Mel Slater
M.Slater@cs.ucl.ac.uk

Amela Sadagic
Martin Usoh
Department of Computer Science
University College London
UK

Ralph Schroeder
School of Technology Management and Economics
Chalmers University
Sweden

Small-Group Behavior in a Virtual and Real Environment:
A Comparative Study

Abstract

This paper describes an experiment that compares behavior in small groups when its members carry out a task in a virtual environment (VE) and then continue the same task in a similar, real-world environment. The purpose of the experiment was not to examine task performance, but to compare various aspects of the social relations among the group members in the two environments. Ten groups of three people each, who had never met before, met first in a shared VE and carried out a task that required the identification and solution of puzzles that were presented on pieces of paper displayed around the walls of a room. The puzzle involved identifying that the same-numbered words across all the pieces of paper formed a riddle or saying. The group continued this task for fifteen minutes, and then stopped to answer a questionnaire. The group then reconvened in the real world and continued the same task. The experiment also required one of the group members to continually monitor a particular one of the others in order to examine whether social discomfort could be generated within a VE. In each group, there was one immersed person with a head-mounted display and head-tracking and two non-immersed people who experienced the environment on a workstation display. The results suggest that the immersed person tended to emerge as the leader in the virtual group, but not in the real meeting. Group accord tended to be higher in the real meeting than in the virtual meeting. Socially conditioned responses such as embarrassment could be generated in the virtual meeting, even though the individuals were presented to one another by very simple avatars. The study also found a positive relationship between presence of being in a place and copresence—the sense of being with the other people. Accord in the group increased with presence, the performance of the group, and the presence of women in the group. The study is seen as part of a much larger planned study, for which this experiment was used to begin to understand the issues involved in comparing real and virtual meetings.

1 Introduction

There is substantial interest in the use of virtual environments (VEs) as a medium for collaboration between remote participants, and several systems and applications have been established to enable this, for example (Carlsson & Hagsand, 1993; Greenhalgh & Benford, 1995; Leigh & Johnson, 1996; Macedonia & Noll, 1997; Major et al., 1997). There is also an explosion of multiuser virtual online worlds and communities and the start of research into the social relations that emerge in such communities, surveyed recently by Schroeder

Presence, *Vol. 9, No. 1, February 2000,* 37–51
© 2000 by the Massachusetts Institute of Technology

(1997), Schiano (1999), and Kollock (1999). However, there has been limited study of what happens when small groups of people actually make use of these systems for collaboration (Bowers et al., 1996). This paper describes an experiment—in fact, part of a much larger planned experiment—that asks the question: What is the experience of participants when carrying out a task with others in a shared VE, and how does that experience compare with working with these others on the same task in the real world?

The experiment was designed to explore the behavior of small groups carrying out a task initially in a virtual and then continuing in a real environment. Each of the ten groups consisted of three people who were unknown to one another beforehand. The group task, to be described fully later, consisted of solving a set of riddles. The task involved only observation and talking, and it could be solved most efficiently by group cooperation.

The focus of the study was not at all on performance, in the sense of how well the task was completed, but rather on how the social relations among the members developed in the virtual environment, and how, if at all, these carried over to their interactions in the real world. In particular, the study was concerned with the following issues:

- Does computational advantage confer social power? One of the group participants was immersed in a virtual environment with a head-tracked, head-mounted display, and the other two were not immersed but used a desktop workstation display. None of the participants had information as to the type of system the others were using. To what extent would the immersed person, given the empowerment bestowed by their computational advantage, become the leader of the virtual meeting, and to what extent would this carry over to the later real meeting?
- Is the sense of presence of being in the virtual place associated with copresence—the sense of being and acting with others in a virtual place? This is a useful question to ask, because, if presence and copresence are associated, this could be a result of common factors influencing both, or because the individual sense of presence influences the chance of an emergent copresence (or vice versa). This was studied using reported presence based on post-experimental questionnaires.
- How does the sense of enjoyment and feelings of group affection vary as between the virtual and the real experience? An attempt was made through questionnaire and post-experimental debriefing to assess the extent to which the experience was ''positive,'' and how this changed in the transition from virtual to real.
- Can reactions such as embarrassment, shyness, and conflict be generated in the virtual environment, and, if so, to what extent does this carry over to the real? In the virtual environment, one of the participants was given instructions, unknown to all others, to closely follow and observe another participant. This could affect group interaction in several ways: the embarrassment of the observer, the annoyance of the observed, and the sense of being left out of things by the third person.

To our knowledge no one before has studied small group meetings in virtual environments with the participants continuing the same task in a real environment (of which the virtual was a simulation). In this experiment, there was an attempt to explore the pattern of relationships within the shared VE, and also to see how these changed in continuing real meetings. The work described in this paper nevertheless makes a limited start in this endeavor. The start is limited for two main reasons. First, the length of time of the meetings was very short (fifteen minutes in the virtual followed by fifteen minutes in the real). Second, the order in which the meetings occurred (first virtual and then real) requires a control situation in which a similar number of groups carry out the experiment first in the real and then continuing in the virtual. This paper describes a study at a certain incomplete stage. Nevertheless, the results stand in their own right as a study of what happens in the transfer from virtual to real meetings.

The details of the experiment are given in Section 2. Results obtained by the use of post-experimental questionnaires are given in Section 3, and results from de-

briefing sessions in Section 4. Section 5 discusses the results in relation to other published work, and the conclusions and way ahead are presented in Section 6.

2 Experiment

2.1 Scenario

The study involved ten groups of three people recruited by advertisement on the university campus. No payment was made for taking part in the study. The experiment took place over a two-week period. Four experimenters were involved in the study, one ''minder'' each to look after one of the subjects, and a ''floor manager'' who maintained overall control and synchronization of the various activities. The experiment took place in a large laboratory divided into partitions, with the three subjects at opposite sides of the laboratory. Care was taken to prevent the subjects from seeing or meeting each other before the start of the experiment.

As each subject arrived, they were assigned to their minder who took them to their assigned workstation, or in one case to the immersive virtual reality room at one end of the main laboratory. Each subject was assigned a color (red, green, or blue), and they were referred to by that color throughout the experiment and later debriefing. The subjects could not see their own avatars (except for Red, the immersed participant, if he or she looked downwards).

Each subject was introduced to the system that they would be using. This was either a desktop system (Green and Blue) or an immersive system with an HMD (Red). The virtual environment that was displayed was actually a rendition of the laboratory in which they were physically located. Each person was represented by an avatar of the same color as their assigned name.

Their first task was to individually learn to move through the environment. Then, at a signal from the floor manager, each subject was given a sheet describing the overall task to be performed. Then again on a signal, they were invited to put on earphones and to introduce themselves to one another. They could refer to themselves and to the others only by their color.

The task was to locate a room that had sheets of paper displayed around the walls. Each sheet had several words in a column, with each preceded by a number. The words across all sheets with a common number combined to form a saying (for example, ''A critic is a man who knows the way but can't drive a car''). The task was first to figure this out and second to unscramble as many of these sayings as possible.

The subjects were asked to find, as a group, the room with the papers and then solve the puzzle. The room with the papers was the rendition of the room with the virtual reality equipment, where Red was physically located.

Green was given an additional task, not revealed to the others. Green was asked to monitor Red as closely as possible, always trying to be in Red's line of vision, although taking part in the puzzle-solving task as much as possible. If Red objected, Green was to comply temporarily with Red's wishes, but then continue anyway with this monitoring task.

The minders sat unobtrusively near the subject throughout the virtual part of the group activity, in case of problems. The minder of Green had an additional job: prompting Green to obstruct Red if Green did not appear to be carrying out this task and instead becoming only involved in the puzzle-solving activity.

After about fifteen minutes, the virtual session was terminated, and the subjects completed a questionnaire, which took about ten minutes. Then each subject was required to put on a waistcoat of their color, and, at a signal from the floor manager, they all met together in real life for the first time just outside the virtual reality room, the room which had the real puzzles on the walls.

They were then invited to continue the task in the physical location, which lasted for about another fifteen minutes. At the end of that time, they completed another questionnaire, and then met with the floor manager for a debriefing.

During the virtual session, the virtual movements of the subjects were automatically recorded, and an audiotape recorded their conversation. The real session was videotaped from above, giving a plan view.

2.2 Materials

The Red (immersed) person was using a Silicon Graphics Onyx with twin 196 MHz R10000, Infinite Reality Graphics, and 64M main memory, running Irix 6.2. The tracking system has two Polhemus Fastraks, one for the HMD and another for a five-button 3-D mouse. The helmet was a Virtual Research VR4, with a resolution of 742 × 230 pixels for each eye, 170,660 color elements, and a field of view of 67 deg. diagonal at 85% overlap.

The total scene consisted of approximately 3,500 polygons which ran at a frame rate of no less than 20 Hz in stereo. The latency was approximately 120 ms.

The Red subject moved through the environment in gaze direction at constant velocity by pressing a thumb button on the 3-D mouse. A virtual body (avatar) responded to hand and head movements.

The Green subject used an SGI High Impact system with 200Mhz R4400 and 64MB main memory. The scene was shown on the full 21 in. display. Navigation was accomplished by using the keyboard arrow keys, with the up and down arrows providing forward and back movement, and left and right keys providing rotation. All movement was on the horizontal plane of the floor.

The Blue subject used an SGI O2 running at 180Mz on Irix 6.3, with an R5000 processor, and 32MB main memory. The scene was shown on a full 17 in. display. Navigation was the same as for the SGI Impact.

The sound system used was the Robust-Audio Tool (RAT) v.3.023. This allows multiple users to talk over the Mbone (Hardman et al., 1995).

The virtual reality software used throughout was DIVE 3.2 (Carlsson & Hagsand, 1993). A DIVE avatar was used for each of the participants, and it was the same for each except for the color. An image of such an avatar is shown in Figure 1.

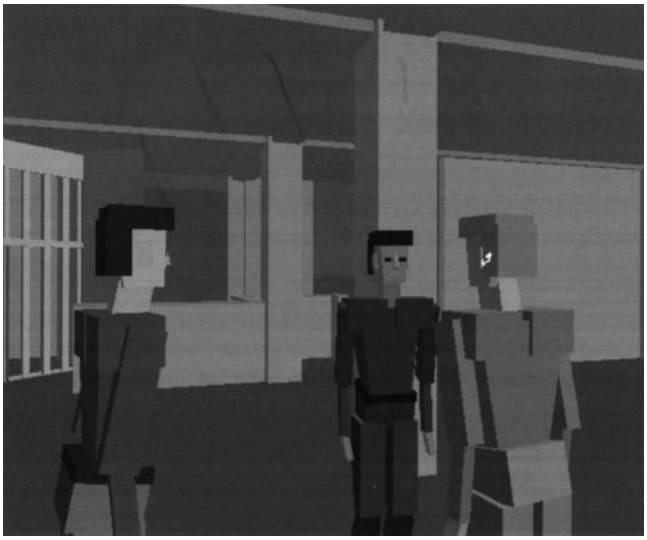

Figure 1. *Avatar representing a participant.*

3 Questionnaire Results

3.1 Leadership

Two questions related to leadership, one directly and the other indirectly. Each subject was asked to score all three subjects on the degree to which that person ''was the 'leader' or main organizer'' in the meeting that had just concluded. The three scores (one for Red, Green, and Blue) had to add to 100. In addition, a similar question concerned who did most of the talking.

Table I. *Mean and Standard Deviation of 'Leadership' Scores*

Person	Score in virtual	Frequency	Score in real	Frequency
Red	46 ± 17	14	33 ± 12	5
Green	34 ± 13	2	35 ± 11	7
Blue	20 ± 13	5	32 ± 10	2

Frequency refers to the number of times out of thirty where the individual had the highest leadership score.

Table 1 shows the results for leadership, after the meeting in the virtual setting, and then after the real setting. The most striking aspect of this is the highly significant difference in leadership rating for Red (the immersed person) compared between the virtual and real. After the real meeting, each participant was assigned approximately the same leadership rating, whereas, immediately after the virtual meeting, Red emerged as the clear leader. In fact, fourteen out of the thirty participants rated Red as the leader immediately after the vir-

Table II. *Mean and Standard Deviation of 'Most Talking' Scores*

Person	Score in virtual	Frequency	Score in real	Frequency
Red	45 ± 17	16	32 ± 12	3
Green	22 ± 11	0	35 ± 11	6
Blue	33 ± 13	6	34 ± 10	5

Frequency refers to the number of times out of thirty where the individual had the highest talking score.

tual session, whereas five rated Red as leader after the real session.

Table 2 shows similar results for ''who did the most talking.'' It is clear that Red was perceived to be the most talkative during the virtual session, but that this did not carry over to the real session. Sixteen of the thirty participants reported Red as the most talkative after the virtual session compared with three of the thirty after the real session.

Two factors distinguish Red from Green and Blue during the virtual session. The first was that Red was monitored by Green. As will be seen later, for the most part, Red was unaware of this, and there is no obvious way that this could have had an effect on leadership behavior displayed by Red. The second difference is that Red was the only one immersed through a head-tracked HMD and a hand tracker. Moreover, Red was on a machine with a faster processor. However, the scene was so small that the frame rate was indistinguishable between the different types of machine. Also, Blue—the one with the least processing power—although scoring least on leadership, had the same level of talkativeness in the virtual and real experiences. The zero score for Green on talkativeness in the virtual part of the experiment probably reflects Green's additional monitoring task.

The first and perhaps most important hypothesis generated from this study is that greater computational resources may enhance leadership capability. The reported leadership behavior of the person who was immersed vanished when all subjects participated on relatively equal terms in the real setting.

3.2 Presence and Copresence

The term *presence* in the virtual environment literature has come to be used to denote the sense of ''being there'' in a place (for example, Held & Durlach, 1992). An orthogonal attribute of presence in a place is the sense of being present with other people. This attribute is logically orthogonal, because, for example, talking on a telephone with someone might give a strong sense of ''being with them'' but not of being in the same place as them. It is useful nevertheless to examine the extent to which these two different types of presence, place presence, and copresence are empirically related. If they are in fact related, then this is either because they influence one another or because there are underlying common factors to both.

The questionnaire asked the following three questions relating to copresence.

1. In the last meeting, to what extent did you have the sense of the *other two people being together with you?*
2. Continue to think back about the last meeting. To what extent can you imagine yourself *being now with the other two people* in that room?
3. Please rate how closely your sense of being together with others in a real-world setting resembles your sense of being with them in the virtual room.

The following two questions related to place presence.

1. To what extent did you have the sense of being in that room which has the pieces of paper with the riddles on the walls? (For example, if you were asked this question about the room you are in now, you would give a score of 7. However, if you were asked this question about whether you were sitting in a room at home now, you would give a score of 1.)
2. Think back now about the meeting and the spatial layout of the room. For example, to what extent in your imagination can you move around that room now?

Each question was rated on a scale from 1 to 7, with 1 equating to ''not at all'' and 7 ''very much so.''

As a conservative measure of the subjective (reported)

Table III. *Responses to Group Accord Questions*

Factor	After virtual	After real	Sig. level for difference (*P*)
1. Enjoyment	4.23 ± 1.19	5.70 ± 1	0.003
2. Meet again	4.23 ± 0.74	4.73 ± 0.41	0.300
3. Isolation	71.50 ± 14.54	44.40 ± 16.08	0.003
4. Meet individuals again	0.66 ± 0.12	0.71 ± 0.09	0.160
5. Comfort with others	0.66 ± 0.09	0.81 ± 0.1	0.002
6. Cooperation	0.77 ± 0.13	0.88 ± 0.12	0.010
7. Embarrassment	0.25 ± 0.09	0.20 ± 0.06	0.110
8. Overall accord	0.62 ± 0.06	0.87 ± 0.06	0.000

level of place presence and copresence, only the high scores were taken into account. The overall measure of place presence is the number of scores of 6 or 7, and hence is a count of 0, 1, or 2. Similarly, the overall measure of copresence is the number of scores of 6 or 7, and hence is a count of 0, 1, 2, or 3. This approach is the same as has been used in previous studies of presence (Slater & Wilbur, 1997).

The correlation between these two scores ($r = 0.59$) is significant ($p = 0.0006$). Considering only the raw scores for the two basic questions (copresence 1 and place presence 1), $r = 0.52$ at a similar level of significance. It is interesting to note that the immersed person (Red) did not report a significantly higher level of presence on any category.

The second hypothesis generated from this study is therefore that presence and copresence are linearly associated, but that the immersed person did not report a higher level of either type of presence than the other two.

3.3 Group Accord

Several questions attempted to assess the group members' appraisals of one another and the group as a whole. All but one question was rated on a 1–7 scale, with 1 the lowest level of the quality concerned (for example, enjoyment) and 7 the highest quality. In each case, the overall group means and standard deviations are given for responses after the virtual and after the real setting.

Table 3 shows the responses to these questions after the virtual session and after the real session. The significance levels are for paired *t* tests over the ten groups.

The corresponding questions are as below.

1. (Enjoyment) Think about a previous time when you *enjoyed* working together in a group. To what extent have you *enjoyed* the group experience just now?
2. (Meet again) Sometimes you meet people in a small group situation, and you'd like to meet them again. To what extent is the current situation similar to that?
3. (Isolation) To what extent was anyone (including yourself) ''isolated'' compared to the other two people? Give a score for each individual out of 100, where a person scores 100 if they were completely isolated from the other two, and where the three scores add to 100. (In this case the maximum degree of isolation was taken as the score for the group as a whole).

The following questions required a response by each subject for each of the other two subjects (for example, Red would give responses with respect to Green and Blue). The score for the group is taken as the sum of the

six scores for the individual members (six because each individual does not self-score), divided by the total possible score for the group, which is 42.

4. (Meet individuals again) Would you like to meet any of the other two people again? (Please put one tick in each column.)
   - (1) I would not like to meet this person.
   - (4) No preference either way.
   - (7) I would very much like to meet this person.
5. (Comfort with others) The extent to which I felt comfortable with each of the other two persons was (please put one tick in each column):
   - (1) I felt very uncomfortable with him / her.
   - (4) Neither comfortable / nor uncomfortable.
   - (7) I felt very comfortable with him / her.
6. (Cooperation) Overall, how cooperative were each of the other two people in the task?
   - (1) S / he was not cooperative at all.
   - (7) S / he was very cooperative.
7. (Embarrassment) Did any of the other two people make you feel self-conscious or embarrassed?
   - (1) S / he did not make me feel this way.
   - (7) S / he did make me feel this way very much.
8. (Overall accord) Finally, each of the seven variables above were combined into one overall score for group accord. In order to make each of the variables result in greater accord in a range from 0 to 1, the scores out of 7 are normalized to be between 0 and 1, non-isolation is taken as 1 − (isolation / 100), and non-embarrassment is embarrassment subtracted from 1.

Taking overall group scores reveals a significant difference between the result after the virtual session and after the real session, with overall group accord higher after the latter. In particular, after the real session, there was greater enjoyment, less isolation of individual members, a greater sense of comfort with the other members, and more cooperation.

The reason for the differences might not be solely due to the nature of a virtual compared to a real encounter. Another factor that was different between the two sessions was that, in the virtual session, Green was asked to monitor Red, while this was not the case in the real session. However, when the responses for the individuals are examined, there are no significant differences between Red, Green, and Blue for any of the accord variables considered above.

There is also simply the question of time. After the real session, the group members had been working on the puzzle altogether for approximately thirty minutes, compared to fifteen minutes after the virtual session. This study should be considered as the first part of a larger experiment in which another ten groups repeat the experiment but with the order of session reversed (real first and then virtual). From this study, it would be possible to see if there were a significant increase in accord after the second session. If so, then the result would be most likely due to time.

3.4 Accord and Presence

A previous study (Barfield & Weghorst, 1993) has found a significant relationship between presence and enjoyment. In order to examine this in relation to the current experiment, a measure of individual accord was constructed on the same lines as in the previous section, except now for each individual rather than for the group as a whole. This was used as the response variable in a regression analysis in which the major explanatory variables were presence, copresence, and combination of the two.

Figure 2 shows a plot of individual accord against the combined count of presence and copresence ($r = 0.72$). Using the combined presence count as an explanatory variable in a regression analysis results in a significant fit; also, gender and the number of riddles solved are significant explanatory variables. Women tend to show higher accord scores than men, and the more riddles solved the greater the accord. This is shown in Table 4.

The copresence aspect of overall presence dominates the relationship. If copresence only is used as the explanatory variable, then a very similar result to Table 4 emerges (with $R^2 = 0.61$). If place presence only is used, then the number of riddles is no longer significant, although gender remains so, with $R^2 = 0.45$).

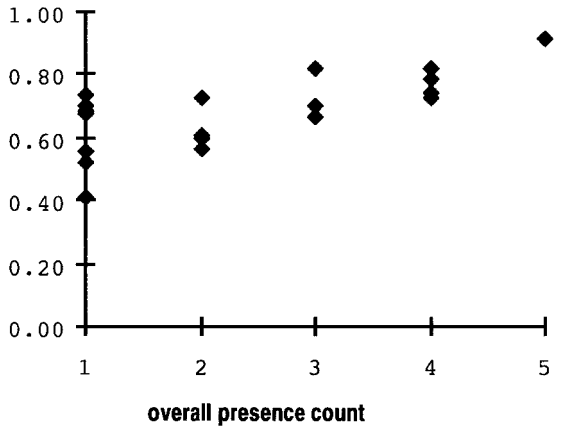

Figure 2. *Accord against overall presence count.*

Table IV. *Multiple Regression for Accord*

Parameter	Estimate	S.E.	*t* value
Constant	0.534	0.023	
Increment in constant for women	0.086	0.034	2.230
Riddles solved	0.015	0.007	2.062
Overall presence	0.056	0.009	6.008

3.5 Analysis of Free Responses

After the virtual session, the questionnaire included an instruction for participants to list any things that hindered them from successfully accomplishing the task.

The report concentrates only on issues that were raised by several people, rather than the more idiosyncratic comments particular to only one person. Three common themes were mentioned by several people that hindered them in the task: poor navigational ability, poor audio, and the discomfort of the immersed group.

- Poor Navigation was recorded as a problem by eight non-immersed and seven immersed people. The problem of ''going through walls'' was especially mentioned as part of this issue.
- Poor Audio was mentioned by ten people. Particular issues mentioned were delays in audio, lack of communication in the sense of being difficult to know if someone were talking to someone else, not being able to hear their own voice at normal level (tending to be too loud) and not being able to easily realize who specifically was talking.
- Immersion Discomfort was reported by five of the ten immersed people. Particular comments mentioned headaches, things being slightly out of focus, feeling sick and sweaty, and, specifically ''it was a physically uncomfortable experience—by the end of the session it was very much distracting me from the task.''

3.6 Summary

This section has examined the results of the questionnaire data. Salient hypotheses that may be generated from this study are as follows.

- Immersion enhances leadership capability. The immersed person was overwhelmingly recognized as the leader in the virtual session, but this disappeared in the real session. This finding was confirmed by a separate question concerning which person did the most talking.
- Presence (being in a place) and copresence (being with other people) were positively correlated.
- Reported presence was not significantly different between the immersed and non-immersed people.
- Group accord increased in the real session compared to the virtual (though it is not possible in this study to rule out the effect of time).
- Higher individual accord was associated with higher overall place presence and copresence.
- Individual accord tended to be higher for women than for men and was positively associated with more-successful performance of the task.
- There was no reported effect of the attempt to deliberately introduce some embarrassment into the virtual session by having one subject monitor another. No differences between the three role colors were reported on any component of accord.

4 Results of the Debriefing Sessions

4.1 Impact of the Monitoring Task

Questionnaires are able to capture rather static and limited information about events. It is often useful to use face-to-face, unstructured encounters in order to look behind the questionnaire data and get a better understanding of what was happening—to allow for possibilities not envisaged during the questionnaire design, and to explore the dynamics of the situation. Therefore, at the end of the experiment, the participants were invited for a debriefing session to allow them to talk freely about their experiences. In each such debriefing, the first issue for discussion was whether Red noticed anything unusual in the behavior of the other two participants, and then the extent to which Green had found the monitoring task awkward or embarrassing.

In three of the ten groups, there was an impact of this additional task by Green. In group 3, Red formed the opinion that Green was being deliberately destructive. Also, in this particular session, the sound from Red was ''crackling,'' and Green thought that Red was doing this deliberately. All three members of this group (Red and Green were male, Blue female) had a high sense of what they described as ''paranoia'' during the virtual session, and agreed that this completely disappeared when they met for real. This group actually never figured out even what the puzzle was and found this to be frustrating.

In group 9 (Red and Green female, Blue male), Red did notice something different, but interpreted this as something being wrong with the avatar configurations. She said that ''everyone was supposed to be looking at the walls, but Green was looking at me.'' In this same group, Green reported that ''I felt I wasn't being me'' and ''What on earth were they thinking of me?''—and found it especially difficult because she was supposed to be doing two tasks at the same time (monitoring Red and helping with the puzzle). She imagined that the other two were ''wondering why I am doing this.'' Sometimes she wondered if Red would think that she were staring at her.

In group 10 (Red and Green male, Blue female), Red did not notice that Green was observing him, but did notice that the way ahead seemed to be frequently blocked. Green was not embarrassed to carry out this task. However, in this group, the major impact was on Blue, who thought that Red and Green ''know where they are—I felt excluded.'' In other words, Blue noticed that Red and Green seemed to be close to one another most of the time, and Blue was left out.

One thing reported by almost all Greens was the difficulty of carrying out the monitoring task at all. Red moved faster than the other two subjects (on the more powerful machine and immersed). Also, it was difficult for Green to know Red's field of view. There being no virtual equivalent of eye contact in any meaningful sense, Green could never know whether or not Red was aware of Green's activities. There could be no exchange of glances. More generally, this lack of feedback about body movements and body language from the avatars was mentioned by several people.

4.2 Relationship to Avatars

A second major issue explored in the debriefings was the relationship of the people to their avatars. The most interesting way in which this was realized was through projection; that is, individuals were respectful of the avatars of the other people and tried to avoid carrying out actions that would cause distress or be impossible in real life.

- In group 1, Blue said that walking through the avatar of another (which happened frequently by accident in the confined virtual space) led to his embarrassment. In the same group, Red reported that walking through another body was ''weird,'' although Red experienced the situation as like being in ''fancy dress,'' the others were ''not quite real people, without a human presence, just pixels.''
- In group 2, Green said that it ''didn't bother me to walk through people. This was the rule of this universe.'' In the same group, Blue found it ''frightening'' to walk through a person.
- In group 4, neither Red nor Green minded about this issue, but Blue had the impression that it was ''rude'' to walk through someone.
- In group 5, Green found it annoying if someone went through him, and Blue also thought that it was ''bad if someone walks through you.''

- In group 7, Red and Green each reported saying ''sorry'' when walking through someone.
- In group 9, Red felt it was ''disconcerting when bodies passed through each other.'' Also, it was ''irritating'' when she ''walked back through someone and didn't know.'' In the same group, Green reported that she ''didn't mind going through things.'' Blue said that, when Red came up close to him, he felt ''really uncomfortable, bloody uncomfortable,'' and backed off.
- In Group 10, Red ''felt like apologizing'' when he went through someone.

Some groups also discussed the impact of the ability to go through walls. (There was no collision detection at all.) In group 1, Red felt himself to be panicking when he seemed to be ''stuck in the wall.'' In group 2, Green reported that it was frightening and if he did so and was outside of the scenario then this induced an agoraphobic feeling. He also did not like the fact that he could not look up or down, but only straight ahead (not being immersed, there was no option to swivel the gaze direction up and down). In group 6, Blue did not like the ability to go through walls (which was easily done by accident). Green reported the same in group 9.

This process of being mindful of the avatars of others was surprising: they were taken seriously in spite of all their shortcomings. This relationship to the avatars was noticed in another way—the surprise that some people experienced on meeting the real person. Some of the group ''reunions''—the moment when they met for real for the first time—can only be described, unscientifically, as somewhat emotional. In group 6, Green reported a ''shock'' when she really met the others. In group 9, Red was surprised to see what Green looked like for real, and Green was similarly surprised by the appearance of Red. In the same group, Blue found surprising the shape of the others' heads; somehow he had expected these to be the same as in the virtual session.

4.3 Summary

This analysis of the post-experimental group discussion revealed a surprising degree of attachment and relationship towards the virtual bodies (avatars). Although, except by inference, the individuals were not aware of the appearance of their own body, they seemed to generally respect the avatars of others, trying to avoid passing through them, and sometimes apologizing when they did so. These were very simple avatars, with limited movement and no capability for any kind of emotional expression. If even these can evoke such responses, it is interesting to wonder what responses more-powerful avatar representations might evoke.

5 Discussion

5.1 Why Shared VEs Are Needed

The need for shared VEs for collaborative working is not obvious. Clearly, multimedia systems with real-time video and audio are capable of bringing people in remote locations together for collaborative work. It could be argued that such multimedia systems are not suitable when there is a requirement to manipulate objects or share design—although whiteboards go a long way in helping with such tasks. A study is considered in this section, where, even though the task does not involve shared design or manipulation, the results strongly suggest that a shared VE might offer substantial benefits. Isaacs et al. (1995) describe an experiment using the Forum system, a desktop-based video and audio system to compare face-to-face presentations with distributed ones. The application involved people giving fourteen presentations to groups, half by the presenter in a lecture hall with the audience in conventional style, and half using the Forum system. The presentations were paired so that the presenter gave the same material twice, once to an audience in a face-to-face lecture hall setting, and the other to a different distributed audience using the Forum system.

The Forum involved live video, audio, and slides presented on a desktop workstation. The audience members could see live video of the presenter, as well as the slides, which could be followed along with the speaker, independently scrolled and annotated by the audience members. The audience members could speak to the whole group and send messages to the speaker and one an-

other. The speaker could not see the audience, and the audience members could not see one another.

The Forum audiences could be using other applications on their desktop machine during the session, whereas of course the lecture attendees had to physically travel to the meeting place and could not easily be engaged in other activities during the course of the lecture. Hence, the Forum audiences tended to be larger, and also self-selected.

The important results from this paper's point of view concerned the perceived quality of the presentation both from the point of view of the presenter and the audience. The Forum talks tended to be longer than the real talks, because the speakers lost track of time. The speakers reported that they were unable to see the usual audience cues of increased restlessness around the time the talks were scheduled to complete. Generally, the speakers had a weak sense of audience reaction, because they were unable to see or hear the usual types of subtle audience responses in the course of a lecture. The experimenters noted that sometimes during the Forum talks audience members did spontaneously chuckle and applaud, but of course neither the speaker nor other audience members were aware of this. Overall, the Forum did not provide sufficient support for the cues that speakers use to monitor and adjust to audiences.

During the course of a face-to-face lecture, a speaker might call on an audience member to help in discussion of a particular point, especially when that audience member was known to have special knowledge of the particular issue. In the case of the Forum, the speaker was reluctant to ask someone in the displayed audience list to contribute in this way, because there was no guarantee that the person was paying attention. They could at that moment have been using some other application on their workstation, and there was no indication of this. More generally, the speakers complained that they did not get the immediate feedback they usually rely upon when answering a particular question for someone, such as seeing them nod or shake their head, or the expression on their face.

The essential point is that, although the audience and speakers are together in a shared system, the space they inhabit together is fragmented between a video representation of the speaker, the audio channel, the lists of audience members, and the workstation environment. There is no unified common space with a metric where participants can vary the distances between one another and become aware of changing spatial relationships and responses to those changes. In particular, although there is visual representation, there is no visual space that all participants simultaneously inhabit. There are no dynamic representations of individuals (except for the presenter) that other individuals can relate and respond to and know that their responses may be experienced by others.

In spite of the current technical shortcomings, shared VEs do offer a common shared visual space, an ideally synchronized audio space, ideally a common haptic space, with ideally multimodal (vision, audio, haptic) personal representations—the avatars.

5.2 Some Characteristics of Shared VEs

The idea of a unified shared space and avatar representations in a shared VE is supported by McQuaid in Nunamaker (1997), in the context of group support systems. In particular, he argues that avatars can give participants a way to judge the focus of attention of others, for example, seeing when two other people are directly communicating. He suggests that avatars can convey information that is given by physical movement in the real world, and that in VEs avatar configurations may take on different social meaning than in everyday reality. For example, sitting on a chair in real life is for comfort and relaxation, and to facilitate certain types of activity. In VEs, there is no inherent need for an avatar to sit, unless this action is directly mapped from that of the real human counterpart. Yet sitting might take on the meaning that the real human counterpart of a seated avatar is currently otherwise engaged and not actually present in the VE. Of course, avatars can also exhibit movements that have a social meaning mapped directly from everyday reality. In the context of a VE lecture, avatars can be made to nod, shake their head, exhibit facial expressions, become fidgety, giving cues to the speaker about audience reactions. The experiment that is the subject of this paper suggests that, even where there are very simple

non-expressive avatars, social conventions may carry over. People can become embarrassed or angry while embodied in very basic avatars, and treat each other's avatars with care. This is a necessary (but not sufficient) condition for social interaction and group working within a shared VE.

5.3 Some Characteristics of Avatars in Shared VEs. Two characteristics in the experimental setup described in this paper can be easily overlooked, but they are actually worth questioning. The first is that the experiment was carried out in a virtual copy of a real laboratory environment, that is, a virtual reflection of a real spatial organization. The second characteristic is that the participants were represented by avatars that had a humanoid resemblance, though with minimal human body functionality. Given the nature of virtual environments, neither of these characteristics are necessary. There is no need to organize virtual space to be anything like real space, and neither is there an intrinsic need for participants to be virtually embodied or embodied with a human appearance. Yet these are characteristics generally employed in shared VEs.

Given that the avatars do inhabit a common space, what characteristics and capabilities should they have? Rich et al. (1994) describe a shared VE system for ''learning by doing,'' where it is possible to explore and learn to use athletic equipment in what is configured as an online community.

There is a virtual body controlled by a user, and also an artificial agent (also embodied as an avatar). Generally, agents (the humans, virtual humans, and other virtual beings such as birds) are able to generate sound and to move as expressive, articulated figures. The human avatars had head, torso, and forearms independently controlled via an actuator system. The goal was to make the users feel as if they were inhabiting a body rather than just operating an animated figure. It was argued that this was achieved by the ability of users to control navigation through hand gestures based on a video recognition system, and to control posture, the changing configuration of different body parts, through a switch box and joystick. No experimental evidence of the outcome was reported.

Benford et al. (1994) discuss extensively the social significance of space as a resource for activity and interaction in VEs. In fact, much of what they say actually concerns the activity of avatars in space, rather than just with space in itself. They argue that continual awareness of others allows people to flexibly modify their own behavior in social situations. (For example, someone heading across the room towards another probably indicates an interest in beginning a verbal communication.) They describe how the use of space—or, rather, the avatars in a meaningful spatial configuration—allows the support or indeed emergence of social mechanisms for control of scarce resources. In a public debate, a line can form around a podium showing to everyone which and how many people are preparing to speak, who indeed the current speaker is (floor control), and the audience reactions (for example, they could all walk out) to an uninteresting talk (something that would be clearly noticed by the speaker, unlike in the Forum system). The authors describe in detail mechanisms that can be provided by the VE to facilitate social interaction above and beyond just copying basic real-life mechanisms, in their notions of aura (the bounding presence of an object), focus (the field in which a user can become aware of others), and nimbus (the field in which others can become aware of the user). They show that social interactions can be seen as a form of negotiation between agents based on their aura, focus, and nimbus fields.

In their discussion, Benford et al. again emphasize the importance of embodiment: how this can provide information about the identity and activity of the participant, how gesture and facial expression can be used for the expression of emotion, and the separation of ''mind'' and ''body''—that is, how the avatar can be used to signal that the real person is currently no longer present in the VE but engaged in other activities (for example, by presenting a ''sleeping'' avatar).

A later paper by Bowers et al. (1996) is an empirical study of what actually happens in an unstructured, small-group, virtual meeting based on the MASSIVE system (Greenhalgh & Benford, 1995). The emphasis was on understanding the relationship between the embodiment of participants through their blocky avatars, and communication issues such as turn taking while talking,

and other aspects of social interaction. The study used conversational analysis to transcribe conversation and was extended to include the simple avatar gestures possible in the system (such as whole body turning or ''ear flapping''). The study found that, in spite of the very limited repertoire of the avatars, the avatars were nevertheless sometimes used to supplement language as an additional mechanism in social interaction. The avatars were not just a means of navigation and representation, but became invested with social meaning, a finding that supports the results of the experiment described in this paper.

The Bowers study also found that participants did move their avatars in socially meaningful ways, for example, to get a better view of those with whom they wished to interact. Participants sought face-to-face communication, even though the use of the audio channel did not actually require this. Although talk was accompanied by the limited repertoire of gestures only to a very limited extent, they did find that there was mutually coordinated movements amongst two or more participants. This suggested that embodiments should support higher-order activities than mere movement, actions of social significance, such as approaches, exchange of glances, turning to, turning away, and other basic expressive actions.

The latter requirements are fully supported by the current study. Recall that, for example, the Green subject found it difficult to know whether their monitoring task was effective because it was hard to tell whether or not they had been noticed by Red. Even more fundamental, it was hard for participants to tell which subject was talking, because there were no accompanying lip movements and no spatialized sound. On this latter point, Rich et al. argued that crude images together with crude audio rendering provides better feedback to participants than better visual or better audio by themselves. They give an example from their system of the avatar walking. Without shadows, it is impossible to tell if the avatar is actually walking along on the floor, and with spatialized sound it is possible to tell only whether the walking noise is coming from the left or right. But, when sound is combined with the visual rendering, the brain integrates the two into a ''foot stepping on the floor'' totality, so that the participant can tell exactly when each foot strikes the floor. In the context of avatars talking, even crude lip movement without spatialized sound is likely to give very strong feedback about who is currently talking.

Vilhjálmsson (1997) provides an elegant approach to avatar functionality in his BodyChat system. He argues that the avatar behavior should be encapsulated into layers, and that at the bottom layer there are what might be described as fundamental or autonomic behaviors that are always happening. This not only gives the sense of ''aliveness'' of each avatar individually, but also enhances the ability of people to interact. So, at the very basic level, avatars visibly breath. Avatars have large black eyes, but with a ''twinkle'' in the center. When one of these avatars looks at you, there is a sense that some presence is there. Each avatar can be in a state of ''being available'' or not being so. When two avatars pass each other while walking, they will involuntary glance towards each other. There is no doubt that each is in the field of view of the other. What happens subsequently is an automatic negotiation based on the state of availability of each. For example, if both are available, then they may stop walking and the potential for a conversation ensues. During a conversation, there are subtle cues (for example, raised eyebrows for questions) and not-so-subtle cues (such as corresponding lip movements). In fact, many of the complaints of the subjects in this paper's experiment would have been overcome through the use of BodyChat, and this without any particularly complex body representations. In a computer-graphics sense, the BodyChat avatars are no more geometrically complex than those available in DIVE. This idea that the avatars systems themselves take care of many autonomic responses, of which in real life we are hardly aware, seems an excellent way forward in the design of personally and socially meaningful embodiments. There is some empirical evidence for this in a study carried out by Thorisson and Cassell (1996), who conclude that ''This supports our claim that what really matters in face-to-face dialogue is, in addition to 'classical information exchange', the supportive behaviors that often have been dismissed as incidental to effective interaction.''

To conclude this section, recall that the current ex-

periment found that the avatars had social significance even with the essentially lifeless avatars that were used. How much more compelling might the experience be with the BodyChat concepts employed?

6 Conclusions

This paper presents an experimental study comparing small-group behavior while carrying out a task in a shared VE and then continuing the task in the real world, where the VE was a virtual copy of the real-world environment. The results of the study, bound as they are by the specific conditions of this experiment, suggest the following hypotheses for future research:

- Leadership capability is enhanced by computational power. In particular, it may be that leadership is enhanced by greater levels of immersion.
- Personal responses to social situations, such as embarrassment and discomfort, can be generated in a shared VE, even though the people involved are experiencing one another through very limited personal avatar representations.
- Even very limited avatars take on social significance, and people have a tendency to be respectful of each other's avatars.
- Presence and copresence are positively associated, although the causality is unknown, and better techniques for eliciting these factors are required.
- Collision detection, enabling avatars to easily obey spatial boundaries (not going through walls) and avoiding one another, must be a crucial component of any shared VE that adopts a conventional spatial representation.

As has been mentioned, this paper reports only a partial study. More groups must be included, the contact time must be extended, the order of presentation varied (some groups should meet first in the real world and then continue in the VE), and the monitoring task (or its equivalent in a future study) should be included in both real and virtual parts (and for some groups only in the real part), rather than just the particular configuration used here. Essentially, this study was conducted to uncover some of the questions that should be asked in a more thorough and extensive experiment, and the results should be considered in that spirit.

Acknowledgments

This research was partially funded through the Inhabit Project, led by the University of Nottingham, and the authors are grateful to Dr. S. Benford for the invitation to take part in Inhabit. The research was also partially funded through the COVEN Collaborative Virtual Environments project, funded by the European Community ACTS framework. We would like to thank Jonathan Howell for his help during the course of the experiment. M. Slater would like to thank Senior Scientist Nat Durlach of the Sensory Communications Group of the Research Laboratory of Electronics at MIT, and other members of the SCG, for providing an ideal environment in which to write this paper during his sabbatical leave from January to April, 1998.

References

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,* Vol B. (pp. 699–704). Elsevier Publisher.

Benford, S., Bowers, J., Fahlen, L. E., Mariani, J., & Rodden, T. (1994). Supporting cooperative work in virtual environments. *The Computer Journal, 37*(8), 653–668.

Benford, S., Greenhalgh, C., Bowers, J., Snowdon, D., & Fahlen, L. E. (1995). User embodiment in collaborative virtual environments. *Human Factors in Computing Systems, CHI'95 Conference*, 242–249.

Bowers, J., Pycock, J., & O'Brien, J. (1996). Talk and embodiment in collaborative virtual environments. *CHI'96 Electronic Proceedings.* http: / / www.acm.org / sigchi / chi96 / proceedings / papers / Bowers / jb_txt.htm.

Carlsson, C., & Hagsand, O. (1993). DIVE: A multi-user virtual reality system. *IEEE Virtual Reality Annual International Symposium*, 394–400.

Greenhalgh, C. M., & Benford, S. D. (1995). MASSIVE: A Virtual Reality System for Tele-conferencing. *ACM Transac-*

tions on Computer Human Interfaces (TOCHI), 2(3), 239–261.

Hardman, V., Sasse, M-A., Handley, M., & Watson, A. (1995). Reliable audio for use over the internet. *Proc. INET'95, International Networking Conference,* 171–178.

Held, R. M., & Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators and Virtual Environments, 1*(4), 109–112.

Kollock, P. (1999). The economies of online cooperation: Gifts and public goods in cyberspace. In P. Kollock & M. Smith (Eds.), *Communities in Cyberspace* (pp. 220–239). London: Routledge.

Isaacs, E. A., Morris, T., Rodriguez, T. K., Tang, J. C. (1995). A comparison of face-to-face and distributed presentations. *Human Factors in Computing Systems, CHI'95 Conference Proceedings,* 354–361.

Leigh, J., & Johnson, A. E. (1996). Supporting transcontinental collaborative work in persistent virtual environments. *IEEE Computer Graphics and Applications, 16*(4) 47–51.

Major, S. B., Stytz, M. R., & Wells, W. D. (1997). Requirements and implementation of multiple modes of collaboration and cooperation within distributed virtual environments, military, government, and aerospace simulation. In M. J. Chinni (Ed.), *Proceedings of the 1997 Simulation MultiConference,* (pp. 175–180).

Macedonia, M. R., & Noll, S. (1997). A transatlantic research and development environment. *IEEE Computer Graphics and Applications, 17*(2), 76–82.

Nunamaker Jr, J. F. (1997). Future research in group support systems: Needs, some questions and possible directions. *Int. J. Human-Computer Studies, 47,* 357–385.

Rich, C., Waters, R. C., Schabes, Y., Freeman, W. T., Torrance, M. C., Golding, A. R., & Roth, M. (1994). An animated on-line community with artificial agents. *IEEE Multimedia, 1*(4), 32–42.

Schroeder, R. (1997). Networked worlds: Social aspects of multi-user virtual reality technology. *Sociological Research Online, 2*(4), http: / / www.socresonline.org.uk / socresonline / 2 / 4 / 5.html.

Schiano, D. (1999). Lessons from LambdaMOO: A social, text-based virtual environment. *Presence: Teleoperators and Virtual Environments, 8*(2), 127–139.

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

Thorisson, K., & Cassell, J. (1996). Why put an agent in a human body: The importance of communicative feedback in human-humanoid dialogue. *Lifelike Computer Characters '96,* 44–45 (abstract).

Vilhjálmsson, H. H. (1997). *Autonomous communicative behaviors in avatars,* master's thesis, Learning and Common Sense Section, Media Laboratory, MIT, http: / / lcs.www.media.mit.edu / people / hannes / msthesis / .