

# LONGITUDINAL STUDY OF HABITUATION AND PARTICIPATORY DESIGN FOR MULTI-USER SHARED VIRTUAL ENVIRONMENTS

N. Ashraf<sup>1</sup>, J. Anderson<sup>1</sup>, C. Douthet<sup>1</sup>, I. McKay<sup>1</sup>, M. Jack<sup>1</sup> and P. Platt<sup>2</sup>

<sup>1</sup>Centre for Communication Interface Research, University of Edinburgh, Scotland

<sup>2</sup>BT Adastral Park, Ipswich, England

## ABSTRACT

This paper details the results of a longitudinal study of user habituation, usage and involvement with a shared virtual 3D environment acting as a meeting space. The study involved investigation of the effectiveness of a range of design features which were included to enhance communication, discussion and social interaction among a group of four users of the shared space. The users took part in six sessions over a period of seven weeks. The paper focuses on usage of the shared space and details user involvement in the design process in terms of reactions to avatar personalisation; avatar life signs, gestures and navigation control; the means for identifying who is talking; and symbolic acting by avatars.

The results indicate the importance of key features for the design of virtual environments. Participants wanted to identify their own protocols for turn-taking in conversation and they wanted simple gesture control. For example, one-click visual buttons for the selection of gestures were preferred to pull-down menus. It took users five or six sessions to complete the participatory design process at which point they were totally comfortable with the use of the virtual meeting space. Finally, symbolic acting was shown to be a viable addition to the shared space to assist group dynamics.

## 1. Introduction

Extensive research has been carried out in the field of 3D virtual environments. The main emphasis of early published work was largely on technical issues such as improvement of run-time performance [Hubbold et al, 1993], the provision of techniques for user viewpoint control [Drucker and Zeltzer, 1994] and improving the understanding of users' perceptual requirements during interaction, e.g. [Rushton and Wann, 1993], [Carr and England, 1993]. The role and importance of 'presence' as a social and communication factor in immersive virtual environments has been extensively researched [Held and Durlach, 1993], [Welsh et al. 1996], [Hendrix and Barfield, 1996a & 1996b], [Sheridan, 1996], [Slater and Wilbur, 1995], [Tromp, 1995].

Recent research in virtual reality conferencing which uses immersive technology or a desktop window onto a 3D virtual environment conventionally assumes an *avatar-as-proxy* model of the relationships between conferees. In this model, participants control avatars which transmit information about their owner (for example, position in the world, orientation, and gestures) by proxy to others using the system. The avatars are therefore the principal conduits for communication between conferees within the shared space. The main problem with this as a design model arises with respect to social interaction between participants. It has been demonstrated, for example, that the absence in shared spaces of signals which are readily available in face-to-face contact to indicate changes of attitude can create communication difficulties [Durlach and Mavor, 1995]. This effect is particularly strong where changes

in 'sense of self' are concerned, an important feature of all interpersonal communication which has long been known to be dependent on social interaction [Mead, 1934]. Guided by the avatar-as-proxy model, research into improving the social communicative capabilities of shared space enhanced audio conferencing has generally concentrated on improving the visual realism of the avatars or increasing the complexity of the mapping between users' actions and avatar responses [Biocca, 1995]. Users strongly associate other people's body states with their mental states [Damasio, 1994], a factor which has led researchers to view the body (or avatar) as *the* physical representation of a person's thought [Johnson, 1987], [Lakhoff & Johnson, 1980], [Lakhoff, 1987].

[Bowers et al, 1996] discussed the design of virtual worlds in relation to the extent to which they encourage social interaction. This work demonstrated how objects in virtual environments (such as tables) can act as focusing devices for participants by giving them the means to co-ordinate their conversation, view each other and organise their mutual body orientations. The work also emphasised the important role played by avatar gestures in turn-taking. [Mortlock et al, 1997] found a range of essential features that need to be considered in the design of virtual conferencing systems. These included the availability of avatar personalisation so that conferees can easily recognise each other and the provision of non-intrusive gesture controls so that conversants can concentrate on the task in-hand and not the underlying technology. The work by Mortlock et al serves as the point of departure for the research presented in this paper, which seeks to extend those results and undertakes a longitudinal study and more detailed investigation of the fine control of personalisation and gesture.

This paper presents results from a longitudinal investigation in the form of a observational study to explore usage effects for a group of four users engaged in a series of communication tasks using a shared space virtual conferencing environment. A participatory design process was followed to allow users to contribute to the definition of features in the shared space which would support and enable effective and efficient communication. The issues of avatar personalisation; effectiveness of avatar life signs; means of gesture and navigation control by the user; talker identification, turn-taking and symbolic acting were all examined from a user perspective. Modifications to the shared space and its avatars were implemented and assessed on the basis of feedback from the users and from observations made during the study sessions.

## 2. Design of the Study

In order to evaluate users' real needs in a shared virtual conferencing space, a longitudinal study was carried out which involved four participants using a customised shared space environment over six sessions. Each session consisted of some forty minutes collaborative use of the shared space followed by a thirty minute group discussion to elicit participants' views on current features and suggestions for modifications and improvements to the environment and the avatars.

Four participants took part in the study (two male, two female, all between twenty and thirty years of age). Three of the participants knew each other; one was unknown to the other three at the start of the study. The participants were post-graduate students in non-computing disciplines. All four had basic computing skills although they had no experience with shared spaces, conferencing systems or avatars. One of the male participants was Canadian, the other participants were British; all were native speakers of English. Participants were fully aware that they were being observed and that their conversation was being recorded.

Five offices, each containing a PC workstation connected to a central server, were allocated for the study, one for each participant and one for a researcher who accessed the shared space during the sessions as an invisible (and silent) avatar in order to monitor interactions and usage. The researcher could eavesdrop on the discussions via a one-way audio link. The only means of communication between participants was via the shared space software.

At the start of each session, the evolved design of the virtual meeting space was explained to the participants as a group. Participants were reminded that, where possible, their feedback comments would be used to modify the virtual environment for the following sessions. Participants were then asked to read editorials from a selection of that morning's newspapers before logging into the shared space. Once in the shared space their task was to discuss the editorials and reach a consensus on the most interesting or important story to be carried forward for inclusion in 'a Sunday newspaper'. This task was chosen because it did not require the participants to have any specialist knowledge. This task was repeated (with different editorial materials) for five of the six sessions. For the sixth session the task was enriched to explore issues raised by the private messaging facilities that had been requested. This modified task involved participants working in pairs to present and discuss their point of view, based on one topical issue.

An observation log was kept during each session by the researcher which included the types and variety of interactions that occurred, and the content of any discussions relating to the technology being used. This information was then used by the researcher to structure the group discussion which immediately followed use of the virtual meeting space.

### 3. Shared Space Resources

Before starting the longitudinal study, extensive work was undertaken to create a starting point for the shared space design to be offered to the participants. Initial considerations included the choice of software platform, focusing on the need for the research team to be able to rapidly introduce features, enhancements and modification called for by participants, and the set of basic communication facilities to be included in the shared space.

#### 3.1 SYSTEM ARCHITECTURE

The custom-built virtual conferencing system used in the study employed a public domain audio conferencing tool (*RAT*<sup>1</sup>) in combination with a multi-user VR client-server package. The latter was implemented using *DeepMatrix*<sup>2</sup>, an open-source Java-based application which operates in conjunction with a standard VRML (Virtual Reality Modelling Language) plug-in<sup>3</sup>. *RAT* ran invisibly in the background during the study sessions so that the participants were only involved with the shared space interface. The availability of public domain and open-source software for these resources allowed rapid modification of features and facilities in the shared space in response to feedback from participants. All visual events within the virtual environment were logged by the software for later analysis.

---

<sup>1</sup> RAT (Robust Audio Tool) V3 for Windows. Set for audio communication using DVI encoding with a 40ms duration. <http://www-mice.cs.ucl.ac.uk/multimedia/software>

<sup>2</sup> <http://www.geometrek.com/products/deepmatrix.html>

<sup>3</sup> The *Contact* plug-in, developed by blaxxun. <http://www.blaxxun.com>

The hardware for the study consisted of four client PCs (one for each participant), one client PC for the observer, and a server. Participants used microphone headsets.

**Server configuration:** Intel dual Pentium II 250MHz; PCI network card (10Mbps); 64MB RAM; Microsoft Windows NT 4.0 Server installed with Microsoft *Internet Information Server* 3.0, Microsoft *SQL Server* 6.5 and *DeepMatrix* software (including Java classes for both server and client).

**Client configuration:** Intel Pentium II 350MHz; PCI network card (10Mbps); 128MB RAM; 128-bit graphics card with 8MB video memory and 3D hardware acceleration; Microsoft Windows NT 4.0 Client installed with Netscape *Communicator* 4.5, blaxxun *Contact* 4.0 VRML browser plug-in and *RAT* audio conferencing tool.

### 3.2 ESTABLISHING THE FEATURES OF THE SHARED SPACE

Before the start of the study an extensive set of tests was carried out on an informal basis to establish general user preferences for the type of environment and layout to be used in the 3D virtual space. For these informal tests, two groups each of four volunteer users (students and researchers) were invited to explore a wide selection of different shared spaces. Usage was observed and their opinions were obtained in a post test interview. The general conclusions drawn from the informal testing led to the following set of facilities being included in the initial configuration for the study:

- A meeting room based on a central table with chairs was identified as most appropriate for the types of tasks involved in the study because it was a setting with which the users were familiar and found easy to use;
- In order to avoid constraining the seating positions for users and also to give a realistic sense of space around the conference table, eight chairs were provided (twice the number of participants);
- To further improve the participants' spatial awareness, various furnishings were added such as pictures, a water dispenser and potted plants. For the same reason, the chairs around the central table were in different colours;
- To allow both close-up view of an avatar and a wide-angle view of the room, zoom-in and zoom-out controls were added to the standard navigation controls;
- To allow each user to view their own avatar and be aware of its representation in the environment, a small self-image was provided at the bottom right hand corner of the 3D view;
- To facilitate navigation around the room, users could select which chair they wished to occupy by clicking directly on it using the mouse and their avatar would then seat itself in that chair;
- Since users need to be able to use either the mouse or the arrow keys on the keyboard to navigate the shared space, both of these were enabled;
- To enhance realism, mouth animation was included as a useful and natural aid to talker identification, with the amplitude of the speech determining the amount of mouth movement;
- Basic avatar life signs such as breathing and blinking were added;
- Users were allowed to choose from a limited range of avatar types, based on the most basic criteria such as gender, formal or informal clothing style and colour (e.g. Mr Blue, formal);
- Avatars were made to display walking motions (as opposed to gliding) to increase their perceived naturalness;
- A range of simple gestures (agree, disagree, greeting) was made available to users.

These features were added by means of modifications to the *DeepMatrix* client applet (e.g. walking motions, avatar self-view); through custom-built helper applications (running invisibly in the background) which could monitor and adjust the audio input/output to RAT and pass relevant information (e.g. mouth movement) to and from the *DeepMatrix* client; and by Script nodes within the VRML implementation of the environment and avatars. The avatars themselves were compliant with the H-ANIM specification to facilitate import and export.

Figure 1 shows the design of the shared meeting space as it was at the start of the study. The zoom controls are at that bottom centre and the avatar self-view is at the bottom right. User names appear above each avatar and a microphone object used to aid turn-taking is visible on the table in front of the male avatar which is shown making a 'greeting' gesture.



*Figure 1: Shared Meeting Space used as Point of Departure for the Study*

#### **4. Evolution of the Meeting Space Design**

The results described in the paper are based on observations and discussions made during the six design discussion sessions. The key features of these sessions are summarised here to aid understanding of the sequence of changes to the environment and avatars on which the results and conclusions are based.

**Session 1.** In the first session eight chairs were offered in the conference room for the four users. A moveable microphone object was also available on the table, to be used as an aid to turn-taking. The avatars were given breathing and blinking as life signs. A pull-down menu for three gestures ('agree', 'disagree', 'greeting') was included, as was direct avatar 'transport' to a chair by mouse-click on the chair. An avatar self-view was provided and the audio connection was continually active whether or not the participant was logged-in to the shared space.

**Session II.** Based on user suggestions, the microphone object was removed (see Section 5.5) as were four of the chairs (see Section 5.2). All avatar life signs were disabled but the mouth movements to indicate the talker were made more pronounced by totally closing the mouth on each cycle (see Section 5.3).

**Session III.** Avatar life signs were re-enabled for this session but with the breathing motion made less pronounced - reduced from 5% chest expansion to 1% (see Section 5.3). Automated arm and hand gesticulations were added to complement the role of the mouth movements for talker identification (see Section 5.5) and an alternative means of gesture invocation was added as a pop-up menu enabled by clicking on the avatar self-view (see Section 5.3). Enhanced avatar personalisation (carried out prior to logging in) was added to allow selection of hair colour and clothing colours, as well as to allow preview of the avatar prior to entering the shared space (see Section 5.2).

**Session IV.** The four extra chairs were re-introduced (see Section 5.1). The avatar breathing motion was increased to a chest expansion of 2% (see Section 5.3), and avatar personalisation options to allow choice of build, height and hair style were introduced (see Section 5.2). The available gestures were re-labelled ('yes', 'no', 'greeting') and the pull-down menu was replaced by a row of three labelled buttons (see Section 5.4).

**Session V.** Avatar personalisation was extended to include choice of skin tone (see Section 5.2). An automatic head-turning facility was introduced to make a user's avatar turn to face whichever avatars were talking (see Section 5.5). This facility could be activated and de-activated by the user at any time during the session. An additional hand raising gesture (labelled 'hand up') which had been suggested by participants as a potential aid to turn-taking was added (see Section 5.5). The waving arm gesture was relabelled 'bye' and a new (open-armed) gesture was introduced (see Section 5.5). At the request of participants, a private text-based messaging facility was introduced together with associated symbolic actions to represent the writing, reading, saving or discarding of messages.

**Session VI.** A second room was added to act as a pre-meeting ante-room (see Section 5.1). The gesture buttons were re-labelled with more generic names ('nod', 'shake', 'hand up', 'bye', 'shrug'). The audio connection (input and output) at each client was modified to be enabled only when the user was logged on.

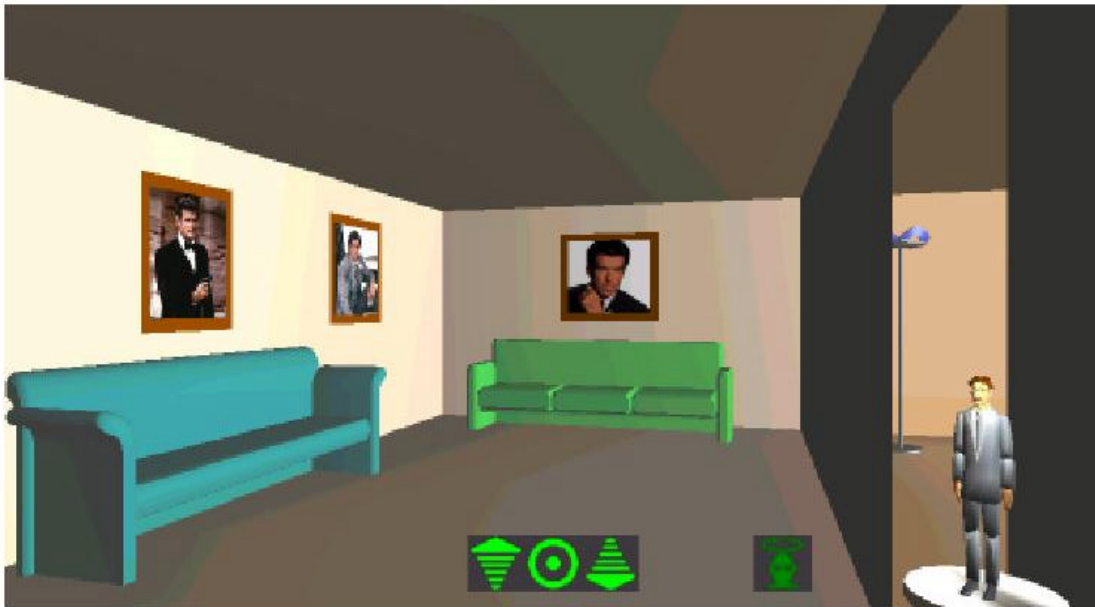
By the end of Session VI the participants felt that they had arrived at a working set of features which supported their communication needs within this particular virtual shared space. The study was therefore concluded at that point.

## **5. Discussion of the Evolving Meeting Space Design**

In the following sections, the changes recommended by participants for each of the various components in the shared meeting space as it evolved over the six sessions are discussed in detail.

### 5.1 CHANGES TO THE ENVIRONMENT

The decision to use a meeting room scenario with a conference table, chairs, pictures and room furnishing was continually re-visited and repeatedly supported by the participants throughout the sessions. Only in the last session was this environment significantly modified with the introduction of an ante-room as a reception area (Figure 2). However, a number of significant modifications to the meeting room facilities and contents were made in response to participants' comments.



*Figure 2. Ante-room Designed as a Pre-meeting Reception Space*

Participants suggested removing four of the eight chairs used in the first session. During this first session a problem was experienced by all participants when attempting to look at the avatar that was talking, since only certain seating arrangements allowed a user to view all other avatars simultaneously. To overcome this problem, participants constantly moved their avatars to sit opposite the avatar of the person they believed was talking. They would then use the zoom facility to see the mouth movements. This led to a great deal of confusion since everyone tried to move seats each time a new person began talking.

Various solutions to this problem were considered. Significant changes to the layout and design of the shared space were rejected because participants felt the realism offered by the table and chairs layout was most appropriate for their discussion task. Automatic software manipulation of the relative positions of the other avatars to make them appear within each user's individual view was rejected as 'unrealistic' and would ultimately cause breakdown of the shared space metaphor since different users would perceive the spatial relationships differently. The solution considered most viable by the participants was to make the table smaller and remove the extra chairs. Consequently, for the second session, four of the eight chairs were removed. The table size was not in fact changed although the participants voiced their belief that it had been reduced in size, perhaps because each user could now see all the other avatars in their field of view.

When participants had been introduced to the use of the keyboard, as an alternative to the mouse, for control of head and body rotation in Session III, and had become familiarised with this feature, they felt they would now be more able to control their viewpoint to look at all avatar positions. In consequence, they suggested that for Session IV, all eight chairs should be restored (to allow for cases where more than four simultaneous users took part). Options such as limiting the arrangement to one chair per logged-in participant, or moving to different rooms with seating capacity dependent on the numbers logged in, were rejected as impractical. From Sessions IV onwards, the participants' skills in using the zoom and head-turn controls developed significantly. This was coupled to their expressed wish to maintain realism since "*in the real world it would be necessary in any case to turn to see the person sitting next to you*".

After Session V, participants suggested the introduction of an entry room. In general, they did not like being 'beamed' directly into the meeting room and the use of an ante-room would also give them a place to gather before the meeting and discussion began. This was implemented for Session VI (Figure 2).

It was interesting to note that on one occasion a participant walked out of the conference room into the ante-room before logging out because she "*wanted to get away from everyone because the meeting had been so intense*". This highlights the power that the shared space metaphor had developed over the five sessions and the high degree of association that participants had with it.

## 5.2 EVOLUTION OF AVATAR PERSONALISATION

Selective realism was a key topic of comment for users throughout the study, especially with respect to the human-like appearance of the avatars. In all the discussions, participants specified that whilst they wanted to be able to personalise their avatars, they still required a certain degree of generalisation, a form of mask, something they were able to '*hide behind*'. This was achieved by limiting the personalisation options to approximate and relative category descriptors (e.g. 'medium', 'taller' or 'shorter' height). Participants felt that such an approach would serve to avoid offending users. As one participant explained, "*if you were fat you could pick a larger build but it wouldn't really look exactly like you – but it doesn't matter... the other people in the meeting aren't going to think you're pretending you're something you're not – they're going to know its you*".

The facility to map individuals' faces onto their avatars was discussed, but all participants agreed that this would not be appropriate because it would be too accurate a level of customisation and would exceed the bounds of the selective realism they judged to be so important. In any case, given the graphical representation of the avatars, a photo-realistic facial rendering "*wouldn't fit with the bodies*".

The degree to which participants started to relate to their avatar became very evident from Session III onwards. In earlier sessions the choice of individual avatars was restricted to selection from a pre-defined fixed list using a pull-down menu available during login. Although the user's name appeared over the head of each avatar, this was judged to be inadequate in satisfactorily personalise the portrayal of the user. In the first session, when the audio connection was continually present, users periodically logged out of the shared space and re-entered with a different avatar. This suggested that at this point in the study, users had not developed any strong association of the other users with their avatar representation.



The role of avatar personalisation was discussed at length after Session II. Participants decided that they needed to be able to add more features to the personalisation of their avatar with representations such as skin tone, hair colour and style, spectacles and facial hair for the men.

Participants were given the facility to select hair colour and clothing colours in Session III. It is interesting to note how such small changes in the personalising of the avatars led participants to perceive a much greater degree of personalisation than had been actually implemented. For example, although participants did not have the facility to change body build in Session III, they insisted that the avatar for one specific user was taller than the others although this was not in fact the case. During this session, the two male avatars stood next to each other to compare their heights and at one stage even went as far as to race around the virtual space, finally concluding that one of the avatars was not only taller, but *“could run faster because its legs were longer”*. The avatar judged to be tallest in fact belonged to the tallest participant, a clear indication that users transferred real-world knowledge to their interpretation of the avatar.

Reacting to the participants wishes for more personalisation options, in Session IV extra features were made available in terms of avatar height and build as well as choice of hair style and length. Another option was provided for selection of facial hair (moustache, beards) which only applied to male avatars. To avoid the situation where participants were logging in multiple times to view different avatar selections, for Session IV the login screen was enhanced to include an option to preview the personalised avatar before entering the shared space.

Further evidence of the degree to which personalisation improved the association of users with their avatars appeared in Session IV when one of the users logged out during the meeting in order to modify his avatar - as all of the users had done in previous sessions. The other users became furious with this action and persisted with their annoyance even through to the discussion period. When he asked why things were different from previous sessions, he was told *“It’s just rude because we know who you are now”*.

Participants generally chose to personalise their avatars with a degree of selective realism in relation to their own physical characteristics, even going so far as to select clothing colour appropriate for that day. However, on the few occasions when participants opted for a clear separation between themselves and their avatars the other participants generally were happy to go along with a fiction created by another user, for example on one occasion when one user chose a different name the others were willing to use the alias. Figure 3 (next page) shows the user interface for avatar personalisation which was used prior to login.

### 5.3 ADJUSTMENT OF AVATAR LIFE SIGNS

The life signs included for the avatars were breathing motions and a random blinking of the eyes. The breathing life sign used in Session I was considered by participants to be over-emphasised, and caused participants to focus on discussing the avatars rather than carrying out the task; indeed, the motion was so exaggerated that participants were not even convinced that it was intended to portray breathing. The effect was felt to be so pronounced that it looked like a gesture (*“they’re shrugging their shoulders”*). In consequence, for Session II all avatar life signs were removed in order to assess user reactions to their absence. It was soon observed that the lack of life signs during Session II had a dramatic detrimental impact on communication in the shared space. At one stage during Session II, the avatars were seated around the table with no-one talking. As a result there were no gestures or lip movements and participants remarked that *“everyone looked frozen and it wasn’t normal”*. Later discussion confirmed the view that avatars did need some form of life signs in order to appear more realistic.

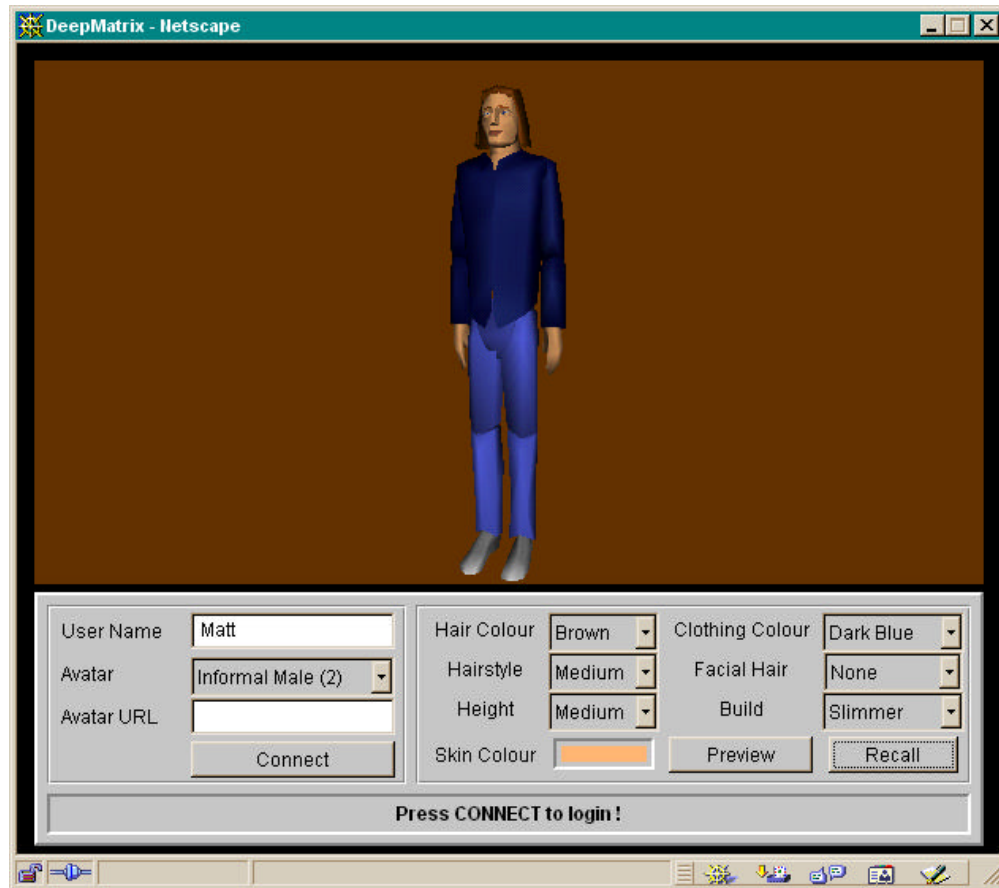


Figure 3: Avatar Personalisation User Interface

Following the re-introduction of avatar life signs in Session III (and onwards), but with breathing much reduced in emphasis, participants then started to notice the presence of the random blinking. This was first noticed by one participant who interpreted it as a facial gesture. It was only when the whole group came to realise that all of the avatars were blinking in this way that they stopped discussing it. When this issue was raised in the discussion group the participants responded “yes, we noticed the eyes blinking but then we just forgot about it because it’s normal isn’t it?”. The video clip [lifesigns.avi](#) shows the life signs used with the avatars.

#### 5.4 OPTIONS FOR AVATAR GESTURE CONTROL

In Sessions I and II, avatar gestures were controlled solely by use of a pull-down menu in the control panel. Three key gestures (‘agree’, ‘disagree’ and ‘greeting’) were identified as important during the pre-study tests. This testing had included pull-down menus with 23 different gestures ranging from arm-folding to toe-tapping and head-scratching, but it was rapidly identified that the real issue to be addressed was the means to activate gestures as opposed to the range of gestures offered.

In Session III, an additional method of gesture activation was implemented: a pop-up menu obtained by clicking on the avatar self-view. This new option was designed to allow users to keep their

focus within the 3D view; it was anticipated that this would facilitate and encourage the use of gestures. However, this did not prove to be the case. Participants preferred to use the pull-down menu from the 2D control panel despite the difficulties they experienced with it. After extended discussion on this issue, a third alternative - the use of labelled buttons in the control panel - was proposed. These were added to the 2D control panel where they could be activated by either mouse click or by the function keys.

The labels used for gestures evolved during the study from their original semantic descriptors (e.g. 'agree') to a more literal set of gesture descriptors (e.g. 'nod'). This was actively requested by the participants because it allowed them to apply their own interpretations to each gesture, rather than being restricted to an interpretation implied by the label.

In the discussion after Session IV, participants decided that the hand-waving gesture labelled 'greeting' was actually more appropriate as a 'goodbye' gesture. The button was consequently re-labelled 'bye' and a fourth gesture button labelled 'hello' (involving an open-armed welcoming gesture) was also introduced for Session V. Following its use in Session V, the 'hello' gesture was re-labelled 'shrug' since participants felt that this was a more appropriate description of how they used the gesture.

Finally, a fifth gesture, consisting of a raised hand, (labelled 'hand up') was introduced in Session V as an aid to turn-taking. This proved to be very successful in both Sessions V and VI.

#### 5.5. TALKER IDENTIFICATION AND TURN-TAKING

In each of the six study sessions, one of the key features discussed was the need for users to be able to identify which avatar was talking, both to support effective social interaction and also to permit efficient control of turn-taking. In the first session, a table-mounted moveable microphone object was included within the space for use as a token to indicate the current talker (see Figure 1). The microphone had no effect on the audio link, and was intended to be used solely as a marker of self-selection or other-selection during turn-taking manoeuvres. After the first session however, it was unanimously agreed by the participants that the microphone was not contributing to turn-taking in the way intended. Participants considered that they had spent too much time moving the microphone around the table and that it was difficult to identify who was trying to move the microphone. Also, given that their discussion task in the shared space was informal (i.e. there was no chairperson to control turn-taking), the microphone did not serve effectively as a device for selecting the next speaker. After discussion, participants agreed that they did not wish to have the use of the microphone object imposed on them; they felt it was *"better to come up with our own protocols of dealing with who was speaking and who wished to speak next"*. The microphone was therefore removed from Session II onwards although, interestingly, participants later asked for the introduction of a turn-taking ('request to speak') gesture. This was implemented in Session V with the addition of a hand-raising gesture via a fifth button labelled 'hand up'. The hand-raising gesture was found to be particularly useful in attracting other participants' attention. At one stage, it was used very effectively by one participant to draw attention to the fact that she had something to say.

The avatar mouth movements had been included on the assumption that they would contribute to realism and also give a clear indication of who was talking. Although the mouth movements were very effective in supporting realism from a zoomed-in viewpoint, they failed in their talker-identification role because the mouth motions were not obvious enough from normal viewpoint.

Participants felt that a more extensive scheme involving gesticulation by a talking avatar might be a more natural and effective form of identifying a talker. In Session III, the avatars were enhanced so that hand and arm gesticulations were activated as soon as the user started talking. However, the absence of a delay resulted in problems for the participants. One participant, who generally had a great deal to say, did

not seem to be involved at all. When asked about this, she explained: “*I didn’t feel like I had any authority to give an opinion during the discussion because I had no control over my own gestures*”. Since it was clear that the gesticulation was judged to be too intrusive, the software was modified (for Session IV) to include a five second delay after talking started. Gesticulation ceased after two seconds of silence by the talker. The heightened level of acceptance of these gesticulations was evident in Session IV when, during a heated discussion, one of the more vocal participants was thought to be gesticulating more emphatically as his voice became louder. This was not in fact the case since gesticulations were not dependent on the loudness of the voice. This finding also highlights the fact that participants were increasingly tending to associate avatars and their owners. This tendency was further evidenced in Session IV when a participant who was talking became angry with another and shouted, “*look at me when I’m talking to you!*” During the subsequent group discussion it emerged that, contrary to previous sessions, participants were now assuming that if an avatar was not looking at them this meant that the user was not listening to what was being discussed. The participant who was not listening, and whose avatar was turned away, had in fact been using this stance as an interruption device. He explained that “*I didn’t like her going on and on and couldn’t get her to stop so I just looked away*”.

As a further aid to talker identification, an automated head-turning facility was introduced in Session V. The mechanism used the loudness and duration of each talker’s audio signal to turn the viewpoint of each user towards the talking avatar(s). The automated head-turning option could be activated or de-activated during a session since it was evident from previous sessions that participants needed to feel in control of their avatars. Indeed, it was noted that users did de-activate during the session in order to return to manual control of the avatar rotation as a mechanism for turn-taking. In Figure 2 the automated head rotation on/off toggle is visible at the bottom centre right.

#### 5.6. THE ROLE OF SYMBOLIC ACTING

The need for private communication facilities was requested periodically in discussion sessions. Two key requirements identified by participants were the need to know when such a facility was being used, and by whom. To accommodate this request a private text messaging facility was introduced in Session IV that allowed the user to send a text message to one or more of the other participants. Received messages could be read immediately, saved for later or discarded. Use of this facility was signalled to participants by the use of symbolic acting by the avatars involved. The symbolic acting consisted of sequences of avatar animations suggestive of writing (hand moves over a sheet of paper), reading (unfolds paper and head moves left-right), saving (folds paper and slips into pocket) or discarding (crumples paper and throws over shoulder).

The symbolic acting sequences are included as the following video clips:

[writing\\_textmessage.avi](#)

[reading\\_textmessage.avi](#)

[saving\\_textmessage.avi](#)

[discarding\\_textmessage.avi](#)

Participants found the private text messaging facility easy to use and appreciated the presence of the associated symbolic acting. This is evidenced by the fact that when private text messaging was originally introduced in Session V, the symbolic acting sequences were not invoked in cases of ‘public’ message passing and this caused confusion since there was no symbolic acting associated with the sender or receiver. For Session VI this problem was overcome by forcing the specification of message recipients so that symbolic acting was always invoked. Such confusions serve to highlight the importance of the feedback function of symbolic acting.

### 5.7. FINAL ENGAGEMENT WITH THE SOFTWARE

By Session V, a convergence of user skills and the mature features of the shared space was very noticeable. Participants were sufficiently habituated and skilled in their use of the shared space and avatar controls that the avatars' gestures were now being used to enhance and enable on-going conversations relating to the task. The participants were now more concerned with using the gestures to support the task rather than exploring the gestures themselves. They remarked that they felt that for the first time they were carrying out the task asked of them and that this was more important than using the software for its own sake. They commented on the fact that Session VI was the first session in which "we were fully confident that we were using the features of the shared space to directly aid the task."

An example of a shared space meeting with four users is shown in Figure 4 and included as the video clip [sample\\_session.avi](#). Figure 4 shows the control panel for the system with its line of gesture control buttons, left-to-right: 'nod', 'shake', 'hand-up', 'bye' and 'shrug'.



Figure 4: Shared Space Session with Four Users

## 6. Conclusions

The results of this study indicate the importance of the following key design features for virtual environments when used for informal meetings:

- *Selective realism.* Users demanded a degree of selective realism from the shared space design. On the one hand users sought a realistic, conference room metaphor for their meeting space; on the other hand they favoured only limited realism in the choice of their self-representation (e.g. choice of taller, shorter or medium height for avatar).
- *Being in control.* Users wanted to identify their own protocols for turn-taking in conversation. Users discarded a microphone object intended for control of discussions and turn-taking in favour of a raised hand gesture.
- *Time to habituate.* It took users five or six sessions to complete the participatory design process and to become totally comfortable with use of the virtual meeting space.
- *Simple gesture control.* One-click, visual buttons for selection of gestures were preferred over pull-down menus. Labelling with a gesture description ('nod') rather than an interpretation ('agree' or 'yes') was also preferred.
- *Symbolic acting.* This feature can be a viable addition to assist group dynamics. Symbolic acting, implemented as an automatic sequence of avatar motions symbolising private text messaging events, added visual portrayal of important relative states for the group.

## 7. Acknowledgements

The authors wish to acknowledge the financial support for this research from BT under their Strategic University Research Initiative. The work has benefited from helpful discussions with Dr. John Foster of the University of Edinburgh and with members of the research staff at BT Adastral Park.

## REFERENCES

- Biocca, F. (1995). "Intelligence augmentation: The vision inside virtual reality" in B. Gorayska & J. Mey (Eds.), *Cognitive Technology*. Amsterdam: North Holland.
- Bowers J., O'Brien J. and Pycock J. (1996). Talk and embodiment in collaborative virtual environments. In: *Computer-Human Interaction '96*, ACM Press, 1996.
- Carr K. T, and England R. D. (1993). The role of realism in virtual reality. In: *Virtual Reality International '93: third annual conference*, London 1993. Proceedings. London: Meckler, pp. 24-33.
- Damasio, A. (1994). *Descartes' error: Emotion, reason, and the brain*. New York: Grosset/Putnam.

Drucker S. M. and Zeltzer D. (1994). Intelligent camera control in a virtual environment. In: Graphics Interface '94 conference, Banff, Alberta, 1994. Proceedings. Toronto: Canadian Information Processing Society, pp. 190-199.

Durlach, N.I. and Mavor, A.S (Eds). (1995). *Virtual Reality: scientific and technical challenges*. National Research Council, National Academy Press, ISBN 0-309-05135-5.

Held R. and Durlach N. (1993). Telepresence, time delay and adaptation. In: Pictorial Communication in virtual and real environments. Ed. by S. R. Ellis, M. Kaiser and A. J. Grunwald. (London: Taylor and Francis).

Hendrix C. and Barfield W. (1996a). Presence within virtual environments as a function of visual display parameters. *Presence: teleoperators and virtual environments* 5(3): 274-289.

Hendrix C. and Barfield W. (1996b). The sense of presence within auditory virtual environments. *Presence: teleoperators and virtual environments* 5(3): 290-301.

Hubbold R., Murta A., West A. and Howard T. (1993). Design issues for virtual reality systems. Paper presented at First Eurographics Workshop on Virtual Environments, Barcelona, September 1993.

Johnson, M. (1987). *The body in the mind*. Chicago: University of Chicago Press.

Lakhoff, G. (1987). *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

Lakhoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Mead, G.H. (1934). *Mind, Self and Society*. Chicago: University of Chicago Press.

Mortlock A., Machin D., McConnell S. and Sheppard P. (1997). Virtual conferencing. *BT Conferencing Journal*, 14, pp. 120-129

Rushton S. and Wann J. (1993). Problems in perception and action in virtual worlds. In: *Virtual Reality International '93: third annual conference*. London 1993. Proceedings. London: Meckler, pp. 43-55

Sheridan T. B. (1996). Further musings on the psychophysics of presence. *Presence: teleoperators and virtual environments* 5(2): 241-246.

Slater M. and Wilbur S. (1995). Through the looking glass world of presence: FIVE: a Framework for Immersive Virtual Environments. In: *FIVE '95: Framework for Immersive Virtual Environments*, London 1995. Proceedings. London: University of London, pp.11-26

Tromp J. (1995). Presence, telepresence and immersion: the cognitive factors of embodiment and interaction in virtual environments. In: FIVE '95: Framework for Immersive Virtual Environments, London 1995. Proceedings. London: University of London, pp. 39-51

Welch R. B., Blackmont T. T. Liu A., Mellers B. A. and Stark L. W. (1996). The effects of pictorial realism, delay of visual feedback and observer interactivity on the subjective sense of presence. Presence: teleoperators and virtual environments 5(3): 263-273.

## AUTHOR BIOGRAPHIES

**Nahella Ashraf** has a BSc degree in Information Technology Systems and an MSc in Human Factors in Information Technology. Her work in the Centre for Communication Interface Research (CCIR) has included trials on automated telephone services, electronic programme guides and interactive television. Her research interests include ergonomics research and development, usability testing and human-computer interface evaluation.



**Dr James Anderson** completed his doctoral thesis in 1998, documenting the development of fast algorithms for modeling virtual clothing. During his time at the CCIR, his work has been concerned with usability issues in telepresence shopping services, virtual mannequin technology, collaborative shared-space shopping services, shared-space technology for virtual conferencing, web-based financial services with multimodal information channels, and the usability of synthetic human-like agents.

### *Contact information:*

Dr James Anderson  
Centre for Communication Interface Research  
Tel: +44 131 650 8230 Fax: +44 131 650 2784  
Web: <http://www.ccir.ed.ac.uk>  
Email: <mailto:jad@ccir.ed.ac.uk>

**Craig Douthier** graduated from Napier University in 1996 with a BSc (Hons) degree in Computer Science. His work at the CCIR included projects examining the design and usage of 3D interfaces for online shopping based on VRML and Java technologies, particularly with a view to keeping users entertained during their shopping experiences.





**Professor Mervyn Jack** is Professor of Electronic Systems at the University of Edinburgh. He leads a multi-disciplinary team of twenty researchers investigating usability engineering of eCommerce services. His main research interests are dialogue engineering and virtual reality systems design for advanced eCommerce and consumer applications.

*Contact information:*

Professor Mervyn Jack, Director, CCIR  
The University of Edinburgh 80 South Bridge  
Edinburgh EH1 1HN  
SCOTLAND  
Phone: +44 131 650 2783. Fax +44 131 650 2784  
Email link: <mailto:maj@ccir.ed.ac.uk>



**Dr Iain McKay** joined the Centre for Communication Interface Research as a Research Associate in 1994 and worked on various projects for commercial sponsors in the areas of electronic commerce, virtual reality (notably techniques for 3D shopping environments) and interactive systems and services to mobile devices. He completed his PhD thesis (“Design for Usability of Interactive Multimedia Services to the Home”) in 1999. He is the author of six research papers in Human Factors, one book (pending publication) and is now a Senior Research Fellow and Deputy Director of CCIR.