Your Place or Mine? Learning from Long-Term Use of Audio-Video Communication

Paul Dourish[*], Annette Adler[†], Victoria Bellotti[‡], and Austin Henderson[‡]

[*]Rank Xerox Research Centre (EuroPARC), 61 Regent Street, Cambridge, CB2 1AB UK
[†]Systems Architecture Group, Xerox Corp., 3333 Coyote Hill Rd, Palo Alto, CA 94304, USA
[‡]Advanced Technology Group, Apple Computer, 1 Infinite Loop, Cupertino, CA 95014, USA

dourish@europarc.xerox.com

Abstract. Workstations and personal computers are increasingly being delivered with the ability to handle multimedia data; more and more of us are linked by high-speed digital networks. With multimedia communication environments becoming more commonplace, what have we learned from earlier experiences with prototype media environments? This paper reports on some of our experiences as developers, researchers and users of flexible, networked, multimedia computer environments, or "media spaces". It focusses on the lessons we can learn from extended, long-term use of media spaces, with connections that last not hours or days, but months or years. We take as our starting point a set of assumptions which differ from traditional analytical perspectives. In particular, we begin from the position that that a real-world baseline is not always an appropriate point of comparison for new media technologies; that a set of complex and intricate communicative behaviours arise over time; and that media spaces connect not only individuals, but the wider social groups of which they form part. We outline a framework based on four perspectives—*individual, interactional, communal* and *societal*—from which to view the behaviour of individuals and groups linked by multimedia environments. On the basis of our long-term findings, we argue for a view of media spaces which, first, focuses on a wider interpretation of media space interaction than the traditional view of persont-to-person connections, and, second, emphasises emergent communicative practices, rather than looking for the transfer of face-to-face behaviours.

1 Introduction

Over the last ten years or so, our research groups and others have been exploring the use of "media spaces"—switched computational and multimedia communication environments supporting cooperative work. A number of these have been built and studied, including PARC's "Media Space" [Stults, 1989; Bly et al, 1993], Bellcore's "Cruiser" [Root, 1988; Cool et al, 1992], University of Toronto's "CAVECAT" [Mantei et al, 1991], and EuroPARC's "RAVE" [Buxton and Moran, 1990; Gaver et al, 1992]. While the research groups differed in their particular concerns, they shared a common focus on *communication* rather than on communication *technologies*. So, these prototype environments were constructed using analogue baseband technology and off-the-shelf components for in-house experimentation. More recently, developments in data network design have begun to bring multimedia communication facilities to the users of Internet-connected workstations through the introduction of multicast media transmission over the IP protocol [Casner and Deering, 1992; Eriksson, 1994], and over high speed networking fabrics such as ATM. Increasingly, workstations and personal computers are being delivered with the ability to manipulate, send and receive multimedia data, and video communication is one of the revolutions which the "information superhighway" is hyped to bring us. Whether or not we are all about to enjoy the delights of interactive media in our living rooms, these trends suggest that multimedia communication environments will become considerably cheaper and more widespread over the coming years.

As these technologies have been developed and deployed, researchers have conducted a number of studies, formal and informal, into the nature of multimedia communication and of interaction in technologically mediated environments. Studies have focussed on aspects such as the impact communication facilities have on group working [Mantei et al, 1991; Tang and Isaacs, 1993]; connection management architectures [Dourish, 1991]; and privacy implications [Fish et al, 1993; Bellotti and Sellen, 1993]. More specifically, we have also seen a number of investigations of the interaction of technological mediation and the mechanisms and practices by which conversation and interaction are typically managed [Gaver, 1992; Heath and Luff, 1992a; Sellen, 1992; O'Conaill et al, 1993].

In this paper, we want to take a different approach to understanding the nature of communication in media space environments. In particular, we want to take as our starting point three positions which differ from the normal assumptions behind these investigations:

1. Face-to-face[1] communicative behaviour in the real world is not always an appropriate baseline for the evaluation of mediated communication. Moving away from this perspective allows us to explore a number of important, intrinsic properties of video as a communicative medium in its own right;
2. A set of communicative practices, tailored to the nature of the medium, arise over time as familiarity with the medium increases. These are related to the specific people and work practices involved—a case of coevolution—and so these practices must be studied in real, long-term use;
3. The use, influence and importance of such technology extend beyond the individuals who are directly engaged with it. Not only does it extend beyond individuals, but also beyond their immediate context and environments, encompassing the wider social groupings in which they are located. This wider group can and should affect the evolution of use and action within the media space.

These three points are central to our view of the use of media space technologies. They encompass wider concerns than are traditionally the focus of usage investigations. By studying how people adapt to various aspects of these environments and make use of them, we can illustrate aspects of their value which may go unnoticed in comparisons with "real-world" interaction, and reveal lessons for the design and evolution of such systems based on a new model of their utility.

2 Background and Related Work

Over the past few years, as multimedia communication environments have become part of working life in various research groups, a number of studies have been published on different aspects of interpersonal communication mediated by the technology. However, for a variety of reasons—not least logistic—none of these studies have focussed on the kind of long-term, everyday use which we will discuss here.

Heath and Luff [1991; 1992a], employing an analytical perspective derived from ethnomethodology and conversation analysis, studied the way in which "video technology... appears to transform the nature of visual conduct". Their conclusions are primarily based on the detailed analysis of video data collected from a number of connections. Some of their data was drawn from a traditional long-distance digital video conferencing system, but the bulk was collected from pairs of remotely-connected colleagues communicating through an in-house analogue media space for periods of a few weeks. Studying various interactions in this data, they point to areas in which video mediation interferes with the techniques by which talk is regulated and managed in face-to-face conversation. In particular, they use examples which illustrate the relative impotence of gesture and the ineffectiveness of gaze in performing their usual functions, helping manage the flow of talk; they have lost their "interactional significance". Less formally, Mantei et al [1991] describe similar issues and problems in the use of their CAVECAT media space.

From a different perspective, Sellen [1992] and O'Conaill et al [1993] also investigated the impact of electronic mediation on patterns of verbal interaction. Looking at a range of cases of interaction in different technological environments and for different tasks, these studies investigate the impact of communication technologies on verbal conduct. In both laboratory studies and naturally-occuring situations, they make direct comparisons between communication face-to-face and mediated by technology, and use face-to-face interaction as an evaluative baseline.

1. The term "face-to-face" can be misleading. We use it here to mean conversation unmediated by electronic technologies. We will use "person-to-person" for the idea that a video connection connects just two people, and "head-and-shoulders" for the restricted, face-centric view of many systems.

Gaver's [1992] study has a different flavour. He looks primarily at the properties of the medium and their implications for perception and action—the "affordances" of the technology. His analysis explores properties inherent to the design of current media spaces, such as the two-dimensionality of video images and the discontinuity of "movement" in media space environments. From these, he derives a set of implications for communication mediated by video technology, organised around the way in which specific properties of the medium "afford" action to individuals (or not). This perspective reflects an orientation towards the redesign of communicative technologies. Indeed, the analysis has been used as the basis of further design which attempts to address problems raised by the affordance perspective, particularly those concerning the exploration of remote spaces [Gaver et al, 1993; Gaver et al, 1995].

While they vary considerably in scope and perspective, these studies focus primarily (but not exclusively) on investigating conversational, person-to-person, head-and-shoulders interaction. In addition, they concentrate on properties of video interaction in comparison to everyday interaction, and so design suggestions tend to be organised around reducing some observed disparity between them. While this is clearly a valuable exercise, especially for future casual users of more widespread multimedia environments, we have a different focus which contrasts with this tradition. Rather than look at the comparisons *between* video and the "real world", we want to look at video *as a part of* the real world, and how people organise everyday, "real world" activities around it.

2.1 Office-shares

Our characterisation and observations are based on our own experiences as long-term users of media spaces. One aspect of our media space use has been particularly significant. Each of the authors has been part of a long-term semi-permanent pair-wise audio and video link known as an "office-share" connection. An office-share connection is one of a variety of connection styles in our environments [Dourish, 1991]. It is a two-way audio and video connection which is technologically similar to a "video phone" connection, but is intended for less focussed and more long-term connectivity. Essentially, an office-share connection uses the technology to connect offices and create a virtual shared space. What makes our office-share connections particularly interesting is that we have had the opportunity to use this technology almost continually over unusually long periods. Bellotti and Dourish had such a connection between their offices, day-in and day-out, for a total of over three years, using EuroPARC's RAVE media space environment; Adler and Henderson used similar facilities at PARC for a total of almost two years.[2]

The technological environments which support our connections are similar; they share a common intellectual heritage. Both are based on analogue communication facilities, and the office "nodes" in each use stand-alone video monitors (separate from our workstation screens), with separate cameras, and omnidirectional microphones with on/off switches or foot-pedals. One major difference between the environments has been in the underlying connection infrastructure. Adler and Henderson's connection was, for some time, a "direct" connection with no switching facilities, although latterly it was provided as part of a switched media space called "Kasmer". Bellotti and Dourish's connection was always part of a switched media space environment; they had the ability to "glance" momentarily elsewhere, switch the connection on and off at will, and be connected to others simultaneously through systems such as Portholes [Dourish and Bly, 1992].

Through informal discussion, we began to realise that there was a considerable overlap in our experiences in these two long-term connections. Our parallel experiences in these separate office-shares have convinced us of the value of investigating and documenting the long-term use of multimedia communications.[3] We have been particularly interested in the ways that our technological "holes in space" have

2. These connections were broken with the departures of participants from our environments; however, we will occasionally lapse into a wistful use of the present tense in our descriptions.
3. In fact, we had already, independently, been keeping our own "video diaries" of interesting facets of multimedia communication.

engendered changes and adaptations in our communicative behaviour. Interestingly, we have also found considerable similarities in the adaptations to the connections made by *colleagues and neighbours* in our working environments, working nearby our connections. We believe that an understanding of these wider-ranging and longer-term aspects of multimedia connectivity is crucial in informing the future use, design and deployment of similar systems.

When looking at patterns of media space interaction, it is important to bear in mind that the role of video communication is not simply to simulate copresence. Instead, media spaces extend the range of poassible encounters between participants, both in form and opportunity. As Mitchell [1995] has observed, in the end there is no irony to the fact that the first message conveyed by telephone was, "Mr Watson, come here; I want to see you." Communicative modalities exist side by side, rather than in competition with each other. This is why our media spaces comprise not simply "desktop videoconferencing" facilities, but a range of connection styles, as well as related applications such as Portholes. (Bellotti and Dourish [forthcoming] discuss this issue in more detail). As such, then, our "office-sharing" activities were somewhat unusual, since the facilities for extended two-way communication were generally not those used most often within our media spaces (after many years' experience, only a few other examples exist, and as discussed below, they were generally not equivalent). However, through this paper, we will argue two points which, we feel, *reinforce* rather than undermine the argument that media spaces should be seen as augmenting, not replacing, other forms of encounter. First, we will emphasise the importance of *continual connection,* which, within organisational and logistical constraints, would be otherwise impossible and therefore arises alongside other forms of connection; and second, we will illustrate that the role of our connections was in fact to create a new social space *within which* encounters could take place, thus, again, expanding the range of potential encounters.

2.2 Presenting Experiences and Reflections

We will illustrate our wider interpretation of media space communication with anecdotes not only from our own experience, but also from questionnaires and discussions with other long-term communication partners in other groups. However, data of this sort is very sparse. Media space research has not been established for long enough, or widely enough, for there to be a large set of experiences to draw on. Where we have been able to find other cases of long-term connections to compare with our own, in our own or other laboratories, we have typically found them lacking in some respect—most commonly that they were video-only (rather than video plus audio), of a duration of months rather than years, established only periodically, or established between common areas rather than private offices[4]. So it's important to be clear on the role of the experiences we will report. They could scarcely be regarded as an ethnography of long-term video use, or the results of participant-observation, and they are not presented as such. Rather, where there is some commonality in our mutual experience, we present anecdotes which help to illustrate the perspectives we wish to highlight. They embody particular views of media space interaction, and help to explain the perspectives we adopt.

The structure of this paper is as follows. Section 3 will introduce and outline a framework for analysing interactions in and through office-share connections. Section 4 will step through each of the four perspectives in our framework, discussing it in more detail and illustrating it with examples from our own experiences with our office-share connections. Drawing on these, we go on in the subsequent sections to discuss some general issues cutting across the analytical perspectives, and some consequences for the future design, deployment and analysis of media spaces.

4. We contacted four prominent industrial and academic laboratories with large-scale, long-term, ongoing media space research programs, and subsequently gathered information on six other long-term connections, through questionaires and interviews; however, none of these had all the characterisics we were looking for.

3 Perspectives on Mediated Interaction

The introductory material has made it clear that we wish to look at the long-term use and impact of media space technologies more widely than simply focussing on the mechanics of pair-wise communication. Our experience has been that a number of other surrounding factors form an important part of the use of these systems. For the purposes of this paper, we will draw on and expand a framework proposed by Adler and Henderson [1994] which characterises different perspectives ("sightings" in Adler and Henderson's terminology) on communication via multimedia technology. The perspectives are complementary, each highlighting particular aspects of use. Together, they form a framework within which to discuss our experiences.

The first is the *individual* perspective. From this perspective, we focus on the interaction between an individual and the technology which provides their interface to the communication environment. This brings up many of the traditional concerns of HCI and ergonomics. Next is the *interactional* perspective on the technology, which focuses on interactions between two individuals at each end of a media space connection, and their communication *through* the technology. From this perspective, we are particularly concerned with the ways in which aspects of their communication are modified or transformed by the technology, and how the participants manage their communication in a mediated environment. Investigations grounded in disciplines such as conversation analysis have generally focussed primarily on this second perspective; although since they often look at cases of short term use (generally less than a few months), they tend not to place an emphasis on the adaptive transformations to the extent that we will here.

These first two perspectives reflect fairly traditional concerns in analyses of video-mediated interaction, although we emphasise a (less traditional) sharp distinction between individual effects (concerning relations *with* technology) and interactional ones (concerning relations *through* techology). The third and fourth perspectives, however, reflect our wider concerns with the implications of long-term office-shares and similar facilities not just for the direct participants but also for others around them.

So, moving on from the interactional view, we will consider a *communal* perspective on mediated interactions. This view arises from our observations that the importance and influence of a semi-permanent connection reaches *beyond* those directly connected, drawing in others from physically or socially proximate groups. Issues of membership, access and "reach" become important when we look at office-share from this perspective. Finally, we will look at a more general set of issues which we refer to as *societal*. From this perspective, we focus on the implications such connections hold for the relationship between the individuals and the larger social groups to which they belong.

In what follows, we use this framework—individual, interactional, communal and societal—to draw out different aspects of mediated communication, based on our observations of long-term media space usage. We will use the framework to *organise* our anecdotal observations, but not to *characterise* them. The framework is not a taxonomy, nor is it complete. Instead, reflecting our use of the term "perspective", we use each component of the framework simply to give us a particular way of looking at issues in media space use, revealing particular facets of use. Typically, at any given moment, elements of each are present simultaneously, and with mutual influence. Design must take account of how aspects of media space interaction appear from each perspective. In the subsequent section, we will focus more on the way in which certain aspects of communication cut across the four perspectives, and discuss general issues with an eye towards implications for the design and deployment of multimedia communication systems.

Figure 1: A video node in an office. Note the size and physical impact of the technology, as well as its position relative to the working position of the room's occupant, at the workstation.

4 Experiences

4.1 Individual

From the individual perspective, we address the relations between an individual and the equipment which provides their interface to a media space. This equipment can take various forms, but in each of our cases we use video monitors separate from our workstation screens, and separate cameras. Sometimes, the audio and video facilities are combined; for Bellotti and Dourish, the microphone, speakers and mixers are separate units, whereas for Adler and Henderson, the microphone and camera are integrated, as are the speaker and video monitor.

There are a number of important issues to be raised even from this fairly restricted perspective. The first is that the equipment clearly takes up a significant amount of desk space. As a result, physical placement, in relation to the usual working position of the room's occupant, is significant and constrained. In an informal survey, we found that the two most common determinants cited for the arrangement of video equipment in an office were desk space and the reach of cables from the wall. So most people worked within these mundane constraints as best they could. It's unsurprising that only a few individuals had actually reorganised their offices so that the relative positions of desks, video node, workstation and working position were appropriately matched to their everyday activities; but it is also significant that some did.

Concerns about cable lengths and desk space seem trivial and inconsequential, but they turn out to have significant impact. An individual's ability to appropriately place equipment within her working environment is key to her ability to flexibly manage video interaction *as part of* her everyday activity. When a monitor isn't within line-of-sight, or a volume control is out of reach, then they are rendered useless; and if a camera cannot be placed so that it gives access to the space in which someone works, rather than just that person's head and shoulders, then there are important consequences for the individuals' interaction[5]. Factors like these are important resources for communication.

Example 1: Directions

The equipment tends to be bulky; it takes up significant desk space, and often must be specifically oriented towards as part of interaction. As a result, it has a significant presence in an office. In the course of a long-term office-share connection, this "presence" can easily become transferred to the remote participant. Essentially, the equipment becomes an interactional surrogate for the video partner.

5. For instance, in one of Heath and Luff's studies, the two individuals connected by video both had their cameras and monitors placed *behind* them. Neither could see the other from their normal working position at their desks. Heath and Luff discuss the considerable difficulty these participants had in attracting each other's attention.

In one of our connections, one participant asked the other for directions to a nearby store. The remote participant began to answer, pointing and using hand gestures to explain the directions. The gestures were misleading, however, because the questioner's orientation in her office was different from her *apparent* orientation in the remote office. Her apparent orientation was determined by the arrangement of the video equipment at the other end; the gesturer gestured *for the equipment* rather than *for the remote participant*. Since there was no consistent mutual alignment between the partners, pointing failed as directions were mistaken. The use of pointing and hand gestures by the remote participant was based on a perceived "mutual" orientation; an orientation based around the view of the video equipment as a surrogate person. Interactionally, this is Gaver's "anisotropism"; *individually*, it highlights technology-as-surrogate.

Once this distinction (between participant and surrogate) is understood, then it is possible to work with it. Indeed, an understanding of the different orientations can become natural and intrinsic to interaction. We learned to point "through" the connection so that our remote image would "point" correctly, and this felt natural since the separation between source and image became familar. In one notable case, one of us pointed easily (or rather, arranged for his remote image to point) towards a mislaid document lying on his colleague's desk, to one side of her video monitor. Understandings of this sort are based on the stable and familiar arrangement of equipment which allow long-term communicative partners to build a picture of the context in which their image is presented, and so tailor their actions for appropriate delivery[6]

Example 2: Noises Off

The "directions" example occurred early in the life of the office-share connection; learning to separate source from image is an example of the sort of adjustment which takes place over time in using this sort of technology. A second example at this individual level concerns the field of view of a camera—that space which is captured in the video frame. Cameras are typically arranged so that the normal working position of the room's occupant is within the field of view. However, it is only over time that users develop an understanding of which parts of the room are in view and which aren't. The first indications that this was problematic at all arose through errors on the part of third parties, less used to the technology. It is not uncommon for a colleague to arrive at an office, and then greet and talk to the remote partner while still standing in the doorway—from where they can see the monitor, and hence the remote partner, but are out of the camera's field of view. For the remote participant, the effect is of a disembodied voice (although, typically, they can recognise and interpret the reorientation of the office's occupant towards his or her visitor). This happens often enough that it quickly becomes recognisable and unremarkable. Experienced users may typically *not* comment on it at the time but simply respond to the other person *as if they could be seen*, especially when dealing with people unfamiliar with the environment. We are familiar with others' unfamiliarity with office-share connections.

The use of a "confidence monitor", showing the outgoing video signal, can help solve the general framing problem, and this is one route towards establishing an understanding of field of view, and hence visual presentation. Bellotti and Sellen [1993] discuss a number of ways that this general problem of "dissociation" can be addressed. A variety of solutions can be used to help smooth the path towards familiarity with aspects of multimedia technology.

4.2 Interactional

The second set of issues concerns the nature of interaction when it's mediated by technology of this sort, and the sorts of practices which emerge for regulating such interactions.

The nature of interpersonal communication in video-mediated environments has been the subject of

6. Clearly, the *specific* understandings of local versus remote orientations are not easily transferrable across configurations of offices and technology; however, the distinction between image and partner *is*.

previous studies described above, such as those of Heath and Luff [1992a] or Sellen [1992]. Again, our observations differ from those studies in that we concentrate on longer-term usage of media space environments—over periods of years—and subsequent adaptations to the nature of the medium.

Another major difference between the connections we have been exploring and those of subjects in Heath and Luff's experiments is the use of continuous open audio. Much research points to the relative importance of audio (usually focussed on audio quality) in audio/video environments (e.g. Tang and Isaacs, 1993). The subjects in Heath and Luff's naturalistics studies used continuous video connectivity, but either used telephones for audio or foot-pedal switches which left their audio disconnected unless explicitly enabled. In contrast, our office-share connections have been based on continual audio accessibility, and our experiences suggest that this enables many important aspects of "office-sharing".

Example 3: The Online Encyclopaedia

In particular, open audio enables very lightweight initiation of conversation, or short bursts of interaction. This feature of the medium supports particular forms of interaction.

First, it supports quick exchanges with very lightweight initiation. For example, we find that we frequently use our connections for simple quick queries—asking about spelling, people's names, research pointers, publication references and so forth. These are extremely short interactions, conducted often without looking up to establish visual contact. The audio channel is the primary means by which we can assess the other's availability for interaction. The remote partner rarely finds these questions intrusive, since they can be answered (or rebutted) with as little overhead as they are made. In a system in which the initiation of conversation is more complex, requiring that both partners explicitly "enable audio" in some way, we feel these queries would be much less likely to take place. The act of turning audio on and off would be much more intrusive than the audio itself[7].

Second, continual audio access lends video partners *peripheral awareness* of each other's activities which can be invaluable. In their studies of collaboration in London Underground control rooms, Heath and Luff [1992b] observe how peripheral monitoring and the specific public rendering of "private" activities, both of which are based largely on continual auditory access, enable fine-grained interaction and coordination of activities. Other groups have observed that these informal, awareness-based facilities are some of the most useful aspects of media space environments [Olson and Bly, 1991; Fish et al, 1992; Bly et al, 1993]; and audio provides a wide range of cues which support this process.

Example 4: Gaze Awareness

We also found a number of specialised patterns of behaviour develop around the video component of our connections. One of the most interesting was originally noticed in reference to Bellotti and Dourish's connection. With cameras sitting on top (or to one side) of our video monitors, our video connections do not support eye contact between remote partners. Early on, after two or three months, we noticed that we had developed the habit, as listeners, of *looking directly into the camera* (rather than the monitor) during periods of extended talk. The result is that the speaker sees the listener apparently looking directly at them, while (in typical configurations), the listener can still watch the speaker from the corner of his/her eye.

This adaptation to the technology—especially since it arose as an automatic behaviour—is interesting and suggestive in itself. However, after greater familiarity developed, we noticed that we had silently abandoned it. It was simply no longer necessary for effective communication. This appears to result from increasing *gaze awareness*, as discussed by Ishii et al [1992] in relation to video-based shared workspaces and by Tang and Isaacs [1993] in relation to workstation-based multimedia communications. Gaze awareness refers to a recognition by one partner in a video connection of the remote partner's focus of

7. In fact, the permanent audio stream itself is not particularly intrusive as the technology tends to "flatten" the audio signal, making it relatively easy to ignore; we will discuss this point later.

Figure 2: Gaze awareness. On the left, the listener looks into the camera (giving simulated eye contact). On the right, he is watching the monitor, attending to the speaker.

attention based on apparent gaze direction. When video partners correlate gaze direction with a specific focus of attention, then it becomes available as a resource for managing their conversation. So the "looking into the camera" behaviour arises initially as a way to give the speaker feedback that the listener is paying attention. However, as the partners develop a greater awareness of each other's gaze patterns, the speaker learns to recognise when the listener's gaze is directed at the video monitor, and hence that the listener is indeed paying attention. The understanding that "I have eye contact" is replaced with the understanding that "she's looking at me"; and it's this more general understanding which is important.

4.3 Communal

Most analyses of multimedia-based communication have concentrated on communication from the individual or interactional perspectives. However, in our experiences with long-term, regular connections, we have found that these perspectives tell at most only half the story. A number of critical issues, especially in the longer term, emerge when we take a wider view of multimedia connections within the organisational and social context of workplaces. This takes place, however, in the context of a connection between two individuals (in contrast to connections explicitly maintained between groups). Our third set of issues, then, is concerned with the interactions between an existing office-share and other people in the local environment or working group—the communal view.

Example 5/6: Communication Through the Shared Office

While an office-share connection provides a link between two individuals, those individuals usually work in environments where they interact with other people who are either physically nearby or organisationally closely related. We found that these factors affected our long-term office-share connections. One interesting starting point is the use of the connection by other people in our environments. The connection becomes a "hole in space", which can be used to communicate with a person at the remote end. At both PARC and EuroPARC, we have found it quite common for colleagues to step into our offices to talk, not with us (the local occupant), but with the person at the *other* end of the office-share connection—or even with other people *near* to the other end. Indeed, on one occasion, Bellotti and Dourish returned to their offices from a meeting to find a person in each office, talking to each other over the link. An even more complex example arose between Adler and Henderson. One was on the phone and the other arrived in their own office, at the same time as two other colleagues were talking over the link—several different forms of interaction concurrently in the same "space". Not only do we, as media space users, come to think of ourselves as "sharing an office", but our colleagues do, too. They orient towards either office as part of a shared space.

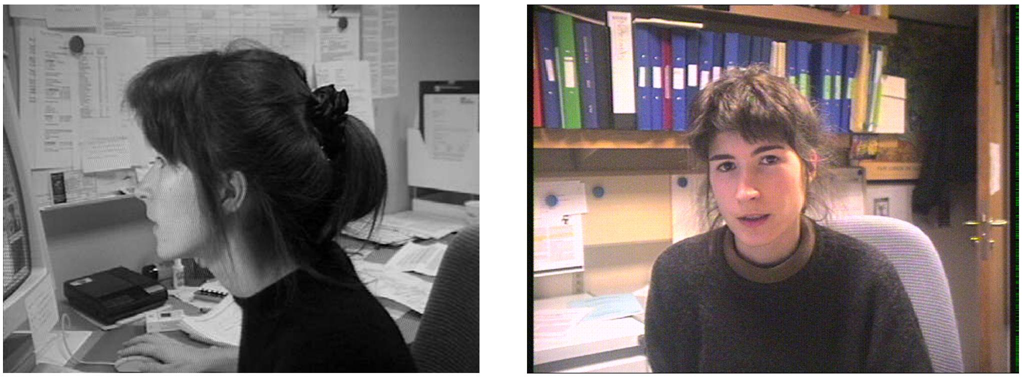

Figure 3: Bellotti physically reorganised her office, rotating it 180 degrees, to give Dourish a view out of her door and into the space beyond

Example 7/8: Presence and Telepresence

Creating a “virtually shared” office can also cause confusion for people not accustomed to the connection being in place, since the effects of “telepresence” also extend beyond the connection’s ostensible participants. A short time after one connection was established, Dourish was asked by a colleague who it was who had been using his office that morning. Dourish was confused, knowing that the office had been empty. It transpired that the colleague had heard the sounds of typing and activity coming from Dourish’s office, although in fact they had originated in Bellotti’s and been transmitted across the connection; hearing these sounds, the colleague had presumed that there was someone physically present in Dourish’s office. Further, those unfamiliar with the presence of a connection can also find that it upsets their intuitions about the organisation of spaces and activities. On a different occasion, a visitor to EuroPARC was on his way to talk to Bellotti, but chose to wait until later because he heard the sounds of conversation coming from her office. Knowing that hers was a one-person office, he assumed that the talk indicated a meeting in progress, and felt he should not interrupt. In fact, of course, the conversation was with her (virtual) office-mate. Had he thought of the office as being shared, as did our colleagues, it seems much less likely that he would have interpreted the sounds of conversation in that way. In the media space, the functional space is no longer isomorphic with the physical space, hence this confusion. Unfamiliarity with the connection lead to his misreading the significance of the talk.

Example 9/10: The Virtual Neighbourhood

The use of open audio leads to more subtle extensions of a link between individuals to the wider group. Audio signals can travel in either direction, to or from the wider environs of an office.

In one direction, this results in the effect of “virtual office neighbours”—sounds reaching one end of a connection which originate, not in the other office, but in a nearby office. This has become an issue at some time for each of our connections. In one, it arose when a “virtual office neighbour” was performing video analysis, which involved repeatedly replaying short video fragments. The sounds on their own, divorced from the activity of analysis, became very strident and distracting across the media space. In the other connection, a similar example arises; in this case, however, the “neighbour” was actually PARC’s fitness centre, with the sounds of the daily aerobics classes being carried to the remote participant. The lack of context turned these everyday background sounds into noise pollution across the media space; but they also convey a sense of activity from the other end.

This effect also operates on the other direction—reaching out into the remote “neighbourhood”. We rapidly found that, as well as wanting our remote partners to position their cameras to focus on their customary working position in their offices, we would want to “see out of the door” into the corridors and public spaces beyond; to get a sense of the activity there, be able to see passers-by and even engage them in conversation. We share not only virtual offices, but virtual neighbourhoods, which impinge on us, and in which we interact.

Example 11: Projecting Audio

The converse effect is where the connection doesn't pick up sounds from the surrounding space, but projects the sounds of mediated conversation into an area beyond the remote office. Gossiping about someone who might be just outside the remote office can be problematic, and confusion over ringing phones is also common.

One problem which gives rise to these effects is that a (near-end) participant is very limited in his or her ability to gauge or manipulate the audio levels at the far end. Volume variations in speech are less effective, since they primarily regulate the amount of signal which is picked up by the microphone; remote amplification, which is not under the sender's control, is the most significant factor in the audio level, and this level is something of which the speaker is generally unaware.

However, as participants become familiar with these sorts of effects of the mediating system, they can become a resource. Audio signals which project into the remote environment may be troublesome during private conversations, but have been effectively used to attract the attention of remotely-observed passers-by. In cases like this, the curious volume distortion effects which mediation introduces can actually become a resource; speaking softly, close up to the microphone, produces an effect which at once sounds intimate (even conspiratorial) but is actually quite loud and carries well.

We can see that audio begins to "open up" the space; again, the reach of a connection extends beyond the individuals who're directly involved. Other users of long-term connections have reported similar behaviours, which suggests that participants in such connections do indeed move from an initial understanding of linked individuals to one of linked spaces. Harrison [1992] has similarly emphasised the importance of this notion of "place" as an important factor in contextualising behaviour and in designing understandable communicative systems. The notion of place, established over time, embodies communally-understood patterns of acceptable and appropriate behaviour. By affording us the opportunity to transform "space", media spaces also affect our sense of place, as illustrated by these emerging patterns of behaviour and adaptation amongst our physical and organisation neighbours.

4.4 Societal

The communal perspective dealt with specific elements of behaviour arising around our connections in local groups. While that takes a wider perspective than the more common person-to-person view of video interaction, it still deals with activities in the immediate social or geographical group. The societal view encompasses wider issues still, looking more broadly at the relationship between media spaces, specific connections and the wider social groups to which we belong, along with the norms and expectations they embody. These larger groupings are less adaptive in a short time frame; they have different rules for sense-making.

Issues of privacy and reciprocity are particularly important from this perspective. While particular aspects of these have been discussed elsewhere (in particular, [Dourish, 1993] and [Bellotti and Sellen, 1993]), looking at them from the societal perspective emphasises the way in which the communicative practices of an individual or a particular pairing relate to the norms which emerge within a group as a whole, or vice versa.

Example 12: Colleagues and Visitors

One obvious aspect of the societal perspective is the way in which the presence of the connections can reveal or highlight the delineations between various groups. An interesting pattern of relationships concerns the direct users of a connection, their immediate colleagues and co-workers, other lab colleagues and visitors, as revealed by their reactions to (and views of) the connections.

For instance, Bellotti and Dourish's connection at EuroPARC was within a small lab of around 30 people, while Adler and Henderson's at PARC was within a large research centre with over 300 employees. So other EuroPARC personnel—that is, the colleagues of the participants—were, largely, familiar with the technology and the presence of this particular connection. This affected the way in which people reacted to the presence of the connection; since it was a known and recognised feature of laboratory life, the level of immediate "social danger" which it represented was reduced. At PARC, as a much larger and hence more diverse and anonymous facility, this is much less true.

This is not simply a comment on familiarity and expectations, but becomes more significant when we look at how people's reactions and understandings serve to act as determinants of social grouping. There is a much more significant and apparent difference between the behaviour and expectations of these two groups—colleagues and visitors—at EuroPARC than at PARC. Since many more people at EuroPARC have, at some point or other, been involved in media space research or usage, the established "cultural norms" are different, incorporating the existence of the connection. At PARC, it is much more likely that laboratory members will share reactions with visitors[8]. Social groups of this sort form the context in which a connection is operated, and in which it is perceived by others.

Example 13: Public Affirmations

To go further, the way that a connection is used can be seen as an explicit demonstration of cultural norms, or of individuals' status within wider groups. A long-term connection is an obvious, highly visible, and hence public element of a person's working environment; and as such it might be regarded as an affirmation of an intellectual, workgroup or personal relationship. When cast against the background of "official" organisation structure, such an affirmation may be more or less surprising; and whether intentional or not, such affirmations may well be perceived by members of the wider social group.

For instance, Adler and Henderson were considerably separated spatially within a large building, and belonged to an organisational group "hosted" by PARC (but not part of it). Their organisational relationship was peripheral to the research centre's structure, and so relatively unfamiliar to most of the centre's staff. In this context, they could use their connection to explicitly affirm their organisational link. Interestingly, Bellotti and Dourish could use the same technological arrangement to do just the opposite. Officially working for different groups within a much smaller laboratory, they could use their link to affirm personal shared research interests (and, indeed, ongoing work) which went *against* the official organisational breakdown. These "readings" rely entirely on the presence of the connection within an evolving social and organisational contex[9].

So a video connection is not socially or politically inert. It can have strong effects on perceived groupings, on membership relations within a group, or even perceptions of the *existence* of particular groups. What's more, multimedia connections can be reorganised much more rapidly than physical working arrangements, and reorganisation is done by the people who are *within* those arrangements. This enables a new set of effects which cannot be replicated in the physical environment.

5 Encompassing Issues

The previous section drew on our own experiences with long-term multimedia connectivity to illustrate our four perspectives on video communication. Our focus, looking at patterns of adaptation not only in individual behaviour, but also more widely in related groups, leads us towards a new set of concerns, which differ from those of more traditional analyses. From the individual and interactional perspectives, we have looked at the emergence of patterns of communicative behaviour over time, organised *around*

---

8. Normally, this reaction can be characterised as, "Ew, don't you find it weird to have a camera pointing at you all day?". The answer is, "No".
9. Of course, the absence of a connection may be as telling as the presence, as we found when one connection was temporarily removed.

the technology and the medium itself. These illustrate an increasing focus on video as a communicative medium in its own right, and a move away from the problems which traditional analyses highlight through comparison with a baseline of face-to-face interaction. From the communal and societal perspectives, we have illustrated that a long-term connection will "draw in" not only those who are nominally "connected", but also others in physical or organisational proximity (as well as others involved in defining the directly-using group). We have found that many of the interesting features of video as a communicative technology located in real organisations and workplaces are revealed through the transformations in conduct of these groups, from local communities to wider societal groupings. A range of examples drawn from two very long-term, semi-permanent connections have illustrated some of these aspects of the technology from each perspective individually. A number of important facets of media space use occur from each perspective, in different guises. In this section, we want to talk about two issues as they arise across the four perspectives used in section 4. The issues which we will focus on are "ownership" and "evolution". We do not accord a particular status to these issues. Clearly, these are just strands of a complex fabric involving other issues such as privacy, responsibility, presence, control, feedback, familiarity, reciprocity and others too numerous to list. We concentrate on ownership and evolution in particular since they arise directly from our experiences, and since they both involve a view of people and technology taken altogether. They provide a useful starting-point from which to develop an understanding the wider interpretations of media space interaction, and to understand the implications of the multi-perspective view.

5.1 Ownership

The issues surrounding ownership manifest themselves from each of our our four perspectives. Here, we'll deal with two aspects: the ownership of technology, and the ownership of space.

At the level of technology, we find that interesting questions of ownership develop as a pattern of regular behaviour becomes established around a permanent or semi-permanent connection. In the absence of such a connection, it is unlikely that another member of my workgroup would be concerned about the state of my media space equipment—such as a broken microphone or flickering camera—since the technology is clearly not "owned" by them. However, my partner in a long-term video connection *would* be concerned about these things, and might become involved in setting them right, even in my absence. In this case, it seems that the ownership of the technology is less clear-cut. Ownership rights over the technology itself have, to an extent, been subsumed by the joint ownership of the connection. Both participants own this window between their offices, and hence both are concerned with the technology which realises it.

Relatedly, it is reasonable to ask what happens when no individuals see themselves as jointly owning a long-term connection. For example, this situation can arise when multimedia technology is used to link public spaces. A number of such experiments have been reported in the research literature. In the case of "the Portland link" discussed by Olson and Bly [1991], successful use of a link between public areas depended upon a common interest and working relationship which encouraged members of the distributed community to become involved with the technology and claim ownership. In contrast, Bellcore's VideoWindow [Fish et al, 1990] linked public areas, neither of which seems to have been an explicit geographical focus for workgroups in the way in which the PARC/Portland areas were[10]. There seems to have been less of a common work focus for the groups involved in the VideoWindow studies, and these connections seems to have engendered less enthusiasm, ownership and responsibility. The differences in the technologies themselves are important factors here. The PARC/Portland connection used small cameras mounted on tripods. Participants would pick them up, move them around, play with them and rearrange them to suit their purposes—it was *their* technology. In the Bellcore case, however, the

10. The common areas linked between PARC/Portland were central common spaces onto which private offices opened.

VideoWindow was realised using expensive, special-purpose videoconferencing technology, including a specially-built camera and large, carefully-configured back-projection displays. Opportunities for the users to become *involved* with the technology were minimal.

Harper and Carter [1994] report on a similar lack of success in another situation where common ownership and mutual interest was missing. In their case, neither the technology, nor the idea of a connection, were "owned" by the participants; and their lack of interest was overwhelming. This is a common phenomenon with technologies of all sorts. Jeanette Blomberg [1988] observes that while departmental photocopiers will be well maintained and kept stocked with consumables, the anonymous "hallway copier" is always someone else's responsibility and frequently sits jammed, empty or broken.

Moving on from ownership of technology, the ownership of space is even more complex in the presence of a multimedia communication environment, as issues of "proximity", for instance, become more malleable. Consider some examples. We have already pointed to the ways in which others in our working groups would use an ongoing connection as a "portal" to the other side. Conversely, local etiquette would also require that they acknowledge the remote individual when coming to talk to the office's "local" occupant. Both of these behaviours suggest that they orient towards either office as a single, shared office space; the shared property of both occupants. Certainly, the occupants can develop this feeling, and various instances have arisen when the remote "occupant" will feel sufficiently "at home" and present to be able to spontaneously make contributions to a local conversation. The nature of open audio means that auditory "space" is implicitly shared; however, the participants in a long-term connection may build this into a larger view in which they jointly own a single office (albeit one with a rather curious topology).

In a singular demonstration of this, one author (Bellotti) completely reorganised her office some months into the connection, in order to support a better mutual orientation for herself and her video partner. In effect, she turned her office around 180°, so that her remote partner would have a better view of her office, and most particularly, a view out of her office door into the public space beyond. The "shared office" was mutual space to such an extent that its organisation should reflect mutual benefit. This powerful example reflects a changing understanding of ownership which cuts across the levels of our framework. It shows a gradual evolution and reorientation from, first, a view of the equipment as a personal resource; through, second, the connection as an interpersonal resource; to, third, a wider view of the shared office as a group resource.

5.2 Evolution

A common theme to many of the observations presented above has been the evolution of a set of behaviours in media space communication.

This evolution takes place at a number of levels. We have talked about an evolution of orientation towards the technology of the media space; the evolution of communicative practices in support of two-way communication and conversation; and the evolution of understandings of the way in which media spaces disrupt the communal resource of "space". The example of eye contact and gaze awareness, which went through three distinct phases—from initial confusion, through simulated eye contact, to the use of gaze awareness—illustrates that such practices continue to evolve over the life of the connection.

The gradual development of understandings about the nature of mediated communication is central to our consideration of its value. This brings us to a position which differs markedly from the "real-world" basis of many other analyses. Heath and Luff [1992a] base their analysis of interactions in media spaces on a comparison with everyday face-to-face interaction. The "losses of communicative impact" which they describe reflect the failure, in the media space context, of behaviours which regulate and coordinate face-to-face conversation. However, we have attempted here to discuss video as a communicative medium in its own right, and so we've shown the emergence and use of *video-specific* mechanisms for interaction. In looking at much longer-term media space use here, we have been able to point towards the development of new behaviours tailored to the nature of the medium. The emergence of these behaviours can restore aspects of communicative impact.

Gaver's analysis focusses more directly on the properties of the medium itself [Gaver, 1992]. He explicitly "contrasts [properties of media spaces] with those of the everyday medium". This form of analysis can provide us with a range of insights into the role which the technology plays in mediating conversation. However, we must be careful about the conclusions we draw, and in particular whether they apply to the *medium* or to *communication*. So while it is quite accurate to observe that "video communication is anisotropic," it's much more problematic to continue, "...interfering with the design of communicative gesture". As we have already seen, communicative gesture in video environments can be quite effective, although, critically, the form and role of gesture is different from that used in face-to-face settings. Dykstra-Erickson et al [1995] observed similar phenomena emerging even over the relatively short period of ten weeks. They document the emergence of a "local visual language"—patterns of stylised gestures developed by media space users, with interactional significances which arose over time out of their use of the system.

Essentially, what we can point to here is a contextualisation of the range of possible action, based on familiarity with various aspects of the medium or even of a specific use of that medium. As familiarity increases, so does the range of activities which can be effectively performed with relative ease. We can only find these through long-term observation. Indeed, they may be quite counter-intuitive if we attempt to predict the effective usages of the medium from the perspective of everyday interaction.

This is not simply a claim that "users should learn to use the technology properly". Instead, on the one hand, it is a claim that analysis should be mindful of the duality of technology and practice[11]; and, on the other, that we are in danger of overlooking many of the important benefits of these sorts of technologies if we do not appreciate the evolving and *creative* uses of them which are *intrinsic* to the medium itself.

We have pointed out one other significant aspect of evolution. This is the evolution within wider communities and organisations of a set of understandings not just *about* video communication, but more broadly *including* video communiction. Office-share connections figure within people's understandings of the world, as demonstrated by their changes in behaviour and perception. One doesn't have to be a direct user of the technology in order to feel its effect and hence be involved in this process and evolution—just as one doesn't have to be watching television (or even own one) to be affected by the role of television in our culture. So people will adapt to the "fact of" video communication, rather than the use of it; their understandings evolve around the "fact of" its presence. The distinction between the "users" of media space technology, and the "others" on whom it has no impact, is clearly called into question. The impact of the technology is communal and societal. Traditional conceptions of "interface" and interaction limit our understandings of impact and significance (c.f. Grudin [1990], and Bowers and Rodden [1993]).

6 Designing Media Spaces

In attempting to emphasise the complex pattern of behaviours which we have found emerge, our focus is similar to that of Hollan and Stornetta [1992]—looking at video technology as a medium on its own terms, rather than by analogy with face-to-face interaction. In particular, we have emphasised "wider" issues in the use of such technology—issues which a focus simply on one-on-one interaction misses. The individual, interactional, communal and societal perspectives highlight aspects of use which are present simultaneously within the context of the sorts of connections we have described. Although clearly depending upon some physical properties of the technology, these aspects are not intrinsic to them, but rather are part of the emerging patterns of use in particular environments.

11. Being mindful of this duality implies that design should not attempt to eliminate it, or to encode the social in the technical.

The traditional design focus for multimedia systems concentrates on the individual and the interactional perspective. In this section, we will also apply the communal and social perspectives, and so get a more rounded and objective assessment of issues in the design and deployment of media space technologies.

Perhaps the crucial implication we can make at this point is also the most general. We have discussed the way that, over time, adaptations take place as partners in long-term communications in media space environments learn effective ways to use the system. We have found that the sorts of problems which new users typically encounter, especially with respect to their ability to manage and regulate conversation, are lessened with time as they learn to employ a new set of resources to regulate interaction.

This is not to diminish the significance of the results of the other studies we have discussed here. However, we have to see them in context—in particular, the context of short-term, casual use, rather than long-term, day-to-day experience. This contextualisation lets us separate *mechanism* from *accomplishment* in the longer term. If we confuse the mechanisms of face-to-face interaction—such as the roles of eye contact and gesture in managing turns at talk—for the accomplishments that they support, then we run two risks. The first is that we may simply devote time, money and effort to the development of innovative solutions to non-problems, supporting the mechanisms when, in fact, the accomplishments are in no danger. The second is that we never look beyond the (inevitably flawed) simulation of copresence, towards the wider harnessing of the properties of video communication.

When people have the technology around enough that it becomes familiar and part of the fabric of daily life, many of the initial breakdowns disappear. The issues are *respecified*, and so our understanding of the requirements on design is deepened. Our concern, then, is with the implications of such patterns of behaviour—patterns designed around the technology and emerging from its use—for the future design of multimedia environments.

6.1 Linking Spaces, Not Just People

One of the points we have repeatedly emphasised in the previous sections is that, despite the person-to-person view suggested by the traditional "video phone" analogy, media space environments link spaces, not just people. What's more, in most working environments, those are populated spaces, and so others in the immediate working environment are also drawn into aspects of the connection. Our concentration, then, is on creating an understanding of space as it is used and regarded by a larger group than simply those directly involved in the connection. This augments understandings of the importance of space from other studies, such as the video-as-data emphasis of Nardi et al [1993] or the context of activity explored by Gaver et al [1993], but it differs in its implications. Nardi et al point out that video can be usefully deployed as a shared focus for collaborative activity, and Gaver et al describe the way in which access to the *space* in which activities occur enhances interaction. For us, however, a crucial aspect of linking spaces is that "spaces" are foci for communal activity; by emphasising the linkage of spaces at each "end" of a connection, we can enhance the ability of individuals to participate in a wider range of remote activities, and to interact "in the space".

This group use of virtual space both supports and emerges around the relationships. It doesn't come from attempts to codify or recapture "real-world sociality" within the design. The experiences we have documented here of casual, informal, social interaction across our connections contrast interestingly with approaches which employ strong real-world metaphors to capture patterns of informal contact (such as the "corridor" model of the early Cruiser work [Root, 1988], or the "doors" metaphor of CAVECAT [Louie et al, 1991]). The patterns of group use and informal interaction we have described arise directly from individuals adapting the space to their needs, rather than from an attempt to create a space and a geography within which it should happen.

However, this view does have significant import for design. For instance, it means that design decisions which are evaluated purely against the criteria of face-to-face communication (such as reduced size images adequate for "head and shoulders" views) may no longer be appropriate. We have discussed the we found we wanted not just "head and shoulders" views, but views which showed the individual in the context of their office, and, if possible, out of the door and into the corridor too. The issue is one not only of technogical design, but appropriate deployment and the flexibility to be able to shape the space.

6.2 Audio

Continuous, "open" audio has been critical to a number of the office-share practices we have discussed. It supports the flexible managament of collaboration and interaction through peripheral monitoring of ongoing activity and the assessment of another's availability for interaction. More interestingly, we have also discussed the role of open audio in supporting very lighteight interactions, as well as in extending the "reach" of a connection beyond the immediate office-space. The provision of high-quality audio has often been a source of difficulty in experimental media spaces, but our perspectives reiterate its importance for the communal and societal perspectives as well as the more traditional individual and interaction ones. There are three points concerning audio which have particular relevance to our experiences, and which we'll discuss here.

The first is the flattening of the frequency spectrum in audio transmission. Since high-quality audio modulation and transmission is difficult to achieve in our environments, one of the effects is a "flattening" and homogenisation of the audio signal. It is this homogenisation which appears to be responsible for the difficulty in distinguishing between different sources of sound (such as between a colleague's voice and the music of the aerobics class). This impairs individuals' ability to "filter" the audio stream and listen selectively. While we have found that it becomes quite easy to distinguish "local" from "remote" sounds[12], making it easy to simply "ignore" the sounds of a conversation at the other end of our connection, it can be much more difficult to separate out mixed sounds at the remote end, making interaction more difficult.

One approach which is used in various environments is to employ microphone headsets to reduce the effect of background noises. Headsets are also intended to keep audio "private" and to reduce interference for others in proximity to the listener. However, we have pointed repeatedly to the way in which an important aspect of media space use is precisely the way it reaches and draws in others in the environment. When the presence of the connection itself is important, for instance in the cases of an explicit affirmation of working relationships, it is the audio connection which is more "visible" (if we can use the term) than the video to others in the environment. So, while headsets may remove problems from the interactional perspective, the effects can be less positive when viewed from the communal or societal perspectives. In addition, headsets distance the wearer from the local environment, separating the media space from the physical space; and it is precisely in the melding of the two that we have found much of the power of media space use.

Finally, there are design and interaction issues concerning directional audio information. We observed at the start that we use omnidirectional microphones. These make it easier to maintain a consistent audio environment, but at the cost of any kind of directional information in the audio signal. It is not possible to hear the subtle variations in sounds associated with a remote participant turning his or her head towards or away from the microphone (and the change in gaze direction which this movement implies). Given that we have observed that the audio channel is the one most commonly used to initiate interaction, this can be problematic. While video conveys little sense of the three-dimensionality of remote space [Gaver et al, 1993], unidirectional audio conveys even less.

12. Electronic sounds are much harder to distinguish this way. It's easy to tell when a human voice is local or being transmitted through the media space; but the sound of a ringing telephone can be quite confusing. It seems likely that this is because of their restricted frequency range, which means that they are less distorted by the media space.

6.3 Digital Transmission and Shared Media

The media environments with which we are most familiar are based on analogue technology. While our focus is on usage, rather than technology, it's clear that there are interactions between the two. Increasingly, digital transmission over data networks is becoming the most cost-effective way to provide multimedia connectivity in many environments, especially over long distances. Many digital transmission mechanisms differ crucially from the analogue systems we have been using in that they are based around a shared medium for which different users contend. Ethernet technology has this property, for instance, as do most long-distance lines based on time-division multiplexing. Analogue systems share switching resource, but not bandwidth.

The result is that, unlike the switched baseband networks we have discussed here, individual actions are no longer independent. Use of the network by one individual affects the level of service which can be offered to others. This can refer not only the question of whether a connection is in use or not, but also to the quality of connection provided. The different parameters of video stream, such as image size, resolution and frame rate, are emphasised to different degrees depending on the role it plays or the task being performed, as illustrated by Pagani and Mackay [1993]. Control over these parameters is a user issue. Flexibly managing connections over this type of shared resource is a major topic in multimedia design. Current approaches address issues of quality of service and resource reservation at the protocol level, removing it from user control and also distancing it from the valuable coding-specific information which could support the sort of control which Pagani and Mackay suggest.

The medium is not only shared amongst different users on a network; it is also shared between the different activities of a single user. Investigations of network management which take an "end-to-end" approach typically define "end" as the termination of the data stream at some host computer; however, it may become necessary to consider the management of multiple streams for the same user; a higher-level end-point. Some interesting work is beginning to emerge on relating high-level patterns of use to lower-level network management issues [Yamaashi et al, 1995].

6.4 No Sense of Place

We have followed Harrison [1992] in talking about the use of media environments in terms of "place" as well as "space". Harrison draws the distinction thus: "Space is the opportunity, place is the understood reality". Key here, then, are the opportunity to flexibly organise activities and structures, giving the means for place to emerge from space, and the mutually recognisable orientation towards spaces which carries with it a sense of appropriate behaviour and expections... a (shared) sense of place.

Place has been a particularly interesting notion in the cases we have discussed, since the spaces we have been dealing with are hybrids of the physical and the electronic; media connections which "punch holes" between physical spaces. This has let us create new spaces, which become distinctive places as sets of appropriate orientations arise within our communities. Our ability to *appropriate*, transform and reuse space is rooted in the flexible switching which media spaces afford[13].

A new range of multimedia environments is emerging in which issues of space and place are even more explicit. Jupiter [Curtis and Nichols, 1993] is a multimedia extension of MOO, a text-based "social virtual reality" which employs a strong spatial metaphor to manage the interaction of a large number of individuals with each other and with objects created in the virtual world. In Jupiter, the purely textual interaction between participants is augmented with the use of digital audio and video. Jupiter is designed as an interaction environment to support distributed workgroups, and hence shares a number of important properties with media spaces. Drawing on the MOO tradition, however, it is organised around a spatially-structured virtual world. MOOs frequently exhibit strong elements of "placeness". Particular plac-

13. As was stated, Adler and Henderson's connection was a fixed link for some time, but its existence was under their control. The fixed link described by Harper and Carter [1994] was not under the control of the participants; and indeed, it transpired that they had no motivation for the creation of a communal place.

es are public, private or somewhere in between; well-worn routes emerge between areas where people congregate, and elements of connectedness indicate aspects of the appropriate uses of particular "rooms", constructed by the community members and recognisable to them.

The experiences described here clearly show that a strong spatial metaphor of this sort—a rigid and explicit "geography"—is not a prerequisite for the emergence of a sense of place; it's the community's orientation towards the space that is critical. However, it also becomes clear from ongoing work with systems such as Jupiter, as well as its purely-textual cousins, that the use of the geographical metaphor engenders the emergence of a shared understanding of the varied appropriate uses of spaces. In integrating the two, then, and in employing spatial models in media spaces generally, it's important that we allow for the way that place-orientations *emerge out of* the flexible, exploratory and *creative* use of the space by its occupants.

7 Conclusions

> "[...] there is a very important social dimension to CMC [computer-mediated communication] which has been neglected by previous approaches in what we have called the 'social cues' tradition. These approaches assume that because many social cues are excluded in CMC, the social dimension is itself partialed out. We have argued that the tendency to neglect the social dimension has arisen from a theoretical blindness deriving from individualistic interdependence conceptions of social behaviour in which the interpersonal dimension of interaction is confused with the social *per se*. Once this equation is made it then becomes logical to assume that the social context impinges less on behaviour within CMC and there is then no reason to consider the distinctive social and normative contexts which we argue shape group behaviour in much experimental work in this field." [Spears and Lee, 1992]

The use of audio and video communication environments to support informal communication and collaboration between remote colleagues is on the increase, and with recent developments in multimedia transmission over data networks, it seems set to continue to rise. In this paper, we have discussed some wider aspects of interaction in environments of this sort.

Most investigations of video-mediated interaction have been based on relatively short-term experimental usage, and have focussed on what we have called the individual or interactional aspects of use. However, we have had the opportunity to look at patterns of behaviour that emerge from continual, everyday use over periods of years. Our experiences have led us to question a number of the basic assumptions which lie behind traditional analyses of media spaces and similar communicative environments.

The first traditional assumption is the *real-world baseline*. We have observed and related the way in which complex patterns of behaviour build up around the interactional details of the video medium; behaviours which support precisely the sort of interactional management which is said to break down in video-mediated interaction. For example, where Heath and Luff [1992a] point out the loss of eye contact in video environments and subsequent confusion in conversational regulation, we observe the use of gaze awareness and consequent recreation of effective conversational practice. When the medium changes, the mechanisms change too; but the communicative achievements remain. The real-world assumption is not only prevalent in media space analysis, but also in media space design, which looks to real-world interactional practice for a set of design guidelines for media environments. Clearly, there are important lessons to be learned from a comparison between mediated communication and face-to-face interaction. However, we argue that to use the real-world baseline to evaluate the efficacy and value of media spaces is to miss the point. The media space world *is* the real world; it is a place where real people, in real working relationships, engage in real interactions. Here we echo Spears and Lee, quoted above, who found a similar perspective towards the relationship between social behaviour and interpersonal action in studies of computer-mediated communication. So we take a more abstract view of the nature of mediated interaction, since we can look at it as a developed skill in its own right, rather than as a modified form of real-world, face-to-face interaction. Just as everyday conduct is organised around the everyday medium, we have found ways in which experienced users of video communication technology learn to employ resources which allow them to overcome some of the problems which have been identified in shorter-term studies as troublesome for new users. They create new patterns of behaviour organised around the nature of the medium itself.

The second traditional assumption is the *person-to-person view*. Media space investigations have, perhaps not surprisingly, focussed on the interaction between two individuals connected through video and audio technology. More recently, a number of researchers have pointed to the importance of providing more than pervasive "head and shoulders" view, and emphasised the value of video-as-data and access to the context of action. Our experiences, however, have led us to look first at the way in which media environments link spaces, rather than people; and secondly, to look outwards towards the groups, related not only physically, but also socially and organisationally, to the people on either end of the link. We found that transformations in their conduct were as remarkable as those in the conduct of the individuals notionally "linked". Many of the most important consequences of our own links have been those affecting colleagues in the spaces around us.

These observations lead to an interesting set of issues for those involved in the design, development and deployment of multimedia communication systems. They suggest not only that we must look more closely in order to decide which interactional problems are design problems, but also, crucially, that it is not merely a pair-wise communication that is supported, but a group within communal and social groups. As we move towards multimedia communication support on wider scales across national networks, these issues are likely to be of increasing importance, and a sensitivity towards them in design, deployment and analysis is critical if we are to realise the potential benefits of media environments beyond the flawed simulation of copresence.

Acknowledgments

We are deeply indebted to the many colleagues who have been involved in the development and study of media space environments over the years. In particular, we would like to thank our colleagues Bob Anderson, Sara Bly, Bill Buxton, Christian Heath, Paul Luff, Wendy Mackay, Allan MacLean and Tom Moran for their insights and contributions to our view of media space usage. We would especially like to thank Bill Gaver and Steve Harrison for their insightful commentary and conspiracy.

The work described here was conducted while Victoria Bellotti was employed at Rank Xerox EuroPARC, and Austin Henderson in the System Architecture Group of Xerox Corp.

References

Adler, Annette and Austin Henderson, 1994. "A Room of Our Own: Experiences from a Direct Office-Share'', in Pr*oc. ACM Conference on Human Factors in Computing Systems CHI'94* (Adelson, Dumais and Olson, eds), pp. 138–144, New York: ACM Press.

Bellotti, Victoria and Abigail Sellen, 1993. "Designing for Privacy in Ubiquitous Computing Environments'', in *Proc. European Conference on Computer-Supported Cooperative Work ECSCW'93* (de Michaelis, Simone and Schmidt, eds), pp 77–92, Amsterdam: Kluwer.

Blomberg, Jeanette, 1988. "Social Aspects of Operability: Ethnography of Photocopiers'', AAAS Symposium on Cognitive Ethnography of Industrial Society, Boston, Mass.

Bly, Sara, Steve Harrison and Susan Irwin, 1993. "Media Spaces: Bringing People Together in a Video, Audio and Computing Environment''*, Communications of the ACM*, 36(1), pp. 28–47.

Bowers, John and Tom Rodden, 1993. "Exploding the Interface: Experiences of a CSCW Network'', in *Proc. of InterCHI'93* (Ashlund et al, eds), pp. 255–262, New York: ACM Press.

Buxton, William and Tom Moran, 1990. *"EuroPARC's Integrated Interactive Intermedia Facility (IIIF): Early Experiences''*, Proc. IFIP Conference on Multi-User Interfaces and Applications, Herakleion, Crete.

Casner, Steve and Steve Deering, 1992. "First IETF Internet Audiocast'', ACM SIGCOMM Computer Communications Review, 22(3).

Cool, C., R. Fish, R. Kraut and C. Lowery, 1992. "Iterative Design of Video Communication Systems'', in *Proc. ACM Conference Computer-Supported Cooperative Work CSCW'92* (Turner and Kraut, eds), pp. 25–32, New York: ACM Press.

Curtis, Pavel and David Nichols, 1993. "MUDs Grow Up: Social Virtual Reality in the Real World'', Proc. Third International Conference on Cyberspace.

Dourish, Paul, 1991. *"Godard: A Flexible Architecture for A/V Services in a Media Space''*, EuroPARC Technical Report EPC-91-134, Rank Xerox EuroPARC, Cambridge, UK.

Dourish, Paul, 1993. "Culture and Control in a Media Space'', in *Proc. European Conference on Computer-Supported Cooperative Work ECSCW'93* (de Michaelis, Simone and Schmidt, eds), pp. 125–137, Amsterdam: Kluwer.

Dourish, Paul and Sara Bly, 1992. "Portholes: Supporting Awareness in Distributed Work Groups'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'92* (Bauersfeld, Bennett and Lynch, eds), pp. 541–548, New York: ACM Press.

Dykstra-Erickson, Elizabeth, Carrie Rudman, Catherine Marshall, Richard Hertz, Kartick Mithal and Jill Schmidt, 1995. "Supporting Adaptation to Multimedia Desktop Conferencing'', Proc. International Conference on Human Factors in Telecommunications, Melbourne, Australian, March.

Eriksson, Hans, 1994. "MBONE: The Multicast Backbone'', *Communications of the ACM*, 37(8), pp. 54–60.

Fish, Robert, Robert Kraut and Barbara Chalfonte, 1990. "The VideoWindow System in Informal Communication'', in *Proc. ACM Conference on Computer-Supported Cooperative Work CSCW'90*, pp. 1–11, New York: ACM Press.

Fish, Robert, Robert Kraut, Robert Root and Ronald Rice, 1993. " Video as a Technology for Informal Communication'', *Communications of the ACM*, 36(1), pp. 48–61.

Gaver, William, 1992. "The Affordances of Media Spaces for Collaboration'', in *Proc. ACM Conference on Computer-Supported Cooperative Work CSCW'92* (Turner and Kraut, eds), pp. 17–24, New York: ACM Press.

Gaver, William, Tom Moran, Allan MacLean, Lennart Lovstrand, Paul Dourish, Kathy Carter and Bill Buxton,1992. "Realising a Video Environment: EuroPARC's RAVE System'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'92* (Bauersfeld, Bennett and Lynch, eds), pp. 27–35, New York: ACM Press.

Gaver, William, Abigail Sellen, Christian Heath and Paul Luff, 1993. "One Is Not Enough: Multiple Views in a Media Space'', in *Proc. InterCHI'93* (Ashlund et al, eds), pp. 335–341, New York: ACM Press.

Gaver, William, Gerda Smets and Kees Overbeeke, 1995. "A Virtual Window on Media Space'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'95* (Katz, Mack and Marks, eds), pp. 257–264, New York: ACM Press.

Grudin, Jonathan, 1990. "interface'', in *Proc. ACM Conference on Computer-Supported Cooperative Work CSCW'90*, pp. 269–278, New York: ACM Press.

Harper, Richard and Kathleen Carter, 1994. "Keeping People Apart: A research note'', *Computer-Supported Cooperative Work,* 2(3), pp. 199–207.

Harrison, Steve, 1992. *"Making a Place in Media Space''*, Technical Report SSL-92-42, Xerox Palo Alto Research Center, Palo Alto, California.

Heath, Christian and Paul Luff, 1991. "Disembodied Conduct: Communication through video in a multi-media environment'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'91* (Robertson, Olson and Olson, eds), pp. 99–103, New York: ACM Press.

Heath, Christian and Paul Luff, 1992a. "Media Space and Communicative Asymmetries: Preliminary Observations of Video-Mediated Interaction", *Human-Computer Interaction*, 7(3), pp. 315–346.

Heath, Christian and Paul Luff, 1992b. "Collaboration and Control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms'', *Computer Supported Cooperative Work*, 1(1–2), pp. 69–95.

Hollan, James and Scott Stornetta,1992. "Beyond Being There'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'92* (Bauersfeld, Bennett and Lynch, eds), pp. 119–126, New York: ACM Press.

Ishii, Hiroshi, Minoru Kobayashi and Jonathan Grudin, 1992. “Integration of Inter-Personal Space and Shared Workspace: ClearBoard Design and Experiments'', in *Proc. ACM Conference on Computer-Supported Cooperative Work CSCW'92* (Turner and Kraut, eds), pp. 33–42, New York: ACM Press.

Louie, Gifford, Abigail Sellen and Marilyn Mantei, 1991. “*Making Contact in a Multi-media Environment''*, unpublished report, Computer Science Research Institute, University of Toronto.

Mantei, Marylin, Ron Baecker, Abigail Sellen, William Buxton, Tom Milligan and Barry Wellman, 1991. “Experiences in the Use of a Media Space'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'91* (Robertson, Olson and Olson, eds), pp. 203–208, New York: ACM Press.

Nardi, Bonnie, Heinrich Schwartz, Allan Kuchinsky, Robert Leichner, Steve Whittaker and Robert Sclabassi, 1993. “Turning Away from Talking Heads: The User of Video-As-Data in Neurosurgery'', in *Proc. InterCHI'93* (Ashlund et al, eds), pp. 327–334, New York: ACM Press.

O'Conaill, Brid, Steve Whittaker and Sylvia Wilbur, 1993. “Conversations over Video Conferences: An Evaluation of the Spoken Aspects of Video-Mediated Communication'', *Human-Computer Interaction*, 8(4), pp. 389–428.

Olson, Margarethe and Sara Bly, 1991. “The Portland Experience: A Report on a Distributed Research Group'', *Intl. Journal of Man-Machine Studies*, 34(2), pp. 211–228.

Pagani, Daniele and Wendy Mackay, 1993. “Bringing Media Spaces into the Real World'', in *Proc. Third European Conference on Computer-Supported Cooperative Work ECSCW'93* (de Michaelis, Simone and Schmidt, eds), pp. 341–356, Amsterdam: Kluwer.

Root, Robert, 1988. “Design of a Multi-Media Vehicle for Social Browsing'', in *Proc. ACM Conference on Computer Support for Cooperative Work CSCW '88*, pp. 25–38, New York: ACM Press.

Sellen, Abigail, 1992. “Speech Patterns in Video-Mediated Comversations'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'92* (Bauersfeld, Bennett and Lynch, eds), pp. 49–59, New York: ACM Press.

Spears, Russel and Martin Lea, 1993. “Social influence and the influence of the ‘social’ in computer-mediated communication'', in *“Contexts of Computer-Mediated Communication”* (Lea, ed), Harvester Wheatsheaf.

Stults, Robert, 1989. *“The Experimental Use of Video to Support Design Activity''*, Xerox PARC Technical Report SSL-89-19, Palo Alto, California.

Tang, John and Ellen Isaacs, 1993. “Why Do Users Like Video? Studies of Multimedia-Supported Collaboration'', *Computer Supported Cooperative Work*, 1(3), pp. 163–196.

Tyson, Philip, 1992. *“The Desk as a Social Institution''*, EuroPARC Technical Report EPC-92-130, Rank Xerox EuroPARC, Cambridge, UK.

Yamaashi, Kimaya, Yukihiro Kawamata and Masayuki Tani, 1995. “User-centered Video: Transmitting Video Images Based on the User's Interest'', in *Proc. ACM Conference on Human Factors in Computing Systems CHI'95* (Katz, Mack and Marks, eds), pp. 325–330, New York: ACM Press.