39.5% of female respondents participate in the environment with a family member—a sibling, parent or child [28]. Open-ended responses from these users indicate that their online relationships shape, influence and allow them to explore their material world relationships.

Many romantic couples who participate in the environment together commented on how the environment highlighted their individual differences. For these users, the MMORPG environment reflected and accentuated differences in their personalities and worldviews.

> Our styles are totally different. For instance, I will rather play in a group just for company, even if the exp gain is minimal, whereas my partner tends to literally AVOID other players. I am often a pushy role-player, forcing others to RP or get out of my face. Thus I am unafraid of starting an argument, whether in /say, /tell or even /shout. This seems to make my partner very uncomfortable. For these reasons, if we are playing together we try very hard to compromise. However, I insist on having "solo" characters that I only play on my own. I tend to find his gaming style restrictive. [female, 23, engaged]

> I would say rather than having learned something new about him, it was more like it emphasized differences between us that I already knew about. He is very patient, I am very impulsive, etc. And these differences are a lot more apparent in a game situation. [female, 27, dating]

For other romantic partners, the MMORPG environment not only reveals individual differences, but it also comes to shape the relationship itself.

> Like children who play dolls to explore social situations and different perspectives, EQ enables us to look at issues of dependence/independence, and gender perceptions. It's increased the equanimity between us, and brought us closer through exercises in trust that transcend in game terms, class, level, and gender. We will discuss game scenarios and learn from each others perception (i.e., when to run). After 3 years of playing together we are a well-oiled machine, and can lead a group, follow or solo together or apart. [female, 34, married]
>
> Our relationship has definitely been enhanced. We're better now at working together towards goals. And we both really enjoy growing, learning and adventuring together. It's exciting to be involved in each other's triumphs. [female, 29, married]

Parents and their children who participate in these environments provide another perspective on how the MMORPG environment interacts with existing relationships. Many parents commented on how the environment allowed them to observe their children in social interactions that they usually had no access to in the material world. For them, the MMORPG environment became a window into parts of their children's identity that they had not known about before.

> I learned that my son is a very good strategist. I knew that to a degree before, but it has been eye opening to watch him lead a group. I did not know he had these skills. [female, 49]
>
> It added depth and clarity to many traits that I knew they had to see how they presented themselves in a different environment. Since I am pretty much removed from their circle of friends and can't watch them at school, EQ provides a window into their behavior outside of the house [female, 37]
>
> I found that my son handles himself in a very mature manner. (He's 13 now). I have also been told by many other players that know of our relationship how courteous and well spoken he is. [male, 49]

Other parents commented that the MMORPG environment has allowed them to transcend the strict roles of parent-child relationships in a rewarding way. The MMORPG environment not only shapes these relationships, but in fact restructures them by allowing the participants to redefine the boundaries of their material world roles.

> I think it has enhanced our relationship, we both treat each other more like equals and partners in our private life. It is much easier to talk to her now and I have found her talking to me about much more of her life and ideas. [female, 40]
>
> Yes, playing EQ with my daughter has been very enjoyable, and I have learned more about my daughters personality as she treats me as a friend on EQ and not a parent. [female, 40]

Thus, MMORPG environments are not only places where new relationships are engineered, but in fact are windows into existing relationships as well as catalysts for the restructuring of roles in those relationships.

14. Role Exploration and Skill Transfer

Turkle [12] articulated how MUDs allow users to explore new roles and identities. MMORPG environments are also used for these purposes.

> In reality I'm an Army Officer, very assertive and aggressive. In MMORPGs I'm more like I wish I could be, quiet, introspective and sensitive of other's feelings. Taking on different roles has also taught me to "walk a mile" in other shoes before judging—not useful as an army officer, perhaps, but very useful in becoming a quality human being. [male, 42]

> When I play my male characters, other male members of the party will listen to me better, take me more seriously. In my male form I could give orders and have them listened to, where as a female, my characters aren't always taken quite as seriously. Also, where my female characters were given many gifts by random players when they were young, I didn't see it happening with my males, which I didn't mind at all. I've enjoyed the higher level of "respect" for my abilities that seems to come with playing in a male body. [female, 22]

But beyond exploring how MMORPGs can shape the identities of individuals, these highly social and structured environments also allow us to explore whether certain valuable skills learned in an MMORPG can transfer to the material world.

Personal advancement in MMORPGs typically involves collaboration among groups of users in an attempt to achieve a challenging task. Thus, a prime candidate for acquired skills is leadership skills. In emergent groups within the MMORPG environment, leaders deal with both administrative as well as higher-level strategy issues, most of which arise and have to be dealt with spontaneously. Administrative tasks include: role assignment, task delegation, crisis management, logistical planning, and how rewards are to be shared among group members. Higher-level strategy tasks include: motivating group members, dealing with negative attitudes, dealing with group conflicts, as well as encouraging group loyalty and cohesion. These issues are even more salient in long-term social groups, such as guilds, which have formalized membership and rank assignments. In other words, MMORPGs provide many opportunities for short-term and long-term leadership experiences. As one user notes:

> I've never been one who is particularly comfortable with a leadership role in real life. In the game, friends and I left another guild that no longer suited us for various reasons and formed our own. I was approached by several of these friends

> to assume leadership of the guild and agreed, even though I was uncertain of my suitability. I've grown more accustomed now to directing various aspects of running the guild and providing a vision and leadership to the members. Follow-up and assertiveness now feel more natural to me even in real life. It has been an amazing opportunity to push myself beyond my boundaries and a rewarding experience. [female, 46]

This sentiment is shared by many users. In the survey study [28], 10% of users felt they had learned a lot about mediating group conflicts, motivating team members, persuading others, and becoming a better leader in general, while 40% of users felt that they had learned a little of the mentioned skills. This is striking given that these environments are not structured pedagogically to teach leadership. Acquisition of leadership skills in these environments is in fact an emergent phenomenon. But more importantly, these findings demonstrate that real-life skills can be acquired or improved upon in these environments. Certainly, self-reported assessments are not robust assessments, but these findings lay the foundation for more controlled studies of the acquisition of complex social skills in these environments.

15. Problematic Usage

As mentioned in the section on time investment, 8% of users spend 40 hours or more in these environments, and 70% have spent at least 10 hours continuously in an MMORPG in one sitting. Both quantitative and qualitative data suggest that a small, but significant, group of users suffer from dependence and withdrawal symptoms [37].

> I am addicted to EQ and I hate it and myself for it. When I play I sit down and play for a minimum of 12 hours at a time, and I inevitably feel guilty about it, thinking there a large number of things I should be doing instead, like reading or furthering my education or pursuing my career. But I can't seem to help myself, it draws me in every time. I have been out of work now for over a month and now find myself in a stressful, depressed state that is only quelled when I am playing EQ, because it's easy to forget about real world troubles and problems, but the problem is when you get back to the real world, problems and troubles have become bigger, and it's a bad, bad cycle. I've tried quitting seriously on several occasions. There are serious withdrawal pangs, anxiety, and a feeling of being lost and not quite knowing what next to do with yourself. [male, 26]

On 5-point Likert scales, 15% of respondents ($n = 3989$) agreed or strongly agreed that they become angry and irritable if they are unable to participate. 30% agreed or strongly agreed that they continue to participate in the environment even when they are frustrated with it or not enjoying the experience. And

18% of users agreed or strongly agreed that their usage patterns had caused them academic, health, financial or relationship problems. Agreement with the mentioned statements was significantly positively correlated with average weekly use of the environment. Even more striking, 50% of respondents (n = 3166) considered themselves addicted to an MMORPG in a direct "yes"/"no" question. While it may be difficult to draw a line between healthy and unhealthy usage of these environments, it is clear that certain users are engaged in problematic usage of these environments.

While the design of these environments, such as the sophisticated reward cycles based on operant conditioning paradigms [38], certainly plays a role in engaging users in problematic usage, it would be overly-simplistic to focus entirely on the architecture of the environment itself. After all, that perspective fails to account for why only certain users engage in problematic usage. It also fails to take into account that different users are motivated to participate in the environment for different reasons. One proposed model of problematic usage [37] approaches the environment as a place where many common anxieties can be overcome. For example, users who have low self-esteem can become powerful and competent in these environments and they are driven to achieve in these environments as a way of overcoming anxieties they have in the material world. Or for example, users who feel undervalued in the material world can become needed and valued members of groups or guilds. Users with poor self-image can choose to be as attractive and physically fit as they desire. Users with low internal locus-of-control gain a stronger sense of agency in these environments. Users with stressful problems can use these environments as escape. In short, these environments are seductive for some users because they empower them in ways specific to their anxieties.

16. Online Environments as Potential Social Science Research Platforms

The structure and design of these environments make them good candidates for a host of alternative uses for social scientists. For example, traditional personality assessment techniques are typically transparent and reactive. Because actions in massively multi-user online environments can be tracked unobtrusively by the server, every users' attitudes and personalities may be tracked using behavioral measures. And because users are personally invested in their avatars and the environment, every decision they make is personally revealing. The length and frequency of utterances, as well as the breadth and depth of a user's social network can all be meticulously measured and tracked over long periods of time. This database of measures provides rich longitudinal profiles of individual users as well as how they rank amongst a large sample of other users. One could think of MMO environments as a gold-mine of personality data as well as a platform to develop unobtrusive personality assessment tools.

The arguments that Blascovich *et al.* [39] make for the use of immersive virtual reality technology as a methodological tool for social psychology can also be applied to MMORPG environments. The movements, interactions and preferences of large numbers of users can all be tracked unobtrusively and recorded. For example, one could implement a transformed social interaction ([40], see also Bailenson and Beall's chapter in this volume), such as non-zero-sum gaze, on one MMORPG server and use another server for control, and track the aggregate changes in mean length of utterances or topology of social networks. The MMORPG environment allows us to answer social psychology questions on a social level rather than an individual level. How does non-zero-sum gaze or other transformed social interactions reshape social networks, alter the flow of information, or affect trust in a social organization? As social organizations proliferate in MMORPG environments, research in transformed social interactions becomes even more important as it will inform us of how designers could engineer these environments to encourage the formation of strong and trusting social networks.

17. Conclusion

As scholars who studied MUDs [11, 12] pointed out, our virtual identities and experiences are not separate from our identities and experiences in the material world. They co-evolve as they shape each other. MMORPGs are not a new form of play as much as a new communication medium that affords new forms of social identity and social interaction.

While typical VR environments try to replicate human avatars in contemporary physical locations, MMORPGs offer fantastical avatars and worlds. After all, if you could be anyone anywhere, would you choose to be exactly who you were? This tension begs the broader question—given that we are not constrained to human forms or modes of movement and interaction that are bound by laws of physics, why do we insist on replicating them? If the body is merely the original prosthesis [41], can we not think "outside of the body"? Insisting on visual veridicality also forces us to abandon interesting issues in self-representations. What might decisions in virtual self-representation tell us about users?

The strong appeal of these environments also has interesting implications. MMORPGs do not only appeal to teenagers. They are online environments where young professional adults, middle-aged home-makers and retirees interact and collaborate on a daily basis. More importantly, the average MMORPG user spends more than half a work week in these environments. As more people engage in online environments instead of watching TV, it raises interesting questions with regard to Gerbner's cultivation theory [42]. Gerbner found that heavy TV viewers have a worldview that overestimates violence and the percentage of legal-enforcement workers in the general population due to their over-representation in TV content. Might certain worldviews be cultivated by heavy exposure to online environments? For example, users are given a high degree of control and agency in MMORPGs, and all events are based on underlying numeric variables. So it might make sense to ask whether heavy users have a stronger internal locus of control, or apply a more closed-system perspective on thinking about events in the material world.

The data presented also explored how virtual environments impact relationship formation in different ways. Not only can these environments facilitate formation of relationships, but they are also windows into and catalysts in existing relationships. More importantly, relationships can be thought of as being engineered by the architecture of the environment. For example, what are the potential effects of transformed social interaction [40] on social interactions at a community level? It also leads us to wonder how a community in the material world could be shaped by allowing them to interact in an engineered virtual environment.

The excessive usage exhibited by certain MMORPG users might appear problematic at first, but in fact forces us to ask whether the mechanisms of appeal in MMORPGs could be harnessed for pedagogical purposes. Story-path curriculums, used in certain schools, embed the syllabus of each term in an ongoing hypothetical setting, such as an iron-forging village in 19th century England. Every student takes on the role of a member of the village, such as blacksmith, pastor or farmer, and the syllabus material is woven into relevant tasks that the villagers encounter. For example, basic algebra may be embedded into a task that tried to optimize ratios of profitable crops, while social policy material may be embedded into a town meeting over a local epidemic of scarlet fever. The goal is to increase interest in learning by making the material personally relevant to students. The structure of MMORPGs are well-suited for story-path curriculums, and in fact, would also allow classes from different schools to inhabit different villages and create a larger social community that worked together to resolve conflicts or achieve common goals.

Finally, MMORPGs also blur the distinction between work and play in intriguing ways. Case studies of virtual real-estate brokers [43] are one of many compelling examples of how digital media blur the distinction between work and play. These users sell virtual real-estate (as well as virtual weaponry and currency) for real-life currency on auction sites such as eBay. More compelling are the "sweatshops" in developing nations that hire youths to generate profit by accumulating these virtual goods and currency and then selling them for real-life currency [44]. In this case, work and play are indistinguishable. As Andrejevic [45] has pointed out, interactive media creates digital enclosures that allow work to be performed under the guise of entertainment. For example, in There.com, brand-name fashion designers use the environment as a marketing test-bed for new clothing designs. Sales of the test products and whether users who have large social networks buy them are aggregated automatically. The irony is that not only do these users have to pay a monthly fee to subscribe to the environment, but they are performing free labor for a third-party corporation. As these environments become more sophisticated, we can imagine them transforming into predominantly sites of economic activity under the guise of interactive entertainment.

We have seen that MMORPG users become highly invested in these environments, and that serious social phenomena occur in these environments that can create, shape and restructure relationships in the material world. Every day, millions of users log on to worlds like Corbantis, performing highly-specialized and complex tasks, interacting and collaborating with each other through avatars. Some of them are accumulating virtual real estate to trade for US dollars. Some are married to people they have never met. Some are collaborating with their children to produce advanced pharmaceuticals, while others are planning a mayoral campaign. Indeed, if we are interested in the social lives of avatars, the citizens of worlds like Corbantis have a great deal they can tell us.

References

1. Woodcock, B. (2003). An Analysis of MMOG Subscription Growth. Available at http://pw1.netcom.com/∼sirbruce/Subscriptions.html
2. Corpnews.com, MMOG Roundup: Depressing 2004 Edition. (2004). Available at http://www.corpnews.com/news/fullnews.cgi?newsid1081411764,6286
3. Bailenson, J.N. & Blascovich, J. (2002). Mutual gaze and task performance in shared virtual environments. *Journal of Visualization and Computer Animation*, *13*: 1–8.
4. Leigh, J., DeFanti, T., Johnson, A., Brown, M., & Sandin, D. (1997). Global telemersion: Better than being there. In *Proceedings of ICAT '97*, no pp.
5. Mania, K. & Chalmers, A. (1998). A classification for user embodiment in collaborative virtual environments. In *Proceedings of the 4th International Conference on Virtual Systems and Multimedia*. IOS Press—Ohmsha, Ltd., pp. 177–182.
6. Normand, V., Babski, C., Benford, S., Bullock, A., Carion, S., Chrysanthou, Y., Farcet, N., Harvey, J., Kuijpers, N., Magneat–Thalmann, Raupp–Musse, S., Rodden, T., Slater, M., & Smith, G. (1999). The COVEN project: Exploring applicative, technical, and usage dimensions of collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *8*: 218–236.
7. Slater, M., Sadagic, A., Usoh, M., & Schroeder, R. (2000). Small group behavior in a virtual and real environment: A comparative study. *Presence: Teleoperators and Virtual Environments*, *9*: 37–51.
8. Zhang, X. & Furnas, G.W. (2002). Social interactions in multiscale CVEs. In *Proceedings of the ACM Conference on Collaborative Virtual Environments 2002*, Bonn, Germany, September 30–October 2, pp. 31–38.
9. Bartle, R. (1990). Early MUD history. Available at http://www.mud.co.uk/-richard/mudhist.htm

10. Koster, R. (2002). Online world timeline. Available at http://www.legendmud.org/-raph/gaming/mudtimeline.html
11. Dibbel, J. (1993). A rape in cyberspace. Village Voice, Vol. XXXVIII, No. 51, December 21.
12. Turkle, S. (1995). *Life on the Screen: Identity in the Age of the Internet*. New York: Simon and Schuster.
13. Nakamura, L., Rodman, G., & Kolko, B. (2000). *Race in Cyberspace*. New York: Routledge.
14. Anderson, C.A. & Bushman, B.J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology*, *1*(1): 19–41.
15. Bushman, B.J. & Anderson, C.A. (2002). Violent video games and hostile expectations: A test of the general aggression model. *Personality & Social Psychology Bulletin*, *28*(12): 1679–1686.
16. Anderson, C.A. & Dill, K.E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality & Social Psychology*, *78*(4): 772–790.
17. Anderson, C.A. & Bushman, B.J. (2001). 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*, *12*(5): 353–359.
18. Ballard, M.E. & Lineberger, R. (1999). Video game violence and confederate gender: Effects on reward and punishment given by college males. *Sex Roles*, *41*(7–8): 541–558.
19. Funk, J.B. & Buchman, D.D. (1996). Playing violent video and computer games and adolescent self-concept. *Journal of Communication*, *46*(2): 19–32.
20. Funk, J.B., Hagan, J.D., Schimming, J.L., Bullock, W., Buchman, D.D., & Myers, M. (2002). Aggression and psychopathology in adolescents with a preference for violent electronic games. *Aggressive Behavior*, *28*(2): 134–144.
21. Griffiths, M. (1999). Violent video games and aggression: A review of the literature. *Aggression & Violent Behavior*, *4*(2): 203–212.
22. Scott, D. (1995). The effect of video games on feelings of aggression. *Journal of Psychology*, *129*(2): 121–132.
23. Ferguson, C.J. (2002). Media violence: Miscast causality. *American Psychologist*, *57*(6–7): 446–447.
24. Castronova, E. (2002). Virtual worlds: A first-hand account of market and society on the cyberian frontier. Available at http://ssrn.com/abstract=294828
25. Castronova, E. (2002). The price of "man" and "woman": A hedonic pricing model of avatar attributes in a synthetic world. Available at http://ssrn.com/abstract=415043
26. Lastowka, G. & Hunter, D. The laws of virtual worlds. (2003). Available at http://ssrn.com/abstract=402860
27. Griffiths, M. (2003). Breaking the stereotype: The case of on-line gaming. *CyberPsychology and Behavior*, *6*: 81–91.
28. Yee, N. (2004). The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. Submitted for publication.
29. Gosling, S.D., Vazire, S., Srivastava, S., & John, O.P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, *59*(2): 93–104.
30. Bailenson, J.N., Shum, M.S., Atran, S., Medin, D.L., & Coley, J.C. (2002). A bird's eye view: Triangulating biological categorization and reasoning within and across cultures. *Cognition*, *84*: 1–53.
31. Yee, N. (2004). The Daedalus project. Available at http://www.nickyee.com/daedalus

32. Bartle, R. (2003). Hearts, clubs, diamonds, spades: Players who suit MUDs. Available at http://www.mud.co.uk/richard/hcds.htm
33. Cassell, J. & H. Jenkins (Eds.). (1998). *From Barbie to Mortal Kombat: Gender and Computer Games*. Cambridge, MA: MIT Press.
34. Walther, J.B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, *23*(1): 3–43.
35. Yee, N. (2003). Inside out. Available at http://www.nickyee.com/daedalus/archives/-000523.php
36. Yee, N. (2003). Engineering relationships. Available at http://www.nickyee.com/-daedalus/archives/000429.php
37. Yee, N. (2002). Ariadne: Understanding MMORPG addiction. Available at http://www.nickyee.com/hub/addiction/home.html
38. Yee, N. (2001). The Virtual Skinner Box. Available at http://www.nickyee.com/-eqt/skinner.html
39. 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*(2): 103–124.
40. Bailenson, J.N., Beall, A., Loomis, J., Blascovich, J., & Turk, M. (2004). Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *13*(4), 428–441.
41. Hayles, K. (1999). How we became Posthuman: Virtual bodies in Cybernetics, Literature, and Informatics. Chicago: University of Chicago Press.
42. Gerbner, G. (1973). *Communications Technology and Social Policy: Understanding the New Cultural Revolution*. New York: Interscience Publication.
43. Dibbell, J. (2003). The unreal estate boom. *Wired*, *11*(01).
44. The Walrus. (2004). Game theories. Available at http://www.walrusmagazine.com/-04/05/06/1929205.shtml
45. Andrejevic, M. (2002). The work of being watched: Interactive media and the exploitation of self-disclosure. *Critical Studies in Media Communication*, *19*(2): 230–248.

Chapter 10

QUESTING FOR KNOWLEDGE—VIRTUAL WORLDS AS DYNAMIC PROCESSES OF SOCIAL INTERACTION

Mikael Jakobsson

1. Introduction

In this chapter I will discuss the nature of social interaction and game-play in the massively multi-player online game *Everquest*. Based on my studies of this particular type of virtual world, I will address the question of how the experience of participating in virtual worlds changes over the course of time and the implications of this on how we conceive virtual worlds from a design perspective. In parallel, I will also address some methodological implications of performing ethnographic studies in an environment where new levels of interacting with the world and its participants continuously reveal themselves—like new levels in a platform game.

For social interaction to exist within a virtual environment there have to be social actors in it. This book as well as its predecessor [1] covers a wide range of cases of social interaction from separate individuals all the way up to large groups of people. In this volume, yet another dimension is added as this and Nick Yee's (chapter 9) look at a fully three-dimensional graphical environment that attracts and supports participants on a massive scale averaging several thousands of participants in the same virtual world.

In a previous essay (in [1]) I argued for the importance of an *inside view* in order to grasp the unique properties of the social environments of virtual worlds. In this chapter my focus is on the vantage point of a participant embedded in the world. I will describe the game from four different points in the process of progressing through the game as a player and through the empirical study as a researcher. These discrete reference points will then be connected in an attempt to reveal a richer picture of the processual qualities of virtual worlds in terms of understanding the object of study as well as the process of studying it.

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 209–225.

While it is generally a bad idea to recount a study in chronological terms, I will do exactly that here with the explicit purpose of showing how the understanding of an object of study can change when looked at over an extended period of time. In this way I wish to uncover the research process rather than hiding it beneath the surface of the text.

Until recently, there were no systems of this scale and complexity in existence—but that changed with the arrival of the game *Everquest* in 1999. Built on a real-time three-dimensional graphics system similar to e.g. Active Worlds but with a closed graphics library, *Everquest* took the role-playing genre of computer games to a new level of technological sophistication. The response from the gaming community was overwhelming and a flood of similar games has followed establishing a new genre of computer games, the massively multiplayer online game MMOG. (This kind of game is sometimes referred to as MMORPGs, or role playing game, but since the role-playing element is often very weak and some games lack it completely, the RP part of the acronym is becoming ever less appropriate).

Everquest continues to be one of the most successful games in the genre with over 2.5 million copies of the game sold and currently having over 400,000 paying subscribers of which around 100,000 play simultaneously at peak times [2]. Operating via a client/server architecture, the world of *Everquest* is a fairly elaborately rendered three-dimensional space in which players battle a variety of characters and creatures. The world of *Everquest* is inhabited by participants who wander a vast terrain covering a number of continents. Much like old-style tabletop role-playing games, players create a character to play by selecting a "race" such as human, halfling or troll, a class such as warrior, shaman or wizard, and setting a few other parameters defining their character. On special servers players can engage in combat against other players or in more formalized role-playing, but the overwhelming majority play against the monsters in the world without bothering with the role-playing aspects of the game.

2. The Beginner Level: Everquest as a Single-Player Game

To say that the *Everquest* world is vast feels like a bit of an understatement. The amount of places to visit and monsters to battle is almost endless. Most of this is, however, out of reach for a beginner. When a player first enters the world, it is in the designated hometown. Before the player has put a few experience levels under her or his belt and acquired some basic equipment, it is impossible to venture very far beyond the city gates without falling prey to the creatures that roam the game world.

The very first obstacle that the player has to overcome is, however, the user interface. Before being able to interact with the environment, with non-player characters, or with other participants, the technical aspects of interaction have to be mastered. A crucial part of this interaction is the communication with other players. All text-based communication is carried out in different chat channels. As long as the player has not selected a channel for the text input, the keys on the keyboard work as shortcuts to different commands. This means that if you target a non-player character and start typing to say something to it without indicating the appropriate channel, the keystrokes will be interpreted as commands. This will in turn prove deadly if the intended message includes the character "A" since that key by default is set to trigger the command "auto-attack", and the non-player characters in the game are powerful enough to quickly eradicate a beginner player. (Measures are continuously taken to make the system more forgiving to the beginners, including a recent move of the default key for auto-attack from typing the "A" key to typing the significantly less used "Q" key.)

One of the promises of three-dimensional graphics was the possibility of making interaction interfaces more intuitive and easy to use for beginners. What we see in the user interface of *Everquest* should not necessarily be interpreted as an indication that these promises were incorrect. Instead, it speaks of the powerful impact of history in technology development. Despite the ocean-wide difference in complexity and experience of play between the early text-based MUDs and today's MMOGs, many designers as well as players have come to today's MMOGs via the MUDs, bringing their history of previous influences and experiences with them. *Everquest* displays a number of technological and cultural traces to these earlier systems.

The command interface is so reminiscent of the DIKU (Datalogisk Institut, Københavns Universitet) MUD system that people coming from a MUD of that variety testify to feeling right at home in *Everquest*. The *Everquest* team at one point had to release a sworn statement that they had not used any actual code from DIKU MUD which resulted in the following response: "The DIKU group is proud that 'the DIKU feeling' has found its way into a game as enjoyable and award winning as Everquest" [3]. *Everquest* also borrows from other sources. Many of the sites and monsters bear a striking resemblance to those in fantasy fiction in general and Tolkien in particular. The game system is heavily influenced by Dungeons and Dragons [4]. There are also technical solutions and modes of navigation originating from predecessors within the field of three-dimensional graphical environments.

The notion of a beginner becomes more complex when the history of the players—and the ways it coincides with the history of the game—is taken into account. A Tolkien reading MUD veteran with previous experience of virtual worlds is a very different kind of beginner than the person without this background.

After a few fatal encounters with merchants and other non-player characters in the hometown, the player will avoid at least the more devastating interaction mistakes. There will, however, still be a barrier between the beginner and the other players in the game. One very palpable factor that keeps beginners separate from other players is the geography. While the beginners cannot stray very far from the guarded cities, the more experienced players have to find other hunting grounds in order to find strong enough monsters to gain experience points from slaying them.

Besides the geographically imposed obstacles to playing the game together with the more experienced players, there is the lack of strategic knowledge. The abilities of different characters are designed in a way to make it beneficial for players to form groups and hunt together. The safety and efficiency of the group is dependent on the participants performing their tasks correctly and being able to efficiently communicate with the other group members. These skills take time to learn and a beginner lacks both the knowledge and the ability to communicate that is required. To accommodate for this, the developers have made it possible for all classes to easily defeat monsters on their own the first ten or so experience levels. A player who decides to start a second character will on the other hand have the required knowledge set and ties to the social networks between players [5] to be able to traverse the isolation of the beginner player.

After having played *Everquest* for a week or so and gained around ten experience levels, I could claim to have gained first-hand experience of the object of study and my notes were full of impressions from the intense experience. But what had I really seen? My perception of the game was that it was a fairly repetitive hack and slash game mainly played solo in a fairly confined space. I knew there was much more of the world to explore out there and that the game was full of other players, but I had no way of knowing if the experience of being in the game world and playing the game would change significantly or if it was going to be more of essentially the same thing. The only way to experience this for myself was to continue playing.

3. The Intermediate Level: Everquest as a World of Personal Communities

Just like a player bursting out in a triumphant "ding" signaling that a new experience level has been reached, I remember feeling a sense of elation when I stepped into the world of interconnected personal networks that signifies participation in the world of *Everquest* at an intermediate level. In my case it was not so much a question of being rewarded after a long hard struggle, as having contacts outside the game that helped me get "connected" on the inside. In [5], Taylor and I give a detailed description of the structure and importance of social networks in the game and the process of socialization. There we describe how social networks share structural properties with the mafia (as we understand its organization from popular fiction). Everything from initiations, pledges of trust and allegiance, vows of silence and favor systems exist both in the mafia and the social networks of *Everquest*. They both also seem to have their origin in a need for protection in an environment of insufficient control of law and order. Here, I will concentrate on describing the way the experience of the game and the world changes as a result of moving from the outside to the inside.

While the beginner stage is passed fairly quickly, the second stage will take months, sometimes years, and ranges approximately from level twenty to sixty. For the majority of the players most of the time is spent in groups at this stage of the development of the character. The gaming session typically begins with the player spending some time looking for a group to join, or starting a group and gathering the other five players needed. These players will then pick a place to set up camp and start killing the monsters there. When a monster is killed a new one will emerge at the same spot after a set time so the group will not run out of things to kill no matter how long they stay. Normally, a session lasts a couple of hours.

Although the players can move around more freely in the world at this stage, they do so to a very limited degree. Since the monsters you kill have to be on par with your own experience level to yield good experience earnings, there are only a few zones to choose from at any given experience level. The activity of camping around the same monsters for hours on end—sometimes referred to as experience grinding—might seem repetitive and even a bit boring. But with a fairly routine task to perform to keep the development of the character in motion, the players have plenty of time to socialize with the other group members and other friends in the game. Some people have social ties to other players already when they come to the game in the form of family members or friends. Others start from scratch in making friends through playing and hanging out together and before long, almost all players have created a social network of friends and acquaintances within the game.

From essentially being a single-player game on the first few levels, the gaming experience transforms into a rich social experience for the intermediate player. The game play stays basically the same. Instead it is the context around the actual monster killing that has changed. Just as in non-game virtual worlds, it turns out that the social interaction is the very foundation for the appeal of the world (see chapter 9 in this volume for more about social interaction as game motivation). At this stage of the game, the social networks are still loose and informal for the most part. Every player has his own web of contacts primarily organized by the friends list function in the game. The players perceive belonging to a community but each player's community looks different; in other words, they are personal communities as defined in [6]. With continued playing this will, however, change.

4. The High-End Game: Organized Play

Somewhere around level fifty to sixty and after perhaps a year or so in the game, players start feeling that there is something—more specifically their epic weapon—that they are missing and that it is time to try to do something about it. All classes in the game can take on an epic quest that will result in an epic weapon as reward if it is completed successfully. The epic quests are designed to require the help of more than just a handful of friends to complete. Besides the monsters that need to be defeated for the epic quest, the high-end game also includes many other tasks that require a raid force of up to seventy strong players to conquer.

The personal communities that players create during their time as intermediate players are not strong enough to support the level of organization that the high-end game requires. There are many reasons to be in a guild—such as belonging to a community and protection of the gaming experience from disruptive forces—but it is not until the high-end game that it becomes more or less a necessity. Guilds are formalized social networks managed by the players with the support of a few in-game tools. To start a guild, a minimum of ten players have to commit to joining. The person starting the guild becomes guild leader and chooses which other players to authorize as officers. The guild gets a dedicated chat channel and a tag under their names showing which guild they belong to.

The role of an officer in a guild differs depending on the size and type of guild. The bigger guilds have officers dedicated to specific tasks such as organizing raids or handling recruitment while the officers of smaller guilds tend to all do a bit of everything. I am a member of a medium sized guild with approximately three hundred registered characters. For about a year I have been involved in managing the guild as an officer, raid leader and for a brief time guild leader. We are a social guild which means that we do not make any particular requirements on our participants in terms of experience level, abilities, play time or attendance and we only raid once a week.

We do, however, require that our members adhere to our guidelines for social behavior within the game. These guidelines—posted on the guild website—cover topics of honorable and fair behavior towards other players and helping guild members achieving their goals. We also keep a fairly strict norm on language use in the guild channel since many of the participants are combining playing with looking after their kids. The social guilds tend to attract a more mature type of players while raiding guilds often appeal to a younger crowd. Older players may still put a significant number of hours into the game every week but often cannot commit to playing in the way a raiding guild demands. Out of these differences in appeal grows a difference in culture between the two types of guilds.

The downside of not belonging to a raiding guild is that some parts of the world require persistent raiding and these zones are the ones where you can find the best items to equip your character with. An alternative to joining a raiding guild is to join a raid channel. These channels are set up and administered by one or more guilds that wish to raid but lack the critical mass of players within their own guild. They therefore offer access to the channel to players who do not have a guild or are in a non raiding guild but still wish to raid occasionally. The raid channels are of course established to benefit the guilds behind them, but work as a kind of community service to social players with a taste for raiding.

At one particular raid we picked up a person along the way who said that he was looking for a guild to join and asked if he could tag along. Since he was a high level character of a class that we had use for, we were happy to bring him with us. While he handled his job of bringing monsters to the rest of us very well, he was also a little too eager to get a cut of the items we were getting by killing the monsters, suggesting that he had special needs for some of the items and that they should be defaulted to him instead of letting the luck of the draw decide who would get it. The officers running the raid and handling the looting and distribution of items that dropped from the monsters, however, paid no attention to these requests.

When we headed home at the end of the evening he mentioned that he wanted to join our guild. At the time the rules for inviting people into the guild were that at least three officers had to agree that a person was suitable for membership before issuing an invitation. Since there were more than three officers present at the raid we could immediately make a decision and he was invited. This was to be the last time someone was invited to the guild this way.

It only took until the next day before our guild leader received a private chat message from another player in the game issuing a complaint against our recently invited guild member. The complaint regarded a case of "ninja looting" which means that a player takes an item from a defeated monster that she or he does not rightfully deserve (see [5] for more on social consequences of the possibility to "ninja loot").

The guild leader then proceeded to interview the accused, the person issuing the complaint and one more person who had been in the same group at the time of the incident. He then posted all this material in the officer-only section of the guild website and requested comments from the other officers. After a few postings back and forth between the officers it was the guild leader who decided that the person should be removed from the guild.

The last thing a guild wants is rumors of questionable behavior to start spreading on the server, since the reputation of a guild affects the reputation of all its members and will make it harder for the guild to form alliances with other guilds. So as a direct consequence of the incident the officers decided to tighten up the recruitment process. Now applicants to the guild have to post an application on the guild website and then spend some time in a special recruitment channel in the game before they can be invited. This increase in rules, guidelines and formalized procedures is typical of guilds growing in numbers and reputation on a server. It also leads to that the officers of the guild spend more and more of their game-time on community management.

As an officer in general and a guild leader in particular, the nature of play shifts dramatically from mostly minding your own business and helping out others when needed, to management responsibilities that take up more time than the actual monster killing. Besides making sure that the guidelines of the guild are adhered to, officers spend much of their time reading up on how to tackle certain monsters, screening guild applicants, managing guild bank money and items and keeping the guild website in shape.

Somewhere in the middle of all this, the actual killing of the monsters still works more or less the same way as at level one, but the scale and complexity of what it means to play the game has reached a level that makes the gaming experience completely different from that of a beginner player.

5. The Endgame: Players Turning Against Players

Before the first expansion of *Everquest* was released there was a very limited selection of monsters to kill at the very high-end level of the game. The guild leader of one of the first raiding guilds on one of the first *Everquest* servers told me that after having done all the other content in the game, there were only two monsters left of interest to them. It was the two dragons Lady Vox and Lord Nagafen. His guild was the first to defeat the dragons but another was not far behind. Despite a fair amount of rivalry between the two guilds they managed to stay out of trouble by taking turns killing the two dragons. The situation became more complicated when a third guild on the server became strong enough to challenge the dragons. The limited resources in the game had created a volatile situation.

At one point, one of the guilds failed in an attempt to kill one of the dragons and when they came back later for a new attempt, the next guild scheduled was already there setting up for their turn. The exact details of what happened that night may differ depending on who you ask, but everybody agrees that things got ugly. Players tried to get players from the other guild killed and the verbal exchange between the raid forces was harsh and abusive. The incident led to an outright war between the guilds and both sides lost members as a result of it. Peace finally came when all the involved guilds agreed to follow a web-based event calendar operated by neutral players. With the aid of this calendar, guilds could make reservations for a monster they wanted to kill and the others would stay away from it until their turn came. The event calendar was successful in making rival guilds share the limited resources equally. But eventually one guild decided that they were not going to honor the calendar any more and would try to kill any of the dragons whenever they would see fit to do so. This guild could be described as a rogue guild, a guild that thrives on gaining a reputation—but not for being good and honorable but for being bad and doing whatever they feel like.

Most servers have one or a few of these guilds and they are always the source of huge amounts of discussions and complaints on the community message boards. While the majority of players harbor negative attitudes towards the rogue guilds, there are also players who are attracted by their boldness and attitude. The impact of these guilds on the servers is undisputable. Once a guild had publicly declared that they were not going to honor the event calendar, the initiative was in effect dead since it required backing from all parties to work as intended.

Things had, however, changed in the game world. New continents and monsters had been added to the game and the two dragons were no longer the most rewarding or prestigious targets. Since there were more possible targets to choose from, the need to take turns for the monsters was no longer as pressing. The issue of bottlenecks still existed though.

When the epic quests were introduced in The Ruins of Kunark—the first expansion of the game—many players wanted to kill the particular monsters that held the pieces they needed to complete the quest and receive the epic weapon for their class of player. At the time, the dragon Zordakalicus Ragefire, needed for the cleric epic, took several days to spawn again after it had been killed. The discrepancy between supply and demand created a severe bottleneck and according to [7] the waiting list for a shot at killing Ragefire was over a year on some servers. Other servers had no waiting list, but instead had to deal with large groups of players waiting around, hoping to engage the dragon first whenever it would spawn, leading to a situation reminiscent of the one with Vox and Nagafen.

While the developers tried to eliminate some of the problems stemming from players fighting over limited resources by simply adding more content and making the end-game more diverse, they were reluctant to shorten the spawn time of Ragefire with the motivation that the cleric epic is a powerful item and that "in Everquest power and rarity usually go hand-in-hand." [7].

The designers continuously have to deal with the trade-off between frustration and sense of accomplishment. If a quest is too easy in terms of time investment or manpower, the items gained from it will not be regarded as very special by the players. On the other hand there are players who know that they will never see large parts of the content of the game since they cannot make the commitment required. This leads to an ever changing environment where the developers try to find new ways of accommodating different styles of play with every new expansion of the game—at the same time as older content is under constant revision.

One of the defining traits of Everquest is that players exclusively fight against non-player characters and not against each other as in many MUDs and other MMOGs. In the endgame, however, when the selection of attractive targets becomes narrower, the guilds tend to turn against each other. While players still cannot directly fight each other, the competition for monsters between raiding guilds can be fierce at times and disputes often lead to trash talking and open animosity between members of the rivaling guilds. Many players who initially were attracted to the game partly because of its focus on collaboration rather than direct competition are deterred by this development. Although their character has become strong enough to pass the entry requirements for a raiding guild, some players choose not to join one, thus excluding themselves from a large number of zones and even larger number of monsters. Others thrive under these conditions and claim that it is at this point that the game truly begins.

6. The Final Level: Death

The following interaction took place in Plane of Knowledge, a hub in the game world that always is full of players, on a Saturday in September 2004. The names of the characters have been replaced with generic names. Since the communication in the game more or less constitutes a language of its own, I have added some explanatory remarks. All of this was said in the out of character channel which is a way of reaching all players in a zone without having to use the more intrusive shout channel.

Monk says, "Bye bye Everquesters . . . monk gona go FD [feign death is a monk specific ability used to fool monsters that you are dead] one last time in Qeynos [one of the hometowns] . . . . if i knew you, well met . . . . . . see you on EQ2 [Everquest 2] or WoW [World of Warcraft]"

Player 1 says, "Bye Monk!"
Player 2 says, "Take Care Monk . . . gl [good luck] man"
Player 3 says, "Fare Well MOnk!!!"
Player 4 says, "Bye Monk"
Player 5 says, "Be safe mon"
Player 6 says, "MONK NOOOOoooo not EQ2"
Player 7 says, "Farewell friend"

Death is a constantly present part of life in *Everquest*. A character that dies in the game respawns at a predetermined location and can run back and reclaim the possessions from the dead body, only suffering a minor penalty to the experience level. But since there is no final goal to the game, the player will sooner or later have to make the decision to stop playing the game altogether.

While the *Everquest* game system is constantly evolving and many additions have been made following requests from the player community, there is not a single line of code or advice written to help players who want to quit.

Once a player stops paying the monthly subscription fee, the character lingers on in a kind of limbo for an undisclosed period of time before eventually risking deletion from the database and being gone forever without a trace. I have come across many examples of how players react to and try to deal with this issue. Some players sell their characters for real money [4]. Others keep paying without playing or give away their characters. I have also encountered players paying for other people's accounts to keep those characters alive. There is even a guild called Zombies of EQ that works as a support group for former *Everquest* players [8].

When Sony announced that the Planes of Power expansion to *Everquest* was going to include something called graveyards, my first thought was that they finally had addressed the issues of the inevitable death of characters in the game, but unfortunately these were only designated areas for corpses to re-spawn in order to make them easier to retrieve. But what if graveyards where players really could bury their characters when they are done with the game were implemented? Maybe a culture of funeral ceremonies would develop within the player community where friends and guild members could say farewell to the character in a way that would provide a sense of closure both for the player who is leaving the game as well as those who have developed a relationship to the player within the game. While a high score list is a fitting way for gamers to leave a trace from their encounter with a coin-op game, and a "hall of fame" website is a suitable addition to a racing game, the persistent nature of MMOGs creates special needs for the kind of traces the gamers should be able to leave behind them when they go.

7. Discussion and Conclusions

7.1. The World as a Process

Everquest is both a game and a virtual world. As a game, it needs to drive the process of playing forward. Exactly what that process is can only be determined individually and at given points in time since different players have different motivations for playing and these motivations change over time. The most important driving force in the game, however, is to increase the abilities and experience level of one's character. The pursuit of experience points could be regarded in terms of a number of possible paths traversing geographical space. Right outside all hometowns are beginner areas with low level monsters that players can kill to gain the first few levels, but soon they need to move on through the world in order to keep the experience points rolling in. The issue

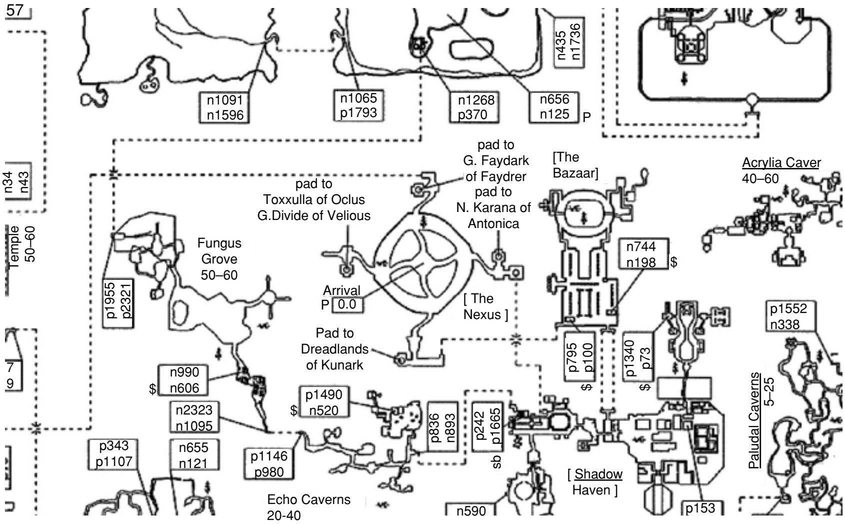


Figure 10-1. Area of a spatially oriented map of the moon Luclin.

of where to go next once the experience gain starts dropping off is a constantly hot topic among the players.

This gives rise to a need from the players to map the game landscape in two different ways. The first is the traditional geographical maps which show which zones are adjacent to each other and how to get from one place to another (figure 10-1). The second is a process-oriented description showing which zones provide the best trade-off between difficulty of the monsters and gain in experience.

While the first map shows the shortest way from one place to another in the virtual landscape, the second way of describing the world focuses on the optimal way to develop the statistics of your character. The process-oriented progression through the game becomes even clearer in the high-end game. In figure 10-2, the high-end zones are ordered in a flow-chart telling players how to progress through the zones to finally get to the Plane of Time, the most rewarding zone in the game at the time when this illustration was made. It also tells us that this world can be understood as a flow of people through the environment working their way towards their goals.

In the case of Everquest, the process that the world is there to support—or provide a pleasurable resistance against, depending on how you look at it—is the development of the player's characters. The process-oriented nature is, however, nothing unique for the MMOG game worlds. In earlier studies of general purpose virtual worlds, I have come to the conclusion that virtual worlds need an activity of some kind to keep them going. The activity provides the driving force that propels the process on which the world depends forward. This should both help us understand the success of MMOGs

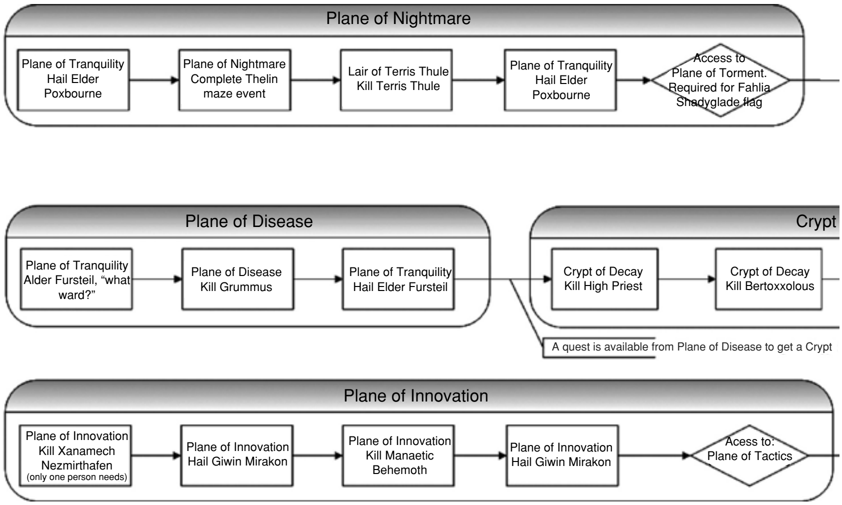


Figure 10-2. Excerpt from a process-oriented map showing how to get to the Plane of Time.

and tell us something about how to create successful virtual worlds for other purposes.

7.2. *One World?*

From the very beginning *Everquest* has struggled with the problem of its own popularity. For one thing, the number of players wanting to be in the same world at the same time leads to technical problems. Without any restrictions to the number of players that can be present at the same location at the same time, the fact that all the player's movements and actions alter the state of the world which has to be distributed to everybody puts a big strain on the networking and processing power of both clients and servers. On top of that, the fairly advanced graphics in the game put a strain on the rendering capabilities of the participants' computers. All these factors contribute to a lag between the issuing of an action by the player and the actual execution of that action. This lag is of course especially annoying if your life is in the balance, which it often is when fighting monsters is on the agenda.

Besides making the technology handle all the people who want to play the game, there is also the issue of having enough content in the world for everyone as was discussed in Section 5. With the release of a number of expansions that all add more territory to explore and more monsters to fight, the developers have tried to reduce some of the bottlenecks on the paths of progression through the virtual landscape and create an alternative solution to the same problem as the event calendar aimed to resolve. In the fourth expansion—Planes of

Power—the trend towards creating alternate progression routes was broken with the introduction of the so called flag progression path: The new zones were divided into tiers and access to higher tier zones was given by killing certain monsters in the lower tier zones giving the character a "flag"—an intangible check mark that works as a key to the next zone. The beginning of the flag progression path can be seen in figure 10-2. In the figure we can for instance see that a player has to kill the Manaetic Behemoth in the Plane of Innovation—a tier one zone—to get access to the tier two zone Plane of Tactics.

Again the focus of many players was turned towards the same monsters with traffic jams as the result. The rogue guilds found a particularly sinister way of benefiting from the situation. Since a guild cannot go to any zones beyond the zone they try flagging their members for, a rogue guild that already has that flag can reduce the competition in the zones they are hunting in by making sure that they kill the monster that gives the flags to get there whenever it spawns, an activity known as "cock blocking."

The developers again reacted to the situation, but this time with a new strategy. Instead of creating a wider range of content, something called instanced zones was introduced. What makes the instanced zones special is that a new copy of the zone is created every time a group of players enter it. In Gates of Discord—the seventh expansion—progression through the new zones is based on the completion of a number of "trials." These trials take place in instanced zones—which means that a guild can begin a particular trial and another guild can arrive five minutes later and start the same trial. For each guild, an instance of the same zone with the same monsters in it is created. The need to wait for particular monsters to re-spawn is thereby eliminated, as is the possibility of blocking other guilds by keeping monsters inaccessible.

The instanced zones, however, break one of the more fundamental principles of virtual world design. This principle—we can call it the singular world rule—states that all the participants are part of the same geographical world and if two participants go to the same place in that world, they will meet each other. This rule was in one sense broken already from the beginning of *Everquest* since there are a number of servers all running their separate copy of the game world. But the instanced zones further weaken the concept of a singular world that all participants are part of. Here the players can still communicate with each other and thus share the same social space, but if they are not part of the same instance, they cannot meet with each other.

In the eighth and latest (as of this writing) expansion—Gates of Discord—the developers have once more adjusted the direction of the development of the world and moved away from instanced zones. Alan Crosby, Community manager of *Everquest*, explains the change like this: "We do feel that the task system [quests given on demand] is a better solution than instanced zones, as they will not remove you from the community and isolate you in a little area by yourself. With tasks you remain a part of the world at large" [9].

The singular world rule has not been debated much in the past. It has always been seen as desirable by the virtual world design community; maybe because it seems to dominate virtual world fiction completely. But when a world becomes as large as the worlds the science fiction authors write about, we realize that there is a distinct difference between design ideals and the best solution in practice. In this regard, *Everquest* works as an interesting test-bed for virtual world design concepts.

7.3. Researchers at the Doorstep

In game reviews and research presentations alike, it is very common to hear references to character creation, exploration of the hometown and endless killing of low-level monsters such as rats, snakes, and spiders. The problem with these descriptions of the game is that they do not capture the typical experience of the game of the people who actually play it. You only create your character once, and although you can make more characters, most players only ever make a few and put substantial play time into even fewer. Once the character is created, however, it can be developed for years and years within the game. By character development I do not just mean the way the statistics of the character such as experience level, skills and abilities are developed. More importantly, the social networks within the game that are slowly developed over time contribute significantly both to the possibilities of success in the game as well as a rich and rewarding gaming experience.

As I have tried to show in this text, the experience of the game changes dramatically based on where in the process the player is. This is easily overlooked since the layers existing beyond the current position are in many ways hidden to the player. I myself have several times thought that I had reached a status quo where the gaming experience would not change dramatically again—only to be proven wrong by continued play. The understanding of the properties of the game world goes hand in hand with a more developed experience of the game as a player. I understand that not everyone can spend years on the same object of study and I do believe that there is a place for snapshot observations of game worlds and game cultures. My point is rather that there are things to be seen that cannot come through any other process than immersion over long periods of time.

7.4. It is All about Learning

If you ask *Everquest* players why the game appeals to them you will find a more or less even mix between exploration, achievement, socialization, empowerment and escapism. This list pretty well corresponds to the typical ideas of what games in general provide to their players. If you instead look at what the players do, you realize that much of the activities revolve around learning more about the game world and passing on that information to other players.

Studying tactics for how to defeat certain monsters or how to get to a certain place is typically thought of as a meta-activity by the gamers themselves. It is something that is needed to do in order to advance in the game, but not considered as gaming as such. It is perhaps not surprising that the learning aspect is diminished when gaming pleasures are discussed since learning can be connected to work, school and studies—which is exactly what a significant portion of the *Everquest* players want to get away from when they play. It is nevertheless always present in the gaming activities and a substantial time is spent on learning more and more about the game world. Reputation in the game is also closely associated with the knowledge a player has about zone geography, the value of tradable items, tactics for killing monsters or the history of other guilds and players.

After two years of playing and studying the game I have started to take interest more and more in questions connected to leaving the game and have begun to interview people who contemplate or have decided to leave the game after having played for a long time. In these interviews the issue of learning plays a strong, almost dominating, role. Most players connect the loss of passion for the game with having seen as much of it as they can hope to see at their level of play and having so much knowledge about the game that it has in some regard become transparent to them and therefore lost its mystical appeal.

This is where my two roles in the game merge. Both the quest of satisfying my needs as a player and my quest to grasp the structures of the game are tightly connected to knowledge and understanding of the world, and to reaching a point where my experiences have reached a state of saturation. I have several times experienced that I have reached a point where I know what I need to know and that it has become time to move on to the next study. But every time I have found new information or some aspect of the game that I had been previously unaware of, and I have been pulled back in again. After years of playing more than an hour every day on average I have to conclude that virtual worlds not only can be seen as processes and places but also that these are processes in a constant state of change and development; they are dynamic. This means that the inside view always can be developed further by continued participation as long as the world continues to exist.

References

1. Schroeder, R. (Ed.) (2002). *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer.
2. Sony Online. Everquest press release. (2003). Available at http://sonyonline.com/corp/press_releases/030503_EQ_growth.html

3. DIKU Group, The DIKU MUD Homepage. (2003). Available at http://www.dikumud.com/everquest.aspx
4. Taylor, T. (2002). Whose game is this anyway?: Negotiating corporate ownership in a virtual world. In F. Mäyrä (Ed.) *Computer Games and Digital Cultures Conference Proceedings*. Tampere: Tampere University Press.
5. Jakobsson, M. & Taylor, T. (2003). The Sopranos meets Everquest—Social networking in massively multiuser networking games. fineArt forum, 17:8, August 2003.
6. Croon, A. (1997). R U out there? On personal communities in cyberspace. In K. Braa & E. Monteiro (Eds.), Proceedings of the 20th IRIS Conference Social Informatics, Oslo: Department of Information and Computer Science, University of Oslo, pp. 591–604.
7. BBC News. (2002). Cyber heroes forced to wait for glory. Available at http://news.bbc.co.uk/1/hi/technology/2129912.stm
8. Zombies of EQ. (2002). Available at http://www.guildportal.com/Guild.aspx?GuildID=1669&TabID=1417
9. Crosby, A. (2004). Omens of War Chat #2. Available at http://eqforums.station.sony.com/eq/board/message?board.id=Crier& message.id=38&no_redir=true

Chapter 11

PLAY AND SOCIABILITY IN THERE: SOME LESSONS FROM ONLINE GAMES FOR COLLABORATIVE VIRTUAL ENVIRONMENTS

Barry Brown and Marek Bell

1. Introduction

While online games have become increasingly popular in recent years, there has been very little overlap between games research and virtual environments researchers. Indeed, one could argue that for a number of years, the design of video games have been ahead of virtual environment research, not only in technical aspects such as graphics or networking, but also in how game designers have managed their online worlds as social environments. Designers of online games have had to take seriously both the details of social interaction between individuals, but also how these interactions play out in the broader socio-economic balance of their online worlds [1].

In this chapter, we explore the lessons which collaborative virtual environments (CVEs) could derive from online gaming environments, focusing on *mundane interaction*.

Our activities and experiences in the real world depend in many ways upon mundane interaction for their operation [2]. Organisations whatever their size, in meetings and elsewhere, rely on talk [3]. Even the market transactions of currency traders depend upon chat for their coherence and reproducibility [4]. In a similar way in virtual environments it is in avatar-to-avatar interaction that experiences are configured. For virtual environments to be successful, we need to be able to interact with others around objects, refer to objects in our talk and share our awareness of other players and their movements [5].

This chapter focuses on these interactions, exploring how in the seemingly simple building blocks of talk and interaction around objects, enjoyable experiences are formed. We focus on the study of one game in detail, the social environment *There*, examining how its flat displays of colour come to form

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 227–245.

meaningful social experiences for its players. As with other virtual environments, *There* moves beyond text chat to support acting together around objects. These shared activities generate a qualitatively different experience to the textual interactions common in MUDs and MOOs.

We focus on how these interactions build two key features of social life: *play* and *sociability*. Play is a prevalent feature of our experiences both in leisure and work. Indeed, although often presented as distinct from work, play is an integral part of work as well as leisure. Play gives us an ability to rest, learn, or experience and experiment with new activities and experiences. In online games, play is a focal concern. In particular, in *There*, we discuss how its non-competitive nature makes it a more playful online environment than other, more competitive online environments. Yet this can present a challenge to players in deciding what to do next, and co-ordinating their activity with others.

The second feature we address is the role of *sociability*. While it is hardly surprising to see online environments as social environments, the more conventional meaning of "social"—as in "sociable"—has been somewhat neglected. We argue that the key issue is not necessarily what sorts or number of social relationships are formed online, but rather how those relationships are performed in online spaces. We use Simmel's discussion of sociability to give traction on understanding what it is about online environments that does (or does not) allow us to socialise online. In particular, we will argue that since in online virtual environments we can *do things with others*, we can "perform" our friendships in these environments.

2. Method: Studying *There*

Our original interest in *There* came from its innovative design for supporting social interactions [6]. *There*'s designers paid special attention to supporting avatar to avatar interactions, and chat more generally. This environment was therefore from the start designed as a social environment. *There* (www.there.com) has been open to the public since October 2003, although at that time it had been in beta testing for over a year, and development for over six. It is an online commercial environment, charging around $4 a month for access. *There* shares many of the features of other online virtual environments, such as "Active Worlds" (www.activeworlds.com), "Second Life" (www.secondlife.com), and other online, persistent role-playing games such as Everquest or Star Wars Galaxies. *There* is a persistent world with objects which can be manipulated, customisable avatars representing each user, and various facilities for interactions between avatars, and between avatars and objects.

Rather than as a competitive game as such, *There* is marketed as a "virtual getaway"—a world where social interaction and play are the main activities. There is no overall goal to *There* and its environment supports a range of

Figure 11-1. A screenshot from There.

activities such as buggy races, paintball, flying jetpacks, treasure hunts, card games and even playing with virtual pets. Avatars are capable of expressing emotional gestures, and chat is displayed in speech bubbles within the game world, word by word, rather than in the complete lines of text displayed in instant messaging (figure 11-1). *There* also has a rich range of features for organisation that sit outside the 3D space in separate 2D web browser windows. This augments *There's* virtual world with support for instant messaging (both text and audio), forums, tools for organising virtual "events", and forming groups. The interface of *There* is therefore split between a conventional set of webpages and a 3D virtual environment.

We have studied *There* for nine months, playing *There* for two hours each week. Our methods are informed by an emerging approach to studying online activity characterised as "virtual ethnography" [7]. This approach involves the familiar techniques of ethnography with a significant amount of time spent online in the research setting, observing, participating and taking field notes. As with all ethnographic work, the researcher's own experiences are taken as key data, yet due to the persistent nature of online data, virtual ethnography also makes extensive use of textual materials such as webpages, screenshots and videos.

A special focus of our study was the interactions between players in the game. Most ethnographic studies of online worlds have sought, in a broadly anthropological way, to study online culture. In comparison in this study we focused much more on the details of the interactions between players. Like any online world, *There* is an unusual conversational environment. We were particularly interested in how the resources (such as speech bubbles) were appropriated by users to support their interactions.

Accordingly, in studying *There* we collected videos of our experiences inside *There*, and our interactions with other players. These videos could be viewed repeatedly to look in detail at how certain interactions were carried out. To analyse the video data we broadly followed an interaction analysis approach [8]. This involved shared data analysis sessions where we observed key incidents recorded from our time in *There* over and over again. One of the key inspirations we used for this work was Sacks' observations on ordinary interaction [2], allowing us some useful comparison between *There* and real-world naturally occurring interactions and conversation.

3. Themes: Play and Sociability

As with other online games, *There* is a complex social environment where a host of different activities and actions take place. However, *There* is distinct in two ways. First, *There* is explicitly presented as a "playful" environment. Unlike games such as Everquest or Star Wars Empires, there is no overall goal to using *There*. The focus is much more on play—non-competitive activities which exploit features of the environment. As we will see, at times this can present problems to players in terms of finding what to do in *There*—actions are not directed by an overall goal. Second, with its focus on social interaction, *There* has been presented, and has some success as, an environment for socialising. Much of the business of *There* is chatting with others, "making new friends" as *There*'s own advertising hopefully presents it.

3.1. Play

As playful experiences, games present a number of challenges to a conventional HCI (human computer interaction) and design focus on optimising behaviour. With games, optimising the efficiency of behaviour makes little sense, since it is the *experience* rather than an end goal that is key. The pleasure of games can come from features such as aesthetics, narrative, or even the connection with a broader culture—topics difficult to address under the troika of effectiveness, efficiency and satisfaction. Yet research into HCI often suffers from a "negative utilitarianism" [9]—the goal in design is usually to minimize the "pain" that a system produces—for example, by replacing a difficult phone collaboration with a more effective interaction using a media-space (e.g. [10]). There is little concern with the "pleasure" that systems can produce.

With games, enjoyment comes from play. Recent research on games (the so called "ludology" approach) has argued that we need to move beyond considering the game itself (and in particular games' narratives) to consider how users "play" with games [11]. That is, studying users' behaviour and exploration of games rather than treating games as a static text. This makes the flow of gameplay, and user activities, into the key research object, rather than seeing games as a cultural object. In some senses, this is a very standard ethnographic move—from studying text to studying action. Within cultural studies it has some similarities with the moves made in the study of television [12], from considering television programs (such as newsbroadcasts) as texts in their own right, to understanding more how television programs were received and understood by viewers.

Focusing on play can move us from the (at times arbitrary) re-description of games to their treatment as social objects. Yet conceptually, play is a difficult activity to define and demarcate. Salen and Zimmenman [13], for example, describe play as "free movement within a more rigid structure" and although this captures the flexibility of play, it gives us very little purchase on the different forms which play might take in different settings. One approach could be to define play by differentiating it from activities done as part of paid employment or work. The problem with this definition is that it ignores much of the variety in work. While we use terms such as "at work" to account for our activity to others, this can hide many aspects of our activity [14]. For example, Roy's classic study of horseplay at work, "Banana Time" [15], shows how monotonous work is made tolerable through "mucking around" or playing at work. Although "work" contains considerable play, professionalism often relies upon *denying* this play since it can conflict with accounts of work as a serious, professional, well managed enterprise. In particular, defining work and play as opposed loses an important area for design—the space of playful work technologies [16].

An alternative approach is to build typologies of play, and a number of authors have attempted to describe play and its different features in that way (e.g. [17]). As ethnographers, however, we would be sceptical of the ability to define play *a priori*. Play can instead be seen as a varied yet similar set of practices which lead into but also conflict with each other—a family resemblance of activities. For example, research into slot machines [18] has shown how these games fit into individuals' lives by occupying time while waiting for others, and how this is not necessarily the same as play in the office, or play amongst children.

Despite these issues with conceptually describing play, and accepting it as a varied set of practices, a focus on play does give us traction on the variety of activities that have been very much ignored in design—the open, flowing and non-utilitarian. As in our other studies of leisure [19, 20], we start our enquiry by asking what leisure is in this specific case—what form does play take for the inhabitants of *There?* How is play organised?

3.2. Sociability

The second theme we address concerns the sociability of *There*. While studies of technology for the last twenty or so years have been deeply concerned with the social (e.g. [21]), research on technology has tended to ignore the more colloquial meaning of social—the experience and enjoyment of companionship with others. Even work on technologies such as instant messaging [22], while considering the pleasures of communication, has seldom addressed what effects IM has on friendship.

One of the few classic sociologists to consider the importance of friendship was Simmel [23]. In his paper about "the sociology of sociability", he argues for the value of "sociability", social experiences where their purpose is not external to that experience but rather *is* that experience itself. We engage in the company of others (in its purest form) when we engage for that company itself. Yet, while in these exchanges we are separated from everyday external conflicts and concerns, sociable encounters have within them a shadow of these conflicts and concerns. As Simmel [23, p. 261] puts it, they intrude in a "play form":

> All sociability is but a symbol of life, as it shows itself in the flow of a lightly amusing play; but, even so, a symbol of *life*, whose likeness is only so far alters as is required by the distance from it gained in the play, exactly as also the freest and most fantastic art, the furthest from all reality, nourishes itself from a deep and true relation to reality, if it is not to be empty and lying.

For Simmel, play, sociability and real life were entangled but at a distance—sociability was the play form of life where we could experience and experiment with the concerns of the world. While sociability never loses its lightness and justification in itself, it does not become sterile and removed from our lives. Indeed, it relies upon reflecting our concerns and battles to maintain a substance.

There are few sociologists who have taken Simmel's lead in considering sociability. While a number of recent ethnographic studies—particularly recent auto-ethnographic studies and research in the sociology of the emotions [24], have explored companionship, there is little discussion of the experience of sociability *per se*. Indeed, the study of friendship in sociology has been dominated by social network approaches which apply more quantitative methods to social relations—the research question becomes "how many friends, what type, and what support do they give you?" [25], rather than the questions of activity or action in friendship. Indeed, discussions of the activity involved in sociability are more prevalent in social psychological discussions of friendship. For example, Duck argues that it is through experiencing enjoyable activities together that we perform our friendships and relationships [26]. This is a view of social relationships as collective action. By choosing to do leisure activities with certain people they become our friends.

Computer gaming has always been a social affair—the original video game "Pong" was, after all, a two player game [13]. The massively multi-player environments of *There* and other games bring this even more to the fore. This is at the heart of what makes *There* enjoyable—it supports, in a crude though still enjoyable form, many of the ordinary activities that we engage in as part of our social lives. Conversation is perhaps the most obvious, but *There* also supports travelling, buying goods, exploring new places and playing conventional games (such as cards). While many of these activities are possible in single player games, multiplayer features make these activities sociable, supporting a community around these activities.

In the CVE community, this has been recognised in the notion of "social presence"—the sense in which users of a CVE feel that there are real people present in the online environment [27]. Alternatively, in the game literature the social ties formed in online games have also been a focus [28]. However, we would argue that both these approaches ignore the importance of *social action* as a focus of CVEs. That is to say, a key advance of CVEs is their support for doing things together with others over the internet. A focus on "social bonds" misses the importance of the shared activity together—such as chat, or interaction around objects, where we perform our friendships. Research on friendship underlies the importance and enjoyment of leisure activities carried out in the company of others. We would argue that it is this shared activity in itself which is pleasurable and a goal for players, rather than necessarily making new social bonds.

4. The Building Blocks of *There*

Play and sociability rely upon a number of interactional "building blocks" for their operation. Interaction around objects, talk, topic and identity, all work together to make play and sociability possible. As an environment distinct from the ordinary and unexplicated, *There* constantly foregrounds simple interactions such as answering a question, or leaving a conversation. Without successful interaction, it is difficult for play to reach its potential, and without conversation, sociability is near impossible. Each of these blocks contributes to the accomplishment of broader activities.

4.1. Interactions around Objects

The starting point to playing *There*, as with all online games, is mastering the interface. Indeed, there are a number of subtle design decisions which the developers of *There* have made which differentiate it from CVEs as conventionally studied and designed. First, and of perhaps surprising importance, is that the avatar control method uses the standard gaming control method, combining the use of the mouse (for direction of gaze) and keyboard (for movement). This dual control method allows users to quickly move their gaze with their mouse, and then slow that movement to rest their gaze on a particular object. As their gaze moves, this action is visible to other players by the turning of their avatar's head and body. By holding down a control key, a mouse pointer can also be made to appear, along with arrows over objects in the scene. These arrows can be clicked to display a menu to carry out different actions on these objects. *There* also places its default view above the head and some distance back. This increases the field of view allowing easier interaction with objects that are close to the avatar. It also goes some way to smoothing the interactions that take place in *There* between avatars, and in particular around objects.

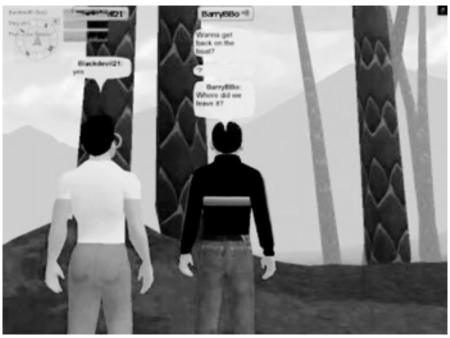
There supports some remarkably smooth, although not unproblematic, interactions around objects. In the example in figure 11-2, the first author (known as "Ba") is giving a tour to a new player ("Bo"). Earlier the users had landed a hoverboat to look round a forest and after deciding to get back in the boat, the users search around to locate the boat once again. After a minute of searching, Ba resorts to "retrieving" the boat—an in-game function that retrieves objects to where the player is. When he sees that Bo has noticed the retrieved boat, he climbs onboard, followed by Bo, and they take off.

While this interaction may not seem a source of enjoyment or a form of play in itself, interactions around objects are a key part of everyday activity [29]. Most real world games feature some sort of interaction around an object—such as a football. The players in this clip co-ordinate their activities (search, retrieval and discovery) without having to explicitly describe to each other what they are doing—both users search the local area, and are seen by each other to be searching. At the end of the clip, after beckoning Bo, Ba sees Bo walking towards the boat—he does not need further confirmation to jump on the boat. The orientation of each avatar in this example allows what each user sees to be predictable. A key part of this short interaction is how the object of the hoverboat exists not as a single user object but as an object shared between the two players.

When seen together, this short episode can be easily described—the two players look for a boat, retrieve it, get in and fly off. Yet a textual description does not do justice to the interaction, since it is much more than a textual experience. It both has unusual features which add to its enjoyability—we

1) Ba: Wana get back on the boat?

2) Bo: yes



3) Ba: Where did we leave it?
Bo: ’?

4) Bo: I don t know
(Ba retrieves the boat)

5) Ba: Hey Bo!

6) (*Ba gets on boat, Bo gets on boat and Ba takes off)*

Figure 11-2. Conversation around a boat.

don't usually fly around in boats—but it is also ordinary (looking for an object collaboratively).

Not all interactions around objects are as unproblematic in *There*. In particular, gaze and orientation is only grossly available so, for example, an avatar's orientation towards a particular seat on the boat would be unavailable (and on other occasions we observed this causes problems with negotiating who sits where). Yet this is an example of a straightforward successful collective social action, analogous to a real world situation where we drop our keys and find them again with a friend's help. That players can reliably interact around objects is a crucial building block in enabling more complex interactions in *There*.

Along with their use in the 3D environment, objects have a less collaborative existence in an e-bay style auction website. This takes the form of a separate 2D web interface that allows players to buy and sell objects. While this 2D web interface is perhaps easier to use than a 3D interface, the move is with some cost. As with real world shopping, aspects of identity creation are performed through the purchasing and display of objects, such as in the form of avatar clothing. Indeed, one of the main activities in *There* is making and selling clothes. Many players have large collections of outfits—one player we spoke to talked about having over one hundred outfits, with each outfit costing around US$ 2 on average. Unfortunately, the buying and selling interface is only a *single* user interface. The move from 3D to 2D interface has been at the cost of the collaborative features of the virtual environment. This makes one of the most important activities in *There*—shopping—a single user activity. Players who wish to shop online together need to constantly switch between the 2D and 3D environments and manage the connection with talk or chat.

The complex uses of objects in *There* builds on the actions shown here—interaction, transportation, trading—to produce complex compound activities such as skydiving into falling cars, or making short films which are recorded in-game. Importantly, objects in *There* have taken on a *social* function, in that they centre the collaborative activities within the game and produce opportunities for social action.

4.2. Chat

Along with interactions around objects, chat—what players say to each other—is an important part of *There*. As evidenced by the popularity of instant messaging, chat can be a valuable and enjoyable part of online interaction. While previous research has uncovered problems with text chat, such as the out of turn sequencing of turns [30], less attention has been given to the temporal aspects of text chat. That is, the timing of chat as it is typed. These aspects are particularly important in *There* since, unlike IRC or IM, text is displayed in speech bubbles which rise above users as they type each word [6]. In its design, *There* takes the use of speech bubbles in systems such as Habitat [31] or Comic Chat [32], and applies this to a 3D environment. Since sentences are shown as they are typed, the system affords some dynamic features of chat, with overlapping typing visible on screen as turns unfold simultaneously. For example, in this extract a group of users are arranging where to meet up in the UK (the square brackets show overlapping typing, the numbers in round brackets are notable pauses):

Jim: We have arranged Internet meets not for this program but for IRC and worlds.com the best place to (0.5 seconds) hold a (1 second) meet is (9 seconds) [Birmingham]

Sam: [Holland?]

One player suggests meeting in Birmingham, but his pause allows a second player to heckle him, completing the sentence with a joke destination. While in many ways this is a very normal interaction, heckling does need some subtlety in timing. Indeed, although text chat is in many ways a less rich experience than voice, small delays in voice conferencing can disrupt this sort of action [33].

More generally, overlapping talk affords replying to chat as it is produced, rather than waiting to the end of a turn, hastening conversation. Chat in *There* is thus unusual in that, rather than having only one speaker at a time, the system supports overlapping chat. In this extract a user replies while the previous turn is still unfolding:

Jo: Something as simple as town hall meetings should be a [requirement of each subcommittee thing]

Sue: [I agree or more vocal] webpage at there.com would be nice

Sue's reply agrees with Jo's suggestion before she has finished, displaying an assumption about how Jo's suggestion will finish. This display of chat as it is produced limits the time in which players are waiting for others' turns to end and mitigates some of the frustrations of slow typing:

Jo: Maybe they read but don't respond

Sue: Or can't [officially respond]

Gail: [well they used to] really respond and gleam info from the forums

In this extract following Sue's collaborative production of a sentence (as in [43, vol. I, p. 144]), Gail's "really respond" acts as a contrast right after Sue's "officially respond", even though her sentence started before Sue had finished. In this way unfolding overlapping text can be used as a live resource by conversation participants.

The slow movement of the speech bubbles up the screen can also work as a resource when entering a new conversation or walking past conversing groups—the bubbles allow users to see at a glance the previous turns in a conversation and thus quickly gain a rough concept of the topic of the conversation.

However, the use of speech bubbles does not come without cost. Speech bubbles occlude much of the environment as they are relatively large and each bubble may only contain a small segment of the conversation. To prevent overlapping speech bubbles, *There* also needs to put users into a specific "conversation" mode when it detects a stable conversation, moving users into positions where their speech bubbles will not overlap (as in figure 11-1). As with all modes, users need to issue an explicit command to leave. While this does solve problems with achieving positioning of avatars during conversation [27], it also puts limits on the number of participants in a conversation as well as the possibility of interaction around objects. Large groups also lead to problems with speech bubbles overlapping on screen. Simultaneous chat can also lose the sequential positioning of turns. This can cause confusion regarding who players are addressing their talk to, and players often need to use naming (such as in figure 11-1) to disambiguate turns.

4.3. Topic

The social interaction in chat rooms often disintegrates into "trivial, useless, sex-oriented" babble (Esther Dyson, quoted in the New York Times [34]).

Much of the conversation in *There* concerns *There* itself. For example, glitches (bugs or mistakes in the system) are common topics of conversation with users, often manipulating the glitches to produce bizarre actions that are otherwise impossible. In one "glitch" two users found themselves superimposed upon each other. This allowed the players to appear as a combination of two people, causing much play and conversation around the bizarre combined body. Glitches have become "local resources" for conversation, topics for discussion as much as problems or bugs in the system (see also [2, vol. 2, p.92].

This topicalisation of the game itself can also be seen in how gestures have moved from a *resource* for conversation to a *topic*. Gestures pervade conversation in *There*, and, as with emoticons [35], players have adopted these gestures to communicate lost aspects (e.g. emotions, illocutionary force) of face-to-face conversation. However, gestures have also become a topic of conversation in themselves. Users discuss the different gestures, and tutor each other in using them effectively. Support for gestures has even led users to exchange ways of producing avatar "dances". Using keyboard macro software, different, otherwise unrelated gestures, can be combined to make avatars "dance" on screen. Indeed, "dancing competitions" have even become popular where avatars compete for prizes for elaborate or inventive "dances".

In this way much of the chat in *There* is actually *about There*. As Sacks remarks, in some cultures there are topics that are inexhaustible as topics: they are "intrinsically rich, in the sense that whatever it is that members of that culture tend to talk about [. . .] they can talk about via that thing." [2, vol.1, p.178] In American youth culture, Sacks observes that topics such as "respect" can be discussed—often at great length—through talking instead about cars. This can also be seen in *There*:

Jo: The advisory board seem really closed [off to the untrained eye.]

Gail: [Well if you have any time Jo get on the IM with them let em] know

Jo: I'm not sure my little voice would make a difference=

Gail: =oh please woman [you are public beta 1 you people get mad respect from *There*]

Sue: [every little bit helps I don't think it is due to a lack of listening] more a slow development turn around.

While the topic here is ostensively the system itself, matters such as showing each other respect and organisational politics are covered through a conversation about the system. So while chat in *There* often concerns *There* itself, like with any group of enthusiasts, other topics are addressed through that topic.

4.4. Status and Identity

In the classic text based MUDs there is often a differentiation between "mortals" [33]—ordinary game players, and "immortals"—players with special abilities who are usually involved in running the game. In newer games such as Everquest or Star Wars Galaxies this has developed into "levelling up"—player characters gaining levels through completing in-game tasks, with different levels having different abilities.

While *There* does have levels for users, such as "socialisation" or "boarding" level, the advantages of higher levels are limited. This makes levelling up less of a goal in *There*. However, one clear visible divide that does exist is between trial users (who play for a limited time for free) and paying users. Trial users have no money to buy clothes, and so are usually adorned with a plain white t-shirt and slacks, in contrast to the colourful outfits of other players.

Indeed, appearance is a major focus in *There*, with avatar clothing used to present an identity in the world and differentiating oneself from others. While much of the research on virtual online worlds has argued that online identity is more flexible and transitory than real world identity [36], with users experimenting with changes such as trans-gendered identities, identity in *There* is a relatively stable and persistent phenomenon. The system itself presents barriers to maintaining multiple identities; users can only have one avatar per account, with a fixed gender and name which is decided before a user even enters the game. *There* also creates a standardized home page for each player (which can be optionally filled in), listing their interests, the clubs of which they are a member, a biography and a virtual and real photo. These features encourage the maintenance and creation of a single online identity, with a fixed gender.

This stability assists the formation of stable relationships—users create virtual relationships listing each other on their "buddy" list. Indeed, our experiences with identity manipulation in *There* was met with considerable resistance from other players. When we experimented with logging into other accounts, this generated hostility from our "buddies" who felt we were misleading them, and subverting the relationships they had built up. Even experimenting with changing our avatar's body shape was a source of complaint.

5. Discussion

Each of these aspects of *There* help to make *There* a reasonably successful environment. In particular, they contribute to making *There* both a playful environment and a sociable one. Through the "open-ness" of action—the ability to talk in different places, interaction around a range of different objects, *There* supports a wide range of activities. Designing for appropriation is a familiar recommendation for design [37], in that technology should be opened in its design sufficiently that new uses and applications can be discovered by users. Yet it is still a rare feature of computer games. Games are normally constrained around a linear plot based on involving the completion of a set number of tasks [38]. Yet in *There*, much of the play has come from the creation of new activities—heckling, joking, dancing, skydiving and the like. As Hughes' work on playground games [39] shows, children invent whole repertoires of new rules around traditional games, provided sufficient resources are available. For children in a playground this can be as simple as chalk markings, free time, and a ball (e.g. [40]). In *There*, these resources were its flexible support for chat and interaction around a range of objects.

Consequently, this means that *There* is as much a *platform* for gaming and play, as it is a specific game in itself. *There* provides resources for users, but it is users themselves who decide what form their play will take. *There* has been appropriated for play by its users in ways not intended by its designers, such as in how accidental glitches in the system have become opportunities for play and conversation. The social online environment, and fan websites around the game, all produce an environment supportive of appropriation and sharing of new activities.

There's open nature also gives it a more "playful" form than many other games. Sacks [2, vol 1, p. 475] remarks that in children's games mistakes are seldom serious, since the games are short and imaginary. Yet due to the persistent and competitive nature of online games, mistakes, such as losing objects, often have serious consequences. In *There*, however, the consequences of losing objects are minor, since missing objects can be easily "retrieved". In the situations discussed in figure 11-2, although the players lose their hoverboat, Ba can "retrieve" the boat easily. The lack of a competitive goal allows for a safe environment where users can play with different activities without conflict with their overall goals or broader enjoyment.

Yet this flexibility does come at some cost—*There* presents a problem to users in that they need to decide what to do next, rather than have a preset goal or task. Users must find their own activities in the game and negotiate their participation in group activities. Enjoying *There* thus involves some of the commitment and organisation of real world activities. While the open nature of *There* supports playful appropriation, it in turn requires commitment from its users. The key problem for users in *There* is "what will I do next?"

Much of the design effort in *There* has gone into its chat environment, and its rich support for talk and interaction. Of course, chat and talk are already massively popular aspects of the internet—seen in the popularity of chat rooms and message boards. Yet social action online—actions shared by others—has tended to be limited to text communication. While text is an obviously powerful medium, it is very different from face to face interaction. By offering a range of collective activities around objects, *There* expands the opportunity for *social action*. What is crucial here is a sense of acting together around objects—such as skydiving, trading, travelling around the world, playing cards and such, building up a shared history of collective experiences. For example, activities could be designed to require the co-ordination of groups of individuals, encouraging collaboration. Skydiving into a car in *There* involves co-ordination, since one player needs to drop the car from high up, while the second player jumps into the car as it falls. The pleasure of this activity comes in part from the difficulty in co-ordinating actions together. Work on other online games such as Everquest has also suggested the importance of shared activities in producing enjoyable experiences, for example Seay et al. [41] argue that a key feature of guilds is their shared activities in the game.

This suggests a key design goal for future environments will be supporting in game social activities. Economic transactions are one activity that could be enhanced in this way—with greater support for negotiation and interaction during the exchange of objects. Systems could also assist in producing a sense of group activity and belonging amongst users. For example, a system could automatically generate a history of what a group does together (such as in the form of a weblog), or of allocating a special game area to a particular group.

There is also unusual in that it encourages and supports interactions between strangers, something which has proved difficult to support in face-to-face settings [42]. One aspect of this is that presence in the environment shows an availability for talk. This suggests a lesson for the design of systems that require interaction between strangers: these systems should require some sort of commitment from participants to a setting which is differentiated from ordinary interaction. In that way a setting may work to produce an availability for interaction amongst participants. This also suggests that interaction between strangers will be harder around technologies which require little commitment to their use—such as, for example, museum exhibits [42], compared to those where there is a clear divide from ordinary interaction—such as in a tour group. More success may be had in settings where participants show their availability for interaction with others.

We started our discussion of sociability with Simmel's arguments concerning the form of pure sociability. Sociability in *There* follows very much Simmel's description, we interact in *There* not for the benefits to our career or goals, but for sociability in itself. Enjoyment is in the conversation and chat in itself. Yet, as Simmel warns, sociability stripped of its connection with the world, with the shadow of life, can become farce. This highlights a key problem with *There*—as with the problem of finding what to do next—its sociability lacks an orientation to an overall goal. There is no competition in *There*, at least at the level of other online games. *There* loses a close connection with the competitive nature of much of the real world. Its design challenge then is to find a way of connecting its rich support for sociability with a "deep and true relationship to the world" [23].

6. Conclusion

This chapter has focused on two themes from *There*. First, we have described the nature of play in *There*, and in particular how this depends upon the building blocks of smooth interaction around objects and chat between players. While these are not dramatic features, they enable many of the richer interactions in this game. The open design of *There*, and in particular how users have appropriated features of *There*, allows support for play in unpredictable ways.

Second, we discussed the importance of social action in *There*, and how players can carry out social activities with other players. By offering a range of collective activities around objects, *There* expands the opportunity for social action, and enjoyment for others. *There* also supports social action with strangers, something that has proved difficult in face-to-face settings.

Virtual environments, perhaps more than any other application, have suffered from hyperbolic accounts of their impact and development. Exaggerated claims of "replacing reality" place unrealistic goals onto systems, and may, as Fraser *et al.* argue, undermine the goal of designing compelling CVEs [43]. The nature of *There* and other virtual environments as games should be kept in mind—they do not replace social encounters or life in the real world. Although for some users they can become obsessions, on the whole they fulfil the traditional roles of enjoyment and enrichment, a role that games have traditionally played along with movies, music and fiction.

It is easy to dismiss games as of little significance, certainly when compared with more weighty topics or activities. Yet games occupy our attention for longer time than many of the traditional office applications studied by HCI. It is not only that games are an interesting new application of collaborative systems, but that in looking at games, HCI has the opportunity to consider new purposes for the systems we design, and new social benefits that they can produce.

References

1. Castronova, E. (2001). Virtual worlds: A first-hand account of market and society on the cyberian frontier. *CESifo Working Paper no. 618, 2001*. Available at http://papers.ssrn.com.
2. Sacks, H. (1995). *Lectures on Conversation: Vol 1 & 2*. Oxford: Basil Blackwell.
3. Boden, D. (1995). *The Business of Talk: Organizations in Action*. London: Blackwell.
4. Knorr-Certina, K. & Bruegger, U. (2003). Global microstructures: The interaction practices of financial markets. *American Journal of Sociology, 107*(4): 905–950.
5. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (1998). Fragmented interaction: Establishing mutual orientation in virtual environments. In *Proceedings of computer supported collaborative work (CSCW) 1998, Seattle, WA.*, New York: ACM Press, pp. 217–226.
6. Clanton, C. & Ventrella, J. (2003). Avatar centric communication in There. *Stanford seminar on people, computers and design*, Stanford University. Available at http://stanford-online.stanford.edu/murl/cs547/.
7. Hine, C. (2000). *Virtual Ethnography*. London, UK: Sage.
8. Heath, C. & Luff, P. (2000). *Technology in Action*. Cambridge: Cambridge University Press.
9. Sinnott-Armstrong, W. (2003). Consequentialism. In E.N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*, Stanford University Press. Available at http://plato.stanford.edu/archives/sum2003/entries/consequentialism/.
10. Gaver, W. (1992). The affordances of media spaces for collaboration. In *Proceedings of Computer Supported Cooperative Work, CSCW'92*. Toronto, Canada., ACM Press, pp. 17–24.
11. Frasca, G. (2003). Ludologists love stories, too: Notes from a debate that never took place. In M. Copier & J. Raessens (Eds.) *Proceedings of LevelUp 2003*, DIGRA.
12. Taylor, A.S. & Harper, R. (2003). Switching on to switch off. In R. Harper (Ed.), *Inside the Smart Home*. London: Springer-Verlag, pp. 115–126.
13. Salen, K. & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge, Mass.: MIT Press.
14. Abramis, D. (1990). Play in work: Childish hedonism or adult enthusiasm? *American Behavioral Scientist, 33*(3): 353–373.
15. Roy, D.F. (1960). Banana time: Job satisfaction and informal interaction. *Human Organization, 18*: 156–168.
16. Boucher, A. (2003). The disuptive office: Mechanised furniture to promote useful conflicts. In *Proceedings of 1AD: First International Conference on Appliance Design*, 6–8 May *Bristol, UK, no pp*.
17. Caillois, R. (1962). *Man, Play and Games*. London: Thames and Hudson.

18. Sutton-Smith, B. (2001). *The Ambiguity of Play*. Cambridge, MA: Harvard University Press.
19. Brown, B. & Chalmers, M. (2003). Tourism and mobile technology. In K. Kuutti, E.H. Karsten, G. Fitzpatrick, P. Dourish, & K. Schmidt (Eds.). In *Proceedings of the European Conference on Computer Supported Collaborative Work (CSCW) 2003, Helsinki, Finland*, Dordrecht: Kluwer Academic Press, pp. 335–355.
20. Brown, B., Geelhoed, E., & Sellen, A.J. (2001). The use of conventional and new music media: Implications for future technologies. In M. Hirose (Ed.), *Proceedings of Interact 2001*, Tokyo, Japan: IOS Press.
21. Button, G. (1993). *Technology in Working Order*. London: Routledge.
22. Nardi, B. & Whittaker, S. (2000). Interaction and outeraction: Instant messaging in action. In *Proceedings of computer supported collaborative work (CSCW) 2000*, Philadelphia, PA: ACM Press, pp. 79–88.
23. Simmel, G. & Hughes, E.C. (1949). The sociology of sociability. *The American Journal of Sociology, 55*(3): 254–261.
24. Ellis, C. (1995). *Final Negotiations: A Story of Love and Chronic Illness*. Philadelphia: Temple University Press.
25. Wellman, B. & Frank, K. (2001). Network capital in a multi-level world: Getting support from personal communities. In N. Lin, R. Burt, & K. Cook (Eds.), *Social Capital: Theory and Research*. Chicago: Aldine De Gruyter, pp. 233–273.
26. Duck, S. (1991). *Friends: The Psychology of Personal Relationships*. London: Harvester.
27. Becker, B. & Mark, G. (2002). Social conventions in computer mediated communication: A comparison of three online shared virtual environments. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. London: Springer, pp. 19–39.
28. Yee, N. (2002). Mosaic: Stories of digital lives and identities. Available at http://www.nickyee.com/mosaic/home.html
29. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (2000). Object-focused interaction in collaborative virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI), 7*(4): 477–509.
30. O'Neill, J. & Martin, D. (2003). Text chat in action. In *Proceedings of Group'03*, New York: ACM Press, no pp.
31. Morningstar, C. & Farmer, F.R. (1991). The lessons of Lucasfilm's Habitat. In M. Benedikt (Ed.), *Cyberspace: First Steps*. Cambridge, MA: MIT press, pp. 273–302.
32. Kurlander, D., Skelly, T. & Salesin, D. (1996). Comic chat. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, ACM Press, pp. 225–236.
33. Ruhleder, K. & Jordan, B. (2001). Co-constructing non-mutual realities: Delay-generated trouble in distributed interaction. *Computer Supported Cooperative Work, 10*(1): 113–138.
34. Marriott, M. (2003). Now playing: Reality without the downside, January 9, 2003. *The New York Times*, New York.
35. Guye-Vuillieme, A. & Capin, T.K. (1999). Non-verbal communication interface for collaborative virtual environments. *Virtual Reality Journal, 4*: 49–59.
36. Turkle, S. (1995). *Life on the Screen: Identity in the Age of the Internet*. New York: Simon and Schuster.
37. Randell, R. (2004). Accountable technology appropriation and use. In *Proceedings of the Third ACM NordiCHI conference*, New York: ACM Press, pp. 161–170.
38. Frasca, G. (2003). Sim sin city: Some thoughts about Grand Theft Auto 3. *Game Studies, 3*(2). Avaliable at http://gamestudies.org/0302/frasca/
39. Hughes, L. (1989). Beyond the rules of the game, why are rooie rules nice? In J. Evans (Ed.), *Children at Play Reader*. Geelong, Victoria: Deakin University Press, pp. 81–89.

40. Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32*: 1489–1522.
41. Seay, A.F., Jerome, W.J., Lee, K.S., & Kraut, R. (2003). Project Massive 1.0: Organizational commitment, sociability and extraversion in massively multiplayer online games. In M. Copier & J. Raessens (Eds.), *Proceedings of LevelUp 2003*, DIGRA.
42. Hindmarsh, J., Heath, C., vom-Lehn, D., & Cleverly, J. (2002). Creating assemblies: Aboard the ghost ship. In *Proceedings of computer supported collaborative work (CSCW) 2002*, New York: ACM Press, pp. 156–165.
43. Fraser, M., Glover, T., Vaghi, I., Benford, S., Greenhalgh, C., Hindmarsh, J., & Heath, C. (2000). Revealing the reality of collaborative virtual reality. In *Proceedings of the Third ACM Conference on Collaborative Virtual Environments (CVE 2000)*, ACM Press, pp. 29–37.

Chapter 12

DIGITAL DYSTOPIA: PLAYER CONTROL AND STRATEGIC INNOVATION IN THE SIMS ONLINE

Francis F. Steen, Mari Siân Davies, Brendesha Tynes, and Patricia M. Greenfield

1. Introduction

Around New Year 2004, a quiet year after the official launch, *The Sims Online* hit the global headlines. Disappointingly, the topic was not a celebration of how this "new virtual frontier," this "daring collective social experiment," had succeeded in bringing "our divided nation" together, as *Time* magazine had blithely prophesied [1]. Instead, the media reported that the online game had turned into a Biblical den of iniquity, a Sin City, a virtual Gomorrah—and that the whistleblower who bore witness to this had his account terminated [2–5]. Rivaling mafia organizations were practicing extortion and intimidation, pimps running brothels where underage girls provided sex for money, and con artists scamming newbies out of their start capital. Maxis, the company that designed and operated the game, maintained a relaxed laissez-fair policy of light and somewhat haphazard intervention. The first mafias and in-game brothels had been established already in the early days of beta-testing; they continued to operate unchecked. Pretend crime paid well and recruitment was good, leading to a rapidly mutating series of inventive scams targeting the inexperienced and the unwary.

It wasn't meant to be like this. Gordon Walton, one of the chief designers, had spoken in glowing terms of the game's potential to provide opportunities for better relations between people. While "all of our mass media positions us to believe our neighbors are psychopaths, cheating husbands, and just bad people," *The Sims Online* would short circuit our distrustful negative stereotypes [6]. Echoing McKenna & Bargh [7], who had found that relationships initiated online benefited from transcending the limitations of spatial proximity and

R. Schroeder and A.S. Axelsson (Eds.), Avatars at Work and Play, 247–273.

physical appearance, leaving more room for a creative identity construction that might act as a guide to a real self, Walton envisioned a game where people would "interact with others anonymously, have physical distance, and not be judged on your outward appearance. You interact with people on a pure intellectual and emotional level, devoid of all those filters." If his team did their job right, *The Sims Online* would feel like Disneyland [6].

Careful steps were taken to forestall "griefers", gamers who derive enjoyment from anti-social behavior. In line with the vision of a privately run amusement park as the contemporary image of the good society, The Sims Online would have no common areas or public property where griefers could harass people. The habitable landscape would be divided into lots, each of which would be owned and controlled by an individual gamer. If a player was giving others a hard time on your lot, you could throw him out or permanently ban him. If that seemed too much trouble, you could restrict access to a registry of friends. The programmers even anticipated the problem of one player blocking another's exit from a room; to avoid this they decided the trapped avatar would be able to escape by walking straight through its would-be captor. Implementing Lessig's (1999) dictum that on the Internet "code is law" [8] the makers of *The Sims Online (TSO)* sought to prevent crime by writing software code that made crime impossible.

These twin elements—the redemptive vision of the game's potential for creating friendships across barriers of distrust, and the proactive, structural legislation within the computer code itself to remove the threat of anti-social behavior—represent two of three main pillars of the game's utopian project. The third is more subtle: the freedom of the players themselves to create and to govern their own virtual society. Its spokesman is Will Wright himself, the lead designer and originator of The Sims family of games. Even before the game was released, he had begun imagining a self-governing society, with local governments and elections. But these features would not be built into the game. "All of this political stuff has to come from the bottom up," he insisted from the start. "We can't do it from the top down and dictate structure" [6]. A key inspiration was the architect Christopher W. Alexander's work on the emergence of communities. Alexander speaks of a "pattern language" that evolves organically from people's small acts. The patterns that define a town or community "can never be 'designed' or 'built' in one fell swoop—but patient piecemeal growth, designed in such a way that every individual act is always helping to create or generate these larger global patterns, will, slowly and surely, over the years, make a community . . ." [9]. Could *TSO* also be designed so that each individual act added up to a whole? After putting a basic framework in place and building some incentives into the game, it must be left to the players to establish their own political and civic cultures. According to Wright, "totally planned cities don't work. It's sort of like the Utopian society movement, where there were these guys who went off and started building planned cities. For the most part the cities were total failures" [6].

The uncertainty that such freedom entailed, however, was also a cause for worry: the team was painfully aware that nobody knew how the virtual world within the game would develop. Walton and Wright envisioned a society in which human relations were—paradoxically—more direct, because they transcended space, physical appearance, and entrenched identities, where crime was banished by the very architecture of the game, and where human freedom would express itself in self-organizing cultures. Could this vision be realized? Within the first year, this utopian dream was shattered, at least provisionally, by persistent reports that human relations had taken a turn towards an eerily familiar catalog of exploitative behaviors, where crime flourished and spread in the face of attempts to remove its very conditions of possibility, and where freedom had led not to democracy but to warring mafias. “How would people act if they were freed from real life laws and social constraints?” the BBC asked rhetorically, responding with reports of “child prostitution, rampant crime, mafia-controlled neighborhoods, shadowy self-declared governments struggling to maintain order and runaway inflation” [3]. “Hobbes in Cyberspace: Life in an online game world proves nasty, brutish, and short,” *Reason Magazine* concluded [10].

In the following, our goal is to understand the role of the unanticipated outbreak of crime in the dynamics of cultural creation in *TSO*. We start by examining the origins of the game in the single-player offline version of *The Sims*, focusing on the key design features that later become incorporated into the multiplayer online version. In the second section, we examine the three main dimensions of player involvement in the game: the immediate-term mechanical control of the avatar, the medium-term control of strategic moves, and the long-term control of the goals and ultimate meaning of the game. We argue that at all three levels, control is inadequate and tends to produce dissociations between player and avatar. In the third section we discuss the architectonic constraints on cultural development, the scope for rebellion against these constraints, and the significance of the strategy of crime. We end with an overall assessment of the game and the lessons to be learned from this vast and exorbitantly expensive experiment.

2. The Origins of The Sims Online

In order to understand the surprising dynamics of the online world of *TSO*, we propose to begin by examining its history. The official birthday of the game is the 17th of December 2002, when the game was first released to the general public. At that point, however, *The Sims Online* was already a bustling world, thanks to the activities of tens of thousands of beta-testers. Starting in mid-September, they had been invited to join the game for free, to uncover any show-stopping programming bugs, to ensure that the system scaled adequately, and to populate and settle the vast and virgin electronic landscapes. Paying users, Maxis reasoned, would prefer to join an existing world to constructing one from scratch. These beta-testers, in turn, encountered a world that was already highly structured, even if this structure would become actualized only by the gamers' own activities. Some of these structural features are odd: key design decisions of *The Sims Online* make little sense until one realizes they are the result of code inherited from the earlier, off-line version of the game, *The Sims*. We will suggest, in our analysis in sections two and three, covering player control and the higher-level cultural dynamics of the game, that the high degree of path dependence on code written for an off-line, single-player environment is a significant contributing factor to the problems that subsequently unfolded.

It was the Oakland fire of 1991 that provided the impetus to what was to become *The Sims*. Will Wright, a game designer at Maxis, lost his house in the blaze, and in the following months he became fascinated with the process by which his family gradually drew up the plans for the new house and its furnishings. Wright abstracted the sequential characteristics of this process and realized it could be turned into computer code. The program would supplement and extend the power of the imagination, simulating the construction and decoration of a house. He began to design *Home Tactics: The Experimental Domestic Simulator*. In 1993, after 2 years of in-house development, he presented a prototype to Maxis executives during a focus session. The idea was so unlike anything that had been done in computer games up to that point that they rejected it outright. For the next 4 years, the game was only worked on by Wright himself in his spare time [6]. When Maxis was acquired by Electronic Arts in 1997, however, Wright's reputation as the designer of *SimCity* earned him a blank cheque with his new bosses. *Home Tactics* got all the attention it needed and was released under the new name *The Sims* in January 2000. Once scorned, *The Sims* and its expansion packs became the most popular computer game of all times, selling more than 20 million units over the next 2 years [11].

At the core of *The Sims* is the act of building, landscaping, and furnishing a suburban house. The natural environment is an invariant subdivision: a rolling meadow by a brook, a road circling a dozen lots ready for construction. The game invites the player to build a house, or a succession of houses, and to move ready-made families into them. The building materials, the plants and trees for landscaping, and the furnishings and interior decorations are selected by the player from virtual shopping windows and assembled onto the three-dimensional canvas of the building site at the click of a mouse. A profusion of bricks and wood sidings, roof shingles, potted plants, wall papers, chairs, and kitchen utensils makes this part of the game straightforwardly enjoyable. The combinatorial possibilities are finite but astronomical; for all practical purposes infinite. You can move the walls around freely, put in a kitchen, a bedroom, a bath, a pool, a TV lounge, an exercise room. These practices build design rather than genuine engineering skills—according to a classic text of structural engineering, "a deep, intuitive appreciation for the inherent cussedness of materials and structures is one of the most valuable accomplishments" of an engineer [12]. In *TSO* the materials are flawless, each brick or chair the spit image of any other of its kind: there is a perfect and predictable match between the real and the ideal. The very act of design is an act of building. This simplification of reality allows the player to focus singlemindedly on the task of architectural design, maneuvering through a gratifyingly vast possibility space. The task of design is further aided by a series of elegant interface features: while you build and decorate, you can zoom in and out and view your progress from different angles. As you adopt a particular view, the walls in front of you becomes selectively transparent, allowing you to see the entire layout of the house. This aspect of the game bears the quality stamp of sustained iterative design [13] and surprises the player with delightful features.

The second layer of code in the game is the sims, or simulated people. They are the dolls that inhabit the houses built by the player. Controlled by artificial intelligence (AI), the sims are imbued with a rudimentary form of agency. They can walk from one location to another, perform a finite but expandable repertoire of tasks, express emotions, and communicate desires. The player interacts with the sims in a manner very different from that of the building materials. Where the latter move only when the player moves them by clicking and dragging her mouse, the sims behave as if they were alive, folding their arms impatiently if you leave them standing, as if chafing at the bit. While they do not actually move from place to place or perform tasks on their own initiative, you also cannot move them directly, by grabbing hold of them with your cursor and dragging them to a new location, as you can the building materials. For a sim to do something, you must give him an instruction to act. This is done by interacting with the sim's virtual environment. Objects in the sims world are endowed with Gibsonian affordances [14], or activities and behaviors that can involve the sim in some way. By right-clicking on the object, a contextual menu in the form of a cloud of affordances is displayed, and you make your choice. Clicking on the sim itself will give you a menu of possible behaviors valid for that particular location and circumstance, or you can click on a distant object and instruct the sim to interact with it. For instance, you can wake your sim up in the middle of his night and tell him to go for a swim in the pool simply by right-clicking on the pool and selecting "Swim". Like a truculent child in the face of a stern parent, the sim will get out of bed and stamp his foot on the floor repeatedly in a bout of displaced aggression and frustration, a fume of anger rising from his mind. He finds his way through the house and out to the pool on his own, without any player intervention: this shows his behavioral routines are pre-programmed and robotic. As he gets into the pool and begins to swim, the image of a bed floats in a thought bubble above his head, informing you that he is tired and needs to sleep. In the morning, having so rudely been deprived of sleep, he may become unresponsive to the instruction to go to work and require time to regain a functional level of comfort.

These responses create the impression, gratefully accepted in pretense, that the sim is a dynamic, homeostatic system, whose behavior is regulated by a simulation of causally connected mental and physical states such as hunger and the need for food. Dominating a sim's life is the physiological needs at the bottom of Maslow's pyramid [15]: the need to sleep, to eat and drink, to go to the restroom, to be comfortable, and—this is after all suburbia—to take regular showers. Each of these needs is tracked by a bar on the player's screen, turning slowly from green to red if it is neglected; taking care of your sim so that his basic needs are being met is called "greening." Higher needs have a rudimentary presence: if you don't provide your sim with regular social company, he will become despondent and slowly refuse to function. With some variations, this is as far as what Wright calls the "economy of motives" has been elaborated [16]. Sims do not live in a world where they need to worry about safety, and although in some versions of the game they can fall in love and marry, they do not complain if this doesn't happen to them. The built-in motives are of a ground-level nurturing kind. The sims have no aspirations to achieve social recognition through outstanding acts—to kill a dragon, to become president: they don't have a life project, a mission. Nor do they show an interest in learning and understanding the world they live in, though if you make them read employment-appropriate books, their skill levels increase and they get promoted at work. They do not seek to actualize themselves artistically or spiritually. They are, after all, dolls.

What are the design goals realized by creating sims with this particular circumscribed degree of autonomy? A sim has a fixed set of clearly defined needs, in part conveyed directly through emotional responses and through the display of the simplistic and generally predictable content of his mind, yet he is entirely incapable of taking care of even the simplest of these needs himself. Devoid of independent initiative, he relies on you, the player, to instruct him how to meet his own needs. For your benefit, a special "greening" panel tracks the precise progress of the sim's needs from moment to moment. If appropriately instructed, the sim will be healthy, energetic, and promptly carry out your commands; if his needs are not consistently met, however, he will drift towards a non-functional state, become unresponsive to your commands, and eventually die. This design creates—and is clearly aimed at creating—a distinctive dependency relation between the sim and the player. The function of this dimension of the game is to encourage the player to care for and nurture the sim. The game in effect provides opportunities for a kind of practice parenting, similar to playing with dolls, but with a more realistic feedback. By nurturing the sim, the player experiences the systematically differential consequences of proper care and neglect, and acquires skills relating to taking care of others. Care and nurture are behaviors with a deep natural history and they remain vitally important for any society, yet they had never before been the target of a sophisticated computer game. *The Sims* tapped into a vast and hitherto neglected audience of young girls who were left cold by the typical competitive or adversary shooter games favored by the boys [6].

Experienced players of *The Sims*, however, take the game far beyond the elementary task of keeping your sims green. As a player, you are in a position of directing the sims' lives as a dramaturg directs his actors, creating dramatic situations and developing extended narratives. Your capacity for absorbing information, for seeing connections, for opening up new possibilities, exceeds that of your charges by orders of magnitude. Like a god you can control whole neighborhoods of sims, staging their marriages, births, quarrels, reconciliations, and breakups. The writer Monique van den Berg's illustrated Sims diary [17], where a dozen families interact in intricate and often comic situations, provides good examples of the game's potential for staging complex narrative scenarios in richly elaborated environments. If you so choose, the godlike power of the player can even be used in the spirit of Gloucester's "As flies to wanton boys are we to th' gods, They kill us for their sport" [18]: you can command your sim with the terrible voice that the God of the Old Testament used to instruct Abraham to sacrifice Isaac, wall him into a closed room, electrocute him when he changes a light bulb, make him drown in the pool. If you adopt the sim's point of view, the situation is distinctly odd: in the middle of the night, with nobody in view, you are commanded to perform some meaningless act, strongly aversive to you, and you feel the anger welling up in you at the pointless imposition. But the command is as ineluctable as it is mysterious, and you are entirely incapable of disobeying. Such is the power of the player in *The Sims*.

In summary, the immensely successful offline, single-player version of *The Sims* developed out of a program designed to simulate the hands-on building and furnishing of a house, and was elaborated to incorporate robotic agents dependent for their basic welfare and continued existence on the constant nurturing behavior of the player. The game facilitated player psychologies ranging from an architect and interior decorator to a doting mother caring for her family and a movie director staging elaborate narratives. The secret of the game's phenomenal appeal lay in providing players with a godlike power to explore and innovate in two complementary and fully developed permutational spaces: that of building, landscaping, and decorating houses, mansions, and castles, and that of caring for, directing, and narrating the lives of simulated humans in evolving tangles of complex social relations. When *The Sims* was to be taken online, Electronic Arts sought to build on the success of the existing game and reuse tested code by preserving the core features of the game. Just like *The Sims*, *The Sims Online* would present a suburban housing division, albeit a much larger one, where players constructed and furnished houses on lots, using the mature code inherited from the offline game. Just as in *The Sims*, the simulated people within *The Sims Online* would respond robotically to instructions selected by the player from clouds of affordances surrounding in-game objects. The secret of the game's phenomenal appeal lay in providing players with a godlike power to explore and innovate in two complementary and fully developed permutational spaces: that of building, landscaping, and decorating houses, mansions, and castles, and that of caring for, directing, and narrating the lives of simulated humans in evolving tangles of complex social relations. *The Sims* would go online.

Yet massively multiplayer online games rely on a very different dynamic in the relationship between players and on-screen characters. Key to a multiplayer environment is that each player is represented in the virtual world by a single character, his or her avatar. A relatively tractable dimension of this change is that the game now needs to have a unique avatar for each player. Users typically want to control their avatar representation and have input into its design, yet the need to download tens or hundreds of thousands of unique avatar designs onto each player's computer would cause severe network and storage problems [19]. *The Sims Online* solves this problem by providing more than a hundred different heads and bodies with outfits for the user to mix and match. This combinatorial space is sufficiently large to minimize the risk of two avatars looking exactly alike, yet avoids storage problems, as all avatars can be represented by a number referring to the graphics in the selection. While this solution provides sufficient differentiation between avatars, it does not allow players to contribute their own graphics, thus limiting the work the representation can do in defining and channeling a particular identity.

More troublesome than avatar design was the requirement of *The Sims Online* that the player identify with his or her avatar. The mantra of the development group was that in TSO, "the sims are real" [16]. Massively multiplayer online games like *EverQuest* and *Ultima Online* had shown that players became emotionally engaged in their own avatar and formed strong bonds with others through their on-screen representations. Such identification would be absurd and inappropriate in *The Sims*, where each simulated human being was endowed with a carefully circumscribed autonomy, expressed in robot behavior, designed to be endearing and to elicit a caring and nurturing stance. As Will Wright himself put it during the alpha phase of the development process, "That's never been an issue in any of my games before. Most of the time I'm dealing with little simulated AI people that pee on the floor all the time" [6]. In *The Sims*, the player was in charge of the entire virtual world and all its inhabitants; in *The Sims Online*, tens and hundreds of thousands of players would interact with each other. If the player was god in *The Sims*, what would be his role in *The Sims Online?*

The conversion of *The Sims* to *The Sims Online*, starting in 2000, was an enormously complex and expensive undertaking, with a development staff of a hundred programmers, three million lines of code, and a rumored budget of $25 million [6]. It was undertaken at breakneck speed, and achieved its ambitious goal of shipping in time for the Christmas season 2002. Subscription numbers shot up instantly to 40,000 and by New Year to 80,000, continuing to mount until topping out around 105,000 in June 2003, far below the projected numbers [20]. In the following section, we examine two players' responses to the game over the course of the first year, focusing on identification and player control. How did the experience of player control change as the core features of *The Sims* were ported to a radically different, multiplayer environment?

3. Player Control in The Sims Online

When *The Sims Online* was released in December 2002, Patricia Greenfield, with the collaboration of Brendesha Tynes and the research team at Children's Digital Media Center at UCLA, had already recruited two players and initiated data-collection. The intention was to treat the game as a laboratory in which the spontaneous emergence and evolution of culture could be documented. By starting participants at the game's beginning, Greenfield hoped to be able to observe not just the adaptation to an existing culture, but the actual creation of a culture from scratch. Steen and Davies joined the project in early 2004, and when the assembled team began the task of examining the collected data, it became clear that we needed to change the research focus. In the participants' diaries and captured gameplay we found little or no evidence of the players creating a shared cultural world of meanings, norms, activities, and physical environments through processes of social interaction and communication [21, 22]. The virtual inhabitants of Alphaville, the first city within *The Sims Online*, appeared to have surprisingly little to say to each other, and the game did not provide our study participants with opportunities or tools to engage in sustained collaborative cultural creation. As we observed their interest in the game slowly fade, we shifted our focus to investigate what had gone wrong. In the following, we present a summary of our findings, with selections from the data and some new analytical points; for the full story, see Steen, Greenfield, Davies, and Tynes [22].

A massively multiplayer online role-playing game (MMORPG) can provide its participants with control along a spectrum of time horizons, ranging from the immediate to the long term (figure 12-1).

At the far left of the spectrum we find the immediate-term mechanical control of the avatar. Is there a one-to-one correspondence between the player's manipulation of the game controller and the behavior of the on-screen avatar? Such synchronous control is a mandatory feature of fast-paced shooter games; Kirk [23] argues it is critical for a low-level, physiologically driven player-avatar identification that powerfully enhances game immersion. The code that *The Sims Online* inherited from *The Sims* handles avatar behavior through artificial intelligence, thus removing a base level of support for an immersive experience. In contrast to other MMORPGs, such as *There* [24], you cannot directly control your avatar—you cannot, for instance, use the keyboard or

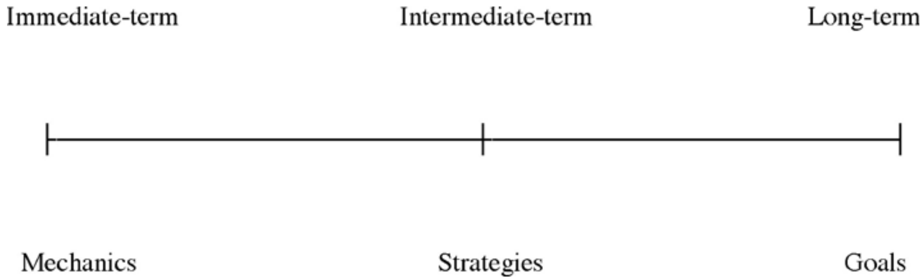


Figure 12-1. The spectrum of player control in a MMORPG.

mouse to turn your avatar to face a certain direction, move its arms, or move it step by step in some direction. All control is done through menu interaction, where you give instructions the avatar acts out.

In the diaries and videotaped gameplay of our first study participant, KM, a 23-year-old female graduate student, we found that the lack of synchronous avatar control generated a series of dissociations between the players and their avatars. KM described herself as a recovered addict of the off-line version of *The Sims* and was thus familiar and comfortable with the sims' robotic behavior in the single-player environment. In the spirit of the new multiplayer environment of *TSO*, she began play on 30 December 2002 by creating an avatar to represent herself, an *alter ego*: "I created this character based on myself. It took me a while to go through all of the hair and outfits to pick one that I thought resembled me." (KM, p. 16). In short order, she was confronted with what for all of us was the most surprising feature of *TSO*: the low level of social interaction and conversation. KM began by identifying the avatar she encounters with a real person: "I said 'hi' to the other person playing chess, but little conversation happened. I noticed that no one in this house was talking everyone was just earning skill points." (KM, p. 17). Skill points can be earned through activities such as standing in front of a mirror ("practicing charisma"), playing chess, and cooking. In order to encourage socializing, a key novel feature engineered into the game is that skills are acquired more rapidly if practiced in the company of other avatars. As with other avatar activities in *TSO*, the behavior itself is robotic, and skills points take hours to build. KM's suspicions were soon aroused: "After about 10 minutes of playing chess, and seeing how long it took for the skill meter to go up I could see why no one was talking—probably no one was there!" (KM, p. 19). The substance of this insight is a dissociation between the avatar and the player: the presence of one frequently does not entail the presence of the other. In the absence of synchronous control, avatars in the online chat rooms of *TSO* can be intermittently attended; it is common practice to go "afk" or "away from keyboard" and leave your avatar, robotic slave that it is, stay behind and accumulate points. An unfortunate downstream consequence of this behavior is that players begin to see the presence of other sims in purely utilitarian terms, as an opportunity for earning skill points more rapidly, rather than as a chance to meet people and socialize. This second dissociation separates your own needs for real human social contact from the avatar's more pedestrian need for the proximity of other avatars, attended or not.

The absence of synchronous control also created anomalies in the immediate-term mechanics of social interaction. We analyzed a recorded conversation at Lucky Luc's Slots, where AJ, the proprietor's roommate, gives KM instructions on how to play a particular gambling game. During most of this conversation, the two avatars have their backs to each other, yet it would be incorrect to infer from this that they had lost interest in the conversation. At one point, AJ refers to a game score result that she could not have seen, as it appeared behind her. These incongruities open up a third dissociation, that between the panoptic perspective of the player and the embedded, in-game perspective of the avatar.

Finally, robotic behavior can generate dissociations of intentional states. In one session of recorded game play, we observed KM giving instructions to her avatar to play the guitar. Once instructed, the avatar persists in the task until it is completed. Other sims talk of leaving; half in jest, as she isn't actually playing, she says, "I am going to go too, can't stand listening to my own music." Zooming out of the building, she ends up with a bird's eye view of Alphaville, from which you can see all the different properties and decide where to go next. Her own intentions have dissociated from those of the avatar, who is left behind playing the guitar (see figure 12-2).

In KM's diaries, we begin to understand why so little conversation takes place in *The Sims Online*: typical gameplay is characterized by long absences from the keyboard, as the robotic work of skilling is itself experienced as boring. Our analysis of her game play captured on video indicates that this boredom forms part of a series of dissociations between player and avatar, which act cumulatively to weaken the bond of identification. These dissociations have a primary cause: the absence of immediate-term, synchronous player control.

At the other extreme of the spectrum of player control is the long-term goal or overall purpose of the game. In the single-player, offline version of the game, *The Sims*, the player is comfortably in charge of the purpose of the game, whether it is to build castles in the air (there's a trick to getting them up there), to raise a sim family, or to start a gay karaoke bar [17]. In keeping with the hallmark of Will Wright's game philosophy, the promise of *The Sims Online* was that it, too, would allow the gamer to formulate his or her own goals within the game.

TSO, however, had introduced something co-designer Chris Trottier called "a real secure economy" [6]. In *The Sims*, the economy is insecure. If you enjoy the constraint, you can abide by the rules and hold off on that swimming pool gazebo until your sim has boned up on his mechanical skills,

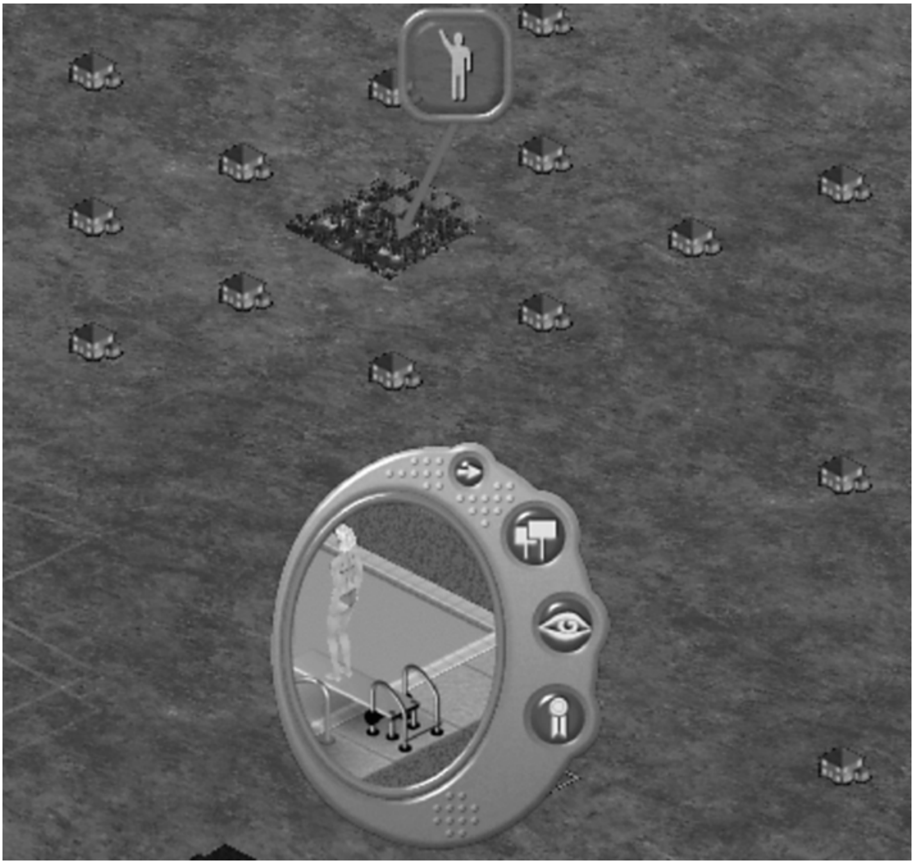


Figure 12-2. Players can leave their avatars behind and—like disembodied spirits—view the world from above. In the meantime, the avatar continues to engage in activities as instructed.

developed some friendships, and can start earning some real money. But if you don't, a Google search for "sims cheat codes" will pull up 283,000 hits. These cheat codes, which are built into the game, allow a player to get unlimited resources with a few keystrokes. The insecure economy of *The Sims* allows players the flexibility of setting their own goals. The secure economy of *The Sims Online*, in contrast, introduces a resource scarcity into the game that tends to swamp all other goals, as illustrated in the following analysis.

Our second study participant was a 28-year-old man (SH), a free-lance producer and part-time actor. On the 19th of January 2003, after spending the first play sessions exploring the game, he creates an avatar, "Sammar", to represent himself:

> I chose to develop this character because he is the closest thing to my alter ego. I needed an outlet for that ego in order to help myself in real day to day life. I'm

> hoping that I'll be able to learn from my other self and take those characteristics that I feel I lack and forge them into my real life. (SH, p. 9)

His entry suggests a very rich vision of the game's potential. By creating a virtual self, he imagines he will be able to explore and to cultivate modes of being and responding to the world that he can subsequently incorporate into his own life in a selective manner. This is a goal he sets himself, demonstrating his confidence in being able to control the overall purpose of the game. After recounting Sammar's first day of play, visiting places, getting a roommate, and earning money—"simoleans"—making pizza, he states a subordinate and interim goal in the game, now projected onto his avatar:

> His main goal at present is to make enough money to build a party pad by the beach. (SH, p. 11).

His entry contains the key elements of a narrative self: a goal (a party pad), an obstacle preventing him from realizing the goal immediately (not enough money), and a strategy for overcoming the obstacle (making money). The goal itself is envisaged in social terms: his house will be constructed in an attractive location and provide a venue for himself and other players to have a good time together. As he continues playing the game, he maintains a strongly positive disposition:

> A lot of the people I've visited at their properties have been exceptionally nice. I imagine it has to do with their visitor bonus. The people I met in the pizza place are not nearly as friendly. It's amazing what greed will do. (SH, p. 13)

The effect of the game's incentive structure on the other players is beginning to transpire. Property owners are paid for the time others spend on their lots; his warm appreciation of their welcoming behavior is qualified by the suspicion that they, just like the unpleasant pizza makers, are driven by greed. A few days later he reports:

> Sammar is feeling accepted in this community. He is still figuring out the finer details but it's coming along well. He aspires to make his skills at their peak and make as much money as possible." (SH, p. 19)

SH continues to view *TSO* primarily as a community, a place regulated by norms and common meanings, in which he can feel accepted and welcomed. At the same time, as he begins to master the rules of the game, his visionary goals narrow. The next day, we find him for the first time thinking of friendships in instrumental terms: "He's building a friendship base that's making him money and skill." (SH, p. 23). When SH returns to the game on 18 March 2003, he reports:

> "Sammar" has built his skill levels, mostly mechanical & logical, and is making a decent amount of money making gnomes. He has a home now and is in the process of building it up to be a place where other sims can come to relax and make money. (SH, p. 25)

He now shows little emotional involvement with his avatar, whose activities in this session are directed not towards forming relationships, but on building skills and making money. However, he sees these activities as a temporary means to a more attractive goal of building a house. The purpose of this house is still to provide a place for others, but he no longer imagines they will come to party. Instead, they will come to his house to hang out, and to make money. His basic motivation remains altruistic:

> My characters main goal at present is to be a viable and successful character who can help other Sims in their money and skill earning endeavors (SH, pp. 25 & 27).

At the same time, his interest in *The Sims Online* is starting to flag. The game, he writes, is "somewhat boring" and does not facilitate the social contact he came looking for:

> The game would be more condusive to chatting if email were accessible while playing to swap pics and personal info. A real possibility of meeting these people off line would get the place buzzing. (SH, p. 39)

Avatar encounters, apparently, do not have the emotional and intellectual qualities of real encounters. His suggestions of introducing e-mail, swapping pictures, and meeting people off-line indicates that he experiences the on-screen characters as poor representatives of the players' social agency: identification has become unattractive and the channeling role of the avatars is failing.

When he decides to give the game another try in April, he abandons Sammar and begins Freakstick, a skeleton-like character to "express my off the wall personality" (p. 37). He attempts to reformulate his strategies:

> Now that I've learned the main tricks and tips in succeeding in The Sims, I have a new way of going about things. I plan on amassing large amounts of mechanical and logical skill. Those skills have the greatest amount of financial profitability with the least amount of constant attention. (SH, pp. 37–39)

He soon learns, however, that *TSO* provides very limited opportunities for rapid progress in the game, gets bored, and plays infrequently. On 27 July 2003 he goes on for a brief session to build skills. There were few people online, so the effort didn't pay off as much as he had hoped. "Maybe there will be more people the next time I log on," he comments (SH, p. 45). He is now speaking of people as a simple means to speed up the gaining of skill points; he is no longer interested in socializing or meeting friends. The incentives of the game appears to have ground sociality out of him.

On 4 August 2003, his goal is subtly reconceptualized, even as he represents it to himself as unchanged:

> My ultimate goal, still, is to gain enough skill and money to build the ultimate house where I won't have to work at making money. Rather I earn money by collecting the revenues given to me by the Sims for visitors coming to my house. Also I will get residuals for every dollar that my guests make. (SH, pp. 47, 49)

Shortly after, complaining that the "time it takes to build skill is a little overwhelming, not [to] mention boring" (26 August 2003, p. 51), the record trails off and he abandons the game.

From start to finish, SH displays an active desire to define the long term, overall purpose as well as the intermediate goals of his gameplay. Step by step, he lets go of his own goals and adapts to the goals built into the game. He relinquishes the desire to become a better person by trying out new modes of relationship within the game, he gives up on the idea of helping others succeed and of becoming appreciated by a community of friends. Unable to remain motivated by the time-consuming and mindless activities of skilling and making money, he is drawn towards the ultimate goal provided by the structure of the game: to become a "sim lord" (SH, p. 41) and live off the labor of others. At that point, human relationships have deteriorated beyond the level of instrumentality to something reminiscent of exploitation. An early review spelled out the emerging culture of the game:

> Since a player can earn money simply by enticing other players to congregate on their property, and because all the other players truly want to do is earn money, the object of the game is reduced to building—not a "house" in which your Sim will live, but a labor camp in which other Sims will come to earn money. Providing beds, showers, food and a pool table persuades your guests to stay longer and spend more of the money they are earning, owing their souls to the company store, so to speak, and never truly needing a place of their own. The result is a "city" in which nearly every house is a sweatshop [25].

What *The Sims Online* failed to do, then, was to provide the players with the tools to control the overall purpose of the game. From the early interviews, it is clear that Wright and his development group fully intended to build this freedom into the game. When large numbers of beta-testing citisims started spending their time making pizza, the developers were distressed. "A few weeks ago, we thought we'd have Disney World. But right now, everyone is just making pizza," Trottier lamented. Wright worried that too many people were chasing money in the game—"we might start to lose the creative players" [6]. However, no satisfactory solution appears to have been found (see figure 12-3).

The Sims online involved two major design challenges, one social and one psychological. The psychological challenge was the shift from a godlike player

Figure 12-3. Avatars practicing “charisma” before mirrors in a lavish interior. Clicking on another avatar gives you a “cloud” of interactive options; the sequence of tasks to be carried out by your avatar accumulates in the upper left corner. A needs panel monitors your avatar’s state along eight dimensions.

to an embedded avatar. In KM’s experience, we see some of the player-avatar dissociations that make this transition unsatisfactory. On the social front, the challenge was to establish some kind of resource scarcity. Kollock’s seminal “Design principles for online communities” argues that scarcity is an important dimension of a vibrant online community, not just to keep things lively, but because “moderate amounts of risk are required for the development of trust . . . and encourage the formation of groups and clubs as a way of managing that risk (or exploiting it, in the case of a guild of thieves)” [26]. Using only a “secure economy” to create scarcity, however, backfired. It set up a dominant incentive gradient that funneled most people’s energies into mindless and boring money-making routines, destroying the fun and creativity of the game.

The effect of the design choices made in these two major transitions was to curtail player control at both extremes of the spectrum—the immediate-term mechanics of avatar movement, as well as the long-term goals and purposes of the game. There remained, however, the central portion: player control of intermediate-term strategic innovation. This area provided the developers with ample possibilities of expansion, but they were not exploited. A rich set of tools and opportunities for medium-term, strategic innovation would have given players something to focus on, compensating in part for the loss of control at the extremes of the spectrum. We see in SH's diaries a mounting frustration at the ineffectiveness of his actions and his inability to come up with strategic shortcuts, indicating a strategic deficit. Player control at the intermediate, strategic level involves the ability to formulate complex and clever sequences of moves that help further the player's goals in effective and original ways. In *TSO*, plodding along a number of roughly equivalent routes was the only legitimate way forward.

When the players of *TSO* found, to everyone's surprise, that they had nothing to talk about, the explanation may lie in the glaring absence of strategic opportunities. In most multiplayer online games, it is strategy development that provides the gamers with something worth communicating about. Indeed, a distressing effect of the failure to develop the central, strategic portion of the spectrum of player control is that players had no way of becoming uniquely valuable to each other. Human beings are irreplaceable resources for each other in part because each one of us gathers our own information, adopts an idiosyncratic perspective, and develops our own strategies—complex sequences of actions that reliably achieve improbable results. In a society rich in strategic opportunities, information becomes the critical scarce resource. In the next section, we examine how a relatively small number of innovative players found a way of mining this middle of the control spectrum, with dramatic effects.

4. Strategies and Limits of Cultural Evolution in The Sims Online

Full-scale social experiments are expensive. In order to put a new model of social organization to a realistic test, you need tens of thousands of subjects over a period of years. Finding a suitable and willing population is almost impossible; funding the enterprise is prohibitive; gaining human subjects approval is impracticable; the logistics a nightmare. In the initial phases, you will need to insulate your state to some degree from its surroundings, to give it time to develop its own economic practices and civic institutions. The continuity of history must be broken for your ideas to be implemented and given full play. If your experiment goes awry and your ideas turn out to produce a monstrous society, you will be held personally responsible, exposed to the murderous ire of your own captive citizens. Until recently, short of a full-scale revolution or a military invasion, the only viable alternative available to social visionaries was the thought experiment. Where past ages were constrained to choose either the expensive realism of war, colonization, and revolution, or the cheap but fallible instrument of utopian thought experiments, we now have a technology that permits us to create imaginary worlds and to populate them with real people.

What was simulated in the mind for millennia can now be simulated in public, in vast simulated online cities.

The utopian question is not "What emerges from a state of nature?" [5]. A utopia is a social experiment that involves the explicit and deliberate manipulation of certain structural parameters in establishing and running the imagined state. The intricate thought experiments of Plato's *Republic*, More's *Utopia*, and Butler's *Erewhon* exploit the imagination as a cheap and readily available vehicle to explore the possibility space of social arrangements. They ask the question, "How do the choices we make in setting up the structure of society affect the behavior of individuals and thus its course of history?" As computer-mediated active worlds proliferate, this question is pivotal not only for understanding the evolutionary trajectory of a particular multiplayer online game, but also for being able to develop the kinds of worlds we want. The utopian question has become reformulated as a challenge of software design: in an active world, "what features must the environment have in order to enable particular types of social interaction?" [19, p. 8]. The evolution of culture in a massively multiplayer online world cannot develop from scratch; it must necessarily emerge from the complex dynamics of interactions between the programmers, the structural framework of the game they create, and the gaming activities of the players, which may take place both within the game itself and—as we shall see in a moment—extend beyond it. A game is not a clean-room implementation of a new society, untainted by preexisting values, beliefs, and conventions, but imports these dimensions, explicitly or implicitly.

In the case of *The Sims Online*, we begin with a richly featured environment, structured on multiple levels. We argued in section one that the secret of *The Sims*' phenomenal appeal lay in providing players with a godlike power to explore and innovate in two complementary possibility spaces: that of building, landscaping, and decorating houses, and that of caring for and directing the lives of simulated humans in evolving tangles of social relations. In section two we argued that the shift to *TSO* involved two major innovations, on the opposite ends of the spectrum of player control: one social and one psychological. The transition from an insecure to a secure economy effectively barred most players from the joys of building, reducing their game play to sweatshop labor. Adding insult to injury, the transition from the dollhouse to the avatar model barred players from the godlike power to control and stage the lives of the sims within the game. These architectural decisions were in large part forced upon the developers: going online meant that you had to remove the godlike power of the player. In this section, we will argue that *TSO's* critical shortcoming was that these constraints on the extremities of the spectrum of player control were not compensated for by new powers in the middle: by new strategic opportunities.

The co-creation of culture within a multiplayer online game is not necessarily cooperative; the goals of developers and gamers may be only partly overlapping. When the implementation of a game fails to provide attractive avenues to success, inventive players may seize the strategic initiative and attempt to move the game in new and unanticipated directions. The game designers' vision may be incoherent or produce unanticipated results, and the developers may falter in the implementation—in May of 2003, EA spokesman Jeff Brown acknowledged that "The people who make The Sims [Online] believe that its execution isn't what it should have been when it was launched" [27]. As we saw in section two, casual players, working with the tools provided to them within the game, met with repeated frustration and boredom in trying to achieve their goals. Creative players, on the other hand, may extend the boundaries of the game, drawing in resources that supplement the game, and find ways to compensate for its weaknesses.

The absence of a mechanism to import customized graphics into the game, for instance, cut off an important dimension of user creativity and constrained the elaboration of in-game identities. Some of the players responded by creating off-game web sites devoted to their *TSO* avatars, in which they were able to utilize their own art work and draw freely on cultural references that added resonance and power to their role play. A striking example was www.thesimsmafia.com by JC Soprano, played by the 25-year-old Sacramento native Jeremy Chase. It sported dramatic flash animations on the mafia theme and a detailed list of available in-game services, from prostitution and gambling to debt collection and assassinations—"Moe Green Specials" as he called them. Another group of players established the Rose Bush Gardens neighborhood in Alphaville devoted to "Bondage, Discipline, and Sadomasochism," amplifying this theme on the external Black Rose Castle Learning Center web site, with detailed instructions by the avatar Lady Julianna on how to be a dominant without being obnoxious, and how to participate in pony submissive races within the game. Urizenus, played by University of Michigan philosophy professor Peter Ludlow, created *The Alphaville Herald*, an online newspaper covering in-game events. By conducting in-character interviews outside of the game, he raised the profile of the inventive players, created a wider audience for their role play, and provided them and other players with an opportunity to reflect on the significance and impact of their in-game behavior. In these and other ways, the limited opportunities for creating an arresting identity within the game were transcended, as the gamers recruited a range of off-game resources to reinforce and heighten their own game experience.

Expanding the game beyond the confines of *TSO* proper also increased the effectiveness of the in-game character. To circumvent Maxis' Online Community Representatives, the Soprano mafia family members chatted with other players on *Yahoo Instant Messaging (IM)* rather than using the game's own chat feature, logged by the company. This allowed JC and his recruits to refine and develop a string of strategies for mimicking mafia-type activities within the game, without giving them away to the developers. As presented on his web site, these strategies were mysterious, vaguely menacing, and clearly fun. Maxis tried to keep up: "Most of the behavior described in stories about these 'mobs' is no longer possible, actually; we've been improving the game with frequent updates," associate director Kyle Brinx claimed in June 2003 [28], but six months later JC confidently proclaimed, "The city is mine . . . I hate to say it, but I got the juice in AV [Alphaville] and have for awhile" [29]. By figuring out ways to "warp" the built-in features of the game for mafia role-play, he had in effect seized the initiative from the developers and taken a strategic level of control of the game.

The blatantly anti-social character of the mafia role, adopted within a game that set out to simulate a real society, raised the question of how such a virtual society can be policed and governed. By creating a publicly available representation of his in-game activities, JC Soprano made it easy for the non-playing world to participate in contemplating this question. In the summer of 2003, the *Associated Press* did a story on "Sex, mob hits: Sims tests virtual morals" [30], reporting on the exploits of two rivaling mafia groups, the Sopranos and the Sims Shadow Government (SSG). CNN followed up a month later with live coverage [31]. In these interviews, both mafia families defended their activities by claiming that they dispensed a rough justice to discourage griefers. EA simply didn't respond effectively to protect innocent players from abuse. "Griefers . . . find and utilize loop-holes within the game," Jennifer Mathieson of the SSG said, "and it happens very, very quickly. So what we do, we just fight back. We use the same tactics . . . against them." [31]. Jennifer and her husband, who jointly played the avatar Mia Wallace, recounted that they had "ransacked apartments, sent out their 'troops' to urinate on others' lawns, and once drove another player from the game" [30]. By extending role-play into the media, these gamers gave people a new reason to play the game: to experience and to explore the ethical dimensions of online worlds. The "darker side of Sims life," Wright himself conceded, "makes the game more interesting. It is pretty playful and harmless"—and the governance of virtual communities "is something our society is grappling with" [32].

At the same time, *The Sims Online* was revealing its potential as a breeding ground for a wide range of humiliating, anti-social, and exploitative behaviors. In a BBC interview, Ludlow later explained that his make-believe newspaper, created off-game but edited in-character, was founded to document "the emergence of economic, social and political structures in the game" [3]. As events unfolded, *The Alphaville Herald* turned out to be perfectly positioned to become the media hub of the seething underworld of TSO. As Urizenus, Ludlow covered the rise of the mafia families and their increasingly hardball tactics, such as harassing a sim by sending her a new roommate, and then asking him to tear down her house. The interviews show that the players invested considerable emotion in the conflicts. In early December, he conducted a series of instant-messaging interviews with Evangeline, a cyber-prostitute, who had set up a brothel early on in the game as a strategy to avoid the boredom of skilling and working. Describing her business in graphic terms, Evangeline let it slip that she was underage herself, and claimed that several of the girls that provided "sexual services" within the game were also minors. Using the proceeds from the prostitution racket, she had purchased the property at the top of the game's welcome list, naming it Free Money for Newbies. Here she cheated newcomers out of their money, humiliated them by caging them in small rooms and ridiculing black avatars as monkeys [24].

Of course it was just a game. The houses are pixels on a screen, "money" is a play currency called simoleans, and sex between avatars is no more than dirty talk in cartoon bubbles. The CNN anchors kept tongues firmly in cheek and concluded that "this is all taking place in a virtual world. We can hope that it stays there and that if you don't like it you can just leave the game and stop playing" [31]. Electronic Arts, the game publisher and Maxis's parent company, who had long been aware of these activities, handled the publicity angle by an appeal to unreality. Confronted with stories of online prostitution in an interview with *The New York Times*, Jeff Brown, vice president for communications at Electronic Arts, said, "If someone says that is going on in cyberspace, is it lost on anybody that it's not actually happening? No law was violated. It's a game" [4].

Yet the interface between game and world does not present a simple and clean-cut boundary. The motto at Maxis, as we saw in section one, was that in *TSO*, "the sims are real" [16]. A subjective act of identification with the sim is an integral aspect of the design of the game: from your perspective as a player, your sim was intended to function as a virtual self onto which your own subjectivity and agency is displaced, and you become emotionally invested in the sim's changing social relations, reputation, and resources. Equally, in order to interact with and understand the other sims, you need to model the sim as an avatar channeling a real person. Since the sims are real in the sense that each one is a real individual once removed, displaced onto an avatar, then ethical questions that could be entirely ignored as fictive in the offline version of the game have inescapably moved much closer to reality. Add to this the fact that creative players actively widened the boundaries of the game, extending their gameplay far beyond the confines of the game itself, and in various ways integrating their in-game character with the real world. Finally, they found ways of taking strategic control by devising their own methods of earning money within the game, bypassing *TSO's* intended gameplay. EA continued to treat the game as if everything were happening within their proprietary controlled world, when this clearly was not the case. Ludlow's research assistant, Candace Bolter, pointed out that although prostitution was remunerated using simoleans, the in-game currency of *TSO*, this "fantasy money" was readily convertible into US dollars through money trading at online auction sites such as Gaming Open Market and eBay [33]. Electronic Arts, they argued, had a moral responsibility of governance that they couldn't simply walk away from: *The Sims Online* was not just a game.

Urizenus' *The Alphaville Herald* provides a case in point: it started out as an in-character blog, but in the course of documenting the conflicts and exploitation taking place within the game, it quickly morphed into a project of serious investigative journalism. The significance of this stance towards the game was underlined by the following incident. In October 2003, roleplaying as the priest Urizenus in a *TSO* church, he was contacted by another player:

> A sim IM'ed me claiming to be a 13-year-old boy and started asking me about God and forgiveness. He claimed that he had beaten his 8-year-old sister because she had annoyed him, and that she had gone to the hospital with a broken jaw. I asked if he had reported this, and he said no, and then broke off contact with me [34].

Responding to the reality of the situation and casting aside an implied in-game seal of confession, Ludlow contacted EA to report a real-life crime, repeatedly requesting they pursue the case with the boy's local authorities. EA responded with a string of boiler-plate customer service replies, advising him that they could only take action on Terms of Service violations [34]. Bolter finally threatened to take the case to the media; EA relented and handed the case over to local police. They also terminated Ludlow's account on a technicality. In the following days and weeks, the termination story and the background corpus of interviews was picked up by the international media, from *The Detroit Free Press* to the *Corriere della Sera* and *Izvestia*, featured as an attempt to suppress the public's knowledge of the truth and a possible violation of freedom of speech. In a recent academic paper, Ludlow notes that the real-world press incongruously treated *The Alphaville Herald* as a real newspaper, even though it was produced as an in-character blog as part of a game, and advances the thesis that "there is no such thing as fiction, and there are no such things as fictional objects" [35].

In this section, we have argued that taking *The Sims Online* entailed a loss of player control at both ends of the spectrum (figure 12-1). While large numbers of casual players, including our study participants, appear to have responded to the highly constrained game by attempting to adapt and eventually losing interest, a small number of highly inventive players found a variety of ways of taking charge of the game. They accomplished this by developing intermediate-term strategic control on two fronts. First, they found original ways of bending and twisting the built-in features of the game into complex series of moves to achieve their own defined goals. Second, they extended their gameplay beyond the confines of the game to include web sites, instant messaging, and media outreach, thus embedding their in-game character within a broader matrix of cultural references and meanings. Moving in-character communication out of the game allowed the players to develop their "warped" use of the programmed rules in a conspiratorial spirit of secrecy, away from prying eyes of EA's monitors, thus deepening the gameplay and staying ahead in the conflict of interest between developers and adventurous gamers. Web sites served the function of reinforcing and advertising social role and player identity, creating a community memory of causally connected historical events, or a cumulative repository of instructions.

Because the regular gameplay within *The Sims Online* was so boring, most of what has been recorded about the game on web sites and discussion boards relates to the events generated by the small number of players who played against the grain. These players took charge of the game and moved it in new directions. Evangeline, for instance, who turned out to be a boy, exposed himself and other underage players to forms of adult sexual imagination that may have been harmful and developmentally inappropriate. Jeremy Chase started playing JC Soprano when he was unemployed; his masterful gameplay got him paid work to play online games and design web sites [36]. Peter Ludlow became world famous as the academic that exposed the seamy side of *The Sims Online*, and his career may take a turn towards active worlds research. In each of these cases, it was the originality of the gameplay that created the real-world consequences. The fictional world blended into the real world because the players dragged the real world into the game.

5. Conclusion and Outlook

The chief creator of *The Sims*, Will Wright, often cites the work of Christopher W. Alexander as an important inspiration for his work, in particular his *A Pattern Language* [37]. Communities, Alexander *et al.* write, emerge out of "a hierarchy of social and political groups, from the smallest and most local groups—families, neighborhoods, and work groups—to the largest groups—city councils, regional assemblies" [9]. The enabling condition that allows such groups to form spontaneously and to constitute communities is that "each group makes its own decisions about the environment it uses in common . . . . Ideally, each group actually owns the common land at its 'level.' And higher groups do not own or control the land belonging to lower groups—they only own and control the common land that lies *between* them, and which serves the higher group" [9]. By assigning resources to be held in common at different levels of organization, people would be challenged to institute appropriate collaborative patterns of governance at each level.

In *The Sims Online*, these architectural principles of community design were never implemented. The community has a flat structure of individual ownership, with no land held in common at any level. In part, this happened by default as the code was ported from *The Sims*; the single-player model of private, suburban lots was simply scaled up to a multiplayer environment. Far from acknowledging this to be a massive design blunder, however, EA championed private ownership of all land as the key to forestalling griefers, as each player could bar any other from their site. The result is distinctly odd: since many players don't own properties, they must meet their avatar's basic needs by entering other player's private houses, using their shower and bathroom, getting food from their fridge, and even sleeping in their beds. This socially anomalous practice is not only tolerated but encouraged, since owners are rewarded for the time other sims spend on their properties. Indeed, as a result, the primary function of a private property is not that of a home planned and decorated to delight its owner and inhabitant. Instead, properties become investments, designed to optimize return on capital by providing the skilling and working equipment, along with food, gyms, motel-style beds, showers, and bathrooms, to attract and retain the maximum number of visitors. As we saw in section two, the incentive structure of this economic model not only produces very boring play but reduces human relationships to one of instrumentality, grinding any real sociality out of the game.

By leaving out the multiple levels of commons recommended by Alexander, the game designers blocked the formation of effective higher-level community structures within the game. It turned out that the hyper-privatized model was not effective in preventing crime and griefing: Evangeline and others set up their bordellos on their own properties, and the mafia leveraged the power of roommates, harassing chat and e-mail, negative reputation links, and a host of other tools within and outside of the game to achieve their ends. Community-based governing structures, in contrast, had a hard time getting off the ground. Consider the case of the Alphaville Government, established by a group of friends in early 2003. This act involved the creation of the avatar Mr-President, played by Arthur Baynes, and the construction of a Capitol in the best neoclassical tradition. Baynes web site at avg.simsgov.com showed an animated graphic of Mr-President waving from the balcony. The focus, then, was on the trappings of power, and since all properties were private, the Capitol, and later the Court, had to belong to an individual player who might at any time decide to quit. The Alphaville Government also aspired to take on the task of maintaining law and order within Alphaville, but they had no legitimate tools with which to enforce city-wide laws and regulations and never became anything like an effective government.

In the fall of 2004, as this manuscript goes to press, there is widespread expectation among the gamers that *The Sims Online* is about to close. The initial prospect of hundreds of thousands of customers paying monthly fees of ten dollars each had made the idea of porting their most popular game to the Internet financially extremely attractive to EA. Suneel Ratan, its former vice president, said informal projections had run as high as a million subscribers for the online version, implying a regular annual revenue of $120 million [27]. Instead, *TSO* subscriptions at the first anniversary were estimated to be around 80,000 [20] and by April of 2004, the company reported 57,500 subscribers [38]. If the attrition rate of 5,000 a month is sustained, *TSO* will dip below 20,000 before the year is out—the level at which EA closed Earth & Beyond [20].

Several of the key players have taken the consequences. After being kicked out of *The Sims Online* in late 2003, Peter Ludlow continued for a while to visit Alphaville through other players' accounts, but soon moved on to *Second Life*, a very differently managed online social space. In July 2004, Jeremy Chase predicted *TSO* would close after a final Christmas season and reported he had joined the online multiplayer game *Star Wars Galaxies*. In August, Simoleanman, one of the main currency traders, announced his closing sale, and The Alphaville Government shut down. It is possible that it's still not too late to rescue the game: the example of the innovative rebels has inspired a host of online newspapers, mafias, and government players, and EA has continued to make significant improvements to the game. A crucial innovation in friendship formation has moved the game closer to a psychology of real relationships. Yet these improvements may have arrived too late.

In its heyday, *The Sims Online* did a spectacular job as a dystopian experiment and remains a rich source of opportunities for social research. Rushed into production, it contains layers of questionable design decisions. Its doll-house ancestry, emphasizing sims driven by artificial intelligence, militated in multiple ways against effective player-avatar identification. The incentive system implicit in its economic structure produced mind-numbingly boring play and purely utilitarian avatar relations. Its lack of strategic opportunities spurred creative players to extend their gameplay beyond the company servers and to bend the built-in features of the game to their own purposes. Finally, the absence of land held in common at different levels prevented the natural emergence of effective layers of governance. Yet precisely by inspiring players to warp the rules and to extend the game to the world-wide web, *The Sims Online* has succeeded in introducing the larger practical and ethical questions of online community governance to the public at large, making us all participants in a virtual utopian experiment.

Acknowledgments

We thank the participants for making a year-long commitment to participating in this study. We also thank the National Science Foundation for funding the Children's Digital Media Center, UCLA, under whose auspices the research was conducted.

References

1. Grossman, L. (2002). Sim Nation. *Time*, 25 November. Available at http://www.time.com/time/magazine/archive/preview/0,10987,1101021125-391544,00.html
2. Seguret, O. (2004). Putsch et Sims. *Liberation*, 27 February. Available at http://www.liberation.fr/page.php?Article=181923#
3. Ward, M. (2003). The dark side of digital utopia. *BBC News*, 22 December. Available at http://news.bbc.co.uk/go/pr/fr/-/1/hi/technology/3334923.stm
4. Harmon, A. (2004). A Real-life debate on free expression in a cyberspace city. *The New York Times*, 15 January. Available at http://www.nytimes.com/2004/01/15/technology/15SIMS.html
5. Manjoo, F. (2003). Raking muck in "The Sims Online". *Salon.com*, 12 December. Available at http://www.salon.com/tech/feature/2003/12/12/sims_online_newspaper/index_np.html
6. Keighley, G. (2002). The endless hours of The Sims Online. *Gamespot.com*, 28 November. Available at http://gamespot.com/gamespot/features/pc/simsonline
7. McKenna, K.Y.A. & Bargh, J.A. (2000). Plan 9 from cyberspace: The implications of the Internet for personality and social psychology. *Personality and Social Psychology Review 4*, *1*: 57–75.
8. Lessig, L. (1999). *Code and Other Laws of Cyberspace*. New York, NY: Basic Books.
9. Alexander, C. W, Ishikawa, S., & Silverstein, M. (1977). *A Pattern Language: Towns, Buildings, Construction*. New York, NY: Oxford University Press.
10. Walker, J. (2004). Hobbes in cyberspace: Life in an online game world proves nasty, brutish, and short. *Reason*, 1 April. Available at http://www.keepmedia.com/ShowItemDetails.do?item_id=387500
11. Electronic Arts (2003). *Annual report*, 31 March. Available at http://ccbn26.mobular.net/ccbn/7/266/277/
12. Gordon, J.E. (1978). *Structures: or, Why Things Don't Fall Down*. New York, NY: Penguin.
13. Salen, K. & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Boston, MA: MIT Press.
14. Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
15. Maslow, A.H. (1943). A theory of human motivation. *Psychological Review*, *50*: 370–396.
16. Green, J. (2002). The Sims Online: Indulging your inner weirdo. Gamers.com, May 1. Available at http://www.gamers.com/game/1016135/previews?page=1
17. van den Berg, M. (2004). Simmery and sluttery . . . the mo pie sims. Available at http://www.mopie.com/sims/simdex.html
18. Shakespeare, W. (2001). *The History of King Lear*. S. Wells (Ed.). Oxford, UK: Oxford University Press.
19. Schroeder, R. (2002). Social interaction in virtual environments: Key issues, common themes, and a framework for research. In R. Schroeder (Ed.), *The Social Life of Avatars: Presence and Interaction in Shared Environments*. London: Springer, pp. 1–18.
20. Woodcock, B.S. (2004). An analysis of MMOG subscription growth. Version 10.0. September. Available at http://pw1.netcom.com/%7Esirbruce/Subscriptions.html
21. Greenfield, P.M. (1997). Culture as process: Empirical methods for cultural psychology. In J.W. Berry, Y. Poortinga, & J. Pandey (Eds.), *Handbook of Cross-Cultural Psychology: Vol. 1. Theory and Method*. Boston, MA: Allyn & Bacon, pp. 301–346.
22. Steen, F.F., Greenfield, P.M., Davies, M.S., & Tynes, B. (In press). What went wrong in The Sims Online? Cultural learning and barriers to identification in a MMORPG. In P. Vorderer & Jennings Bryant (Eds.), *Playing Video Games: Motives, Responses, and Consequences*. Mahwah, NJ: Erlbaum.

23. Kirk, C. (2004). Culling external sensory response: How we feel videogames. Unpublished manuscript, Department of Communication Studies, UCLA.
24. Brown & Bell, this volume.
25. White, A.A. (2003). Chatting for dummies. Game-Revolution.com, January. Available at http://www.game-revolution.com/games/pc/sim/sims_online.htm
26. Kollock, P. (1996). Design principles for online communities. Harvard Conference on the Internet and Society. Available at http://www.sscnet.ucla.edu/soc/faculty/kollock/papers/design.htm
27. Ratan, S. (2003). Sim flop dogs game developers. *Wired*, 12 May. Available at http://www.wired.com/news/games/0,2101,58749,00.html
28. Hopkinson, T.B. (2003). The Sims Online mob interview. *GameGossip.com*, 11 June. Available at http://www.gamegossip.com/comment.php?id=208
29. Ludlow, P. (2004). JC Soprano: Now the city is mine! *The Alphaville Herald*, 14 January. Available at http://www.alphavilleherald.com/archives/000089.html
30. Wadhams, N. (2003). Sex, mob hits: Sims tests virtual morals. *Associated Press*, 5 July. Available at http://www.cnn.com/2003/TECH/fun.games/07/05/misbehaving.online.ap/index.html
31. Sieberg, D. (2003). Next@CNN, *CNN*, 2 August. Transcript available at http://www.cnn.com/TRANSCRIPTS/0308/02/nac.00.html
32. Twist, J. (2004). Simulating life, love and the Universe. *BBC*, 17 September. Available at http://news.bbc.co.uk/1/hi/technology/3645552.stm
33. Ludlow, P. (2003). Evangeline: Interview with a child cyber-prostitute. *The Alphaville Herald*, 8 December. Available at http://www.alphavilleherald.com/archives/000049.html
34. Ludlow, P. (2003). Comment. *The Alphaville Herald*, 11 November. Available at http://www.alphavilleherald.com/archives/000016.html
35. Ludlow, P. (In preparation). From Sherlock and Buffy to Klingon and Norrathian platinum pieces: Pretense, contextalism, and the myth of fiction. Available at http://www.alphavilleherald.com/archives/Fiction.rtf
36. Chase, J. (2004). Bombshell! JC Soprano retires from TSO, but confesses ALL on the way out. *The Second Life Herald*, 20 July 2004. Available at http://www.alphavilleherald.com/-archives/000331.html
37. Cambron, M. (2002). A chat with Will Wright. *GIGnews.com*, April. Available at http://www.gignews.com/goddess_wright.htm
38. Glassman, M. (2004). Braving bullying hecklers, simulants run for president. *The New York Times*, 1 April. Available at http://tech2.nytimes.com/mem/technology/techreview.html

Index