Chapter 4

Designing for collaboration and communication

4.1 Introduction

Imagine going into school or work each day and sitting in a room all by yourself with no distractions. At first, it might seem blissful. You'd be able to get on with your work. But what if you discovered you had no access to email, phones, the Internet and other people? On top of that there is nowhere to get coffee. How long would you last? Probably not very long. Humans are inherently social: they live together, work together, learn together, play together, interact and talk with each other, and socialize. It seems only natural, therefore, to develop interactive systems that support and extend these different kinds of sociality.

There are many kinds of sociality and many ways of studying it. In this chapter our focus is on how people communicate and collaborate in their working **and** everyday lives. We examine how collaborative technologies (also called groupware) have been designed to support and extend communication and collaboration. We also look at the social factors that influence the success or failure of user adoption of such technologies. Finally, we examine the role played by ethnographic studies and theoretical frameworks for informing system design.

The main aims of this chapter are to:

- Explain what is meant by communication and collaboration.
- Describe the main kinds of social mechanisms that are used by people to communicate and collaborate.
- Outline the range of collaborative systems that have been developed to support this kind of social behavior.
- Consider how field studies and socially-based theories can inform the design of collaborative systems.

4.2 Social mechanisms in communication and collaboration

A fundamental aspect of everyday life is talking, during which we pass on knowledge to each other. We continuously update each other about news, changes, and developments on a given project, activity, person, or event. For example, friends and families keep each other posted on what's happening at work, school, at the pub, at the club, next door, in soap operas, and in the news. Similarly, people who work together keep each other informed about their social lives and everyday happenings—as well as what is happening at work, for instance when a project is about to be completed, plans for a new project, problems with meeting deadlines, rumors about closures, and so on.

The kinds of knowledge that are circulated in different social circles are diverse, varying among social groups and across cultures. The frequency with which knowledge is disseminated is also highly variable. It can happen continuously throughout the day, once a day, weekly or infrequently. The means by which communication happens is also flexible—it can take place via face to face conversations, telephone, videophone, messaging, email, fax, and letters. Non-verbal communication also plays an important role in augmenting face to face conversation, involving the use of facial expressions, back channeling (the "aha's" and "umms"), voice intonation, gesturing, and other kinds of body language.

All this may appear self-evident, especially when one reflects on how we interact with one another. Less obvious is the range of social mechanisms and practices that have evolved in society to enable us to be social and maintain social order. Various rules, procedures, and etiquette have been established whose function is to let people know how they should behave in social groups. Below we describe three main categories of social mechanisms and explore how technological systems have been and can be designed to facilitate these:

- the use of conversational mechanisms to facilitate the flow of talk and help overcome breakdowns during it
- the use of coordination mechanisms to allow people to work and interact together
- the use of awareness mechanisms to find out what is happening, what others are doing and, conversely, to let others know what is happening

4.2.1 Conversational mechanisms

Talking is something that is effortless and comes naturally to most people. And yet holding a conversation is a highly skilled collaborative achievement, having many of the qualities of a musical ensemble. Below we examine what makes up a conversation. We begin by examining what happens at the beginning:

A: Hi there.

B: Hi!

C: Hi.

A: All right?

C: Good. How's it going?

A: Fine, how are you?

C: Good.

B: OK. How's life treating you?

Such mutual greetings are typical. A dialog may then ensue in which the participants take turns asking questions, giving replies, and making statements. Then when one or more of the participants wants to draw the conversation to a close, they do so by using either implicit or explicit cues. An example of an implicit cue is when a participant looks at his watch, signaling indirectly to the other participants that he wants the conversation to draw to a close. The other participants may choose to acknowledge this cue or carry on and ignore it. Either way, the first participant may then offer an explicit signal, by saying, "Well, I must be off now. Got work to do," or, "Oh dear, look at the time. Must dash. Have to meet someone." Following the acknowledgment by the other participants of such implicit and explicit signals, the conversation draws to a close, with a farewell ritual. The different participants take turns saying, "Bye," "Bye then," "See you," repeating themselves several times, until they finally separate.

Such conversational mechanisms enable people to coordinate their "talk" with one another, allowing them to know how to start and stop. Throughout a conversation further "turn-taking" rules are followed, enabling people to know when to listen, when it is their cue to speak, and when it is time for them to stop again to allow the others to speak. Sacks, Schegloff and Jefferson (1978)—who are famous for their work on conversation analysis—describe these in terms of three basic rules:

- rule 1—the current speaker chooses the next speaker by asking an opinion, question, or request
- rule 2—another person decides to start speaking
- rule 3—the current speaker continues talking

The rules are assumed to be applied in the above order, so that whenever there is an opportunity for a change of speaker to occur (e.g., someone comes to the end of a sentence), rule 1 is applied. If the listener to whom the question or opinion is addressed does not accept the offer to take the floor, the second rule is applied and

someone else taking part in the conversation may take up the opportunity and offer a view on the matter. If this does not happen then the third rule is applied and the current speaker continues talking. The rules are cycled through recursively until someone speaks again.

To facilitate rule following, people use various ways of indicating how long they are going to talk and on what topic. For example, a speaker might say right at the beginning of their turn in the conversation that he has three things to say. A speaker may also explicitly request a change in speaker by saying, "OK, that's all I want to say on that matter. So, what do you think?" to a listener. More subtle cues to let others know that their turn in the conversation is coming to an end include the lowering or raising of the voice to indicate the end of a question or the use of phrases like, "You know what I mean?" or simply, "OK?" Back channeling (uh-huh, mmm), body orientation (e.g., moving away from or closer to someone), gaze (staring straight at someone or glancing away), and gesture (e.g. raising of arms) are also used in different combinations when talking, to signal to others when someone wants to hand over or take up a turn in the conversation.

Another way in which conversations are coordinated and given coherence is through the use of adjacency pairs (Shegloff and Sacks, 1973). Utterances are assumed to come in pairs in which the first part sets up an expectation of what is to come next and directs the way in which what does come next is heard. For example, A may ask a question to which B responds appropriately:

A: So shall we meet at 8:00?

B: Um, can we make it a bit later, say 8:30?

Sometimes adjacency pairs get embedded in each other, so it may take some time for a person to get a reply to their initial request or statement:

A: So shall we meet at 8:00?

B: Wow, look at him.

A: Yes, what a funny hairdo!

B: Um, can we make it a bit later, say 8:30?

For the most part people are not aware of following conversational mechanisms, and would be hard pressed to articulate how they can carry on a conversation. Furthermore, people don't necessarily abide by the rules all the time. They may interrupt each other or talk over each other, even when the current speaker has clearly indicated a desire to hold the floor for the next two minutes to finish an argument. Alternatively, a listener may not take up a cue from a speaker to answer a question or take over the conversation, but instead continue to say nothing even though the speaker may be making it glaringly obvious it is the listener's turn to say something. Many a time a teacher will try to hand over the conversation to a student in a seminar, by staring at her and asking a specific question, only to see the student look at the floor, and say nothing. The outcome is an embarrassing silence, followed by either the teacher or another student picking up the conversation again.

Other kinds of breakdowns in conversation arise when someone says something that is ambiguous and the other person misinterprets it to mean something else. In

such situations the participants will collaborate to overcome the misunderstanding by using repair mechanisms. Consider the following snippet of conversation between two people:

A: Can you tell me the way to get to the Multiplex Ranger cinema?

B: Yes, you go down here for two blocks and then take a right (pointing to the right), go on till you get to the lights and then it is on the left.

A: Oh, so I go along here for a couple of blocks and then take a right and the cinema is at the lights (pointing ahead of him)?

A: No, you go on this street for a couple of blocks (gesturing more vigorously than before to the street to the right of him while emphasizing the word "this").

B: Ahhhh! I thought you meant that one: so it's this one (pointing in same direction as the other person).

A: Uh-hum, yes that's right, this one.

Detecting breakdowns in conversation requires the speaker and listener to be attending to what the other says (or does not say). Once they have understood the nature of the failure, they can then go about repairing it. As shown in the above example, when the listener misunderstands what has been communicated, the speaker repeats what she said earlier, using a stronger voice intonation and more exaggerated gestures. This allows the speaker to repair the mistake and be more explicit to the listener, allowing her to understand and follow better what they are saying. Listeners may also signal when they don't understand something or want further clarification by using various tokens, like "Huh?", "Quoi?" or "What?" (Schegloff, 1982) together with giving a puzzled look (usually frowning). This is especially the case when the speaker says something that is vague. For example, they might say "I want it" to their partner, without saying what it is they want. The partner may reply using a token or, alternatively, explicitly ask, "What do you mean by it?"

Taking turns also provides opportunities for the listener to initiate repair or request clarification, or for the speaker to detect that there is a problem and to initiate repair. The listener will usually wait for the next turn in the conversation before interrupting the speaker, to give the speaker the chance to clarify what is being said by completing the utterance (Suchman, 1987).

ACTIVITY 4.1 How do people repair breakdowns in conversations when using the phone or email?

Comment

In these settings people cannot see each other and so have to rely on other means of repairing their conversations. Furthermore, there are more opportunities for breakdowns to occur and fewer mechanisms available for repair. When a breakdown occurs over the phone, people will often shout louder, repeating what they said several times, and use stronger intonation. When a breakdown occurs via email, people may literally spell out what they meant, making things much more explicit in a subsequent email. If the message is beyond repair they may resort to another mode of communication that allows greater flexibility of expression, either telephoning or speaking to the recipient face to face.

Kinds of conversations

Conversations can take a variety of forms, such as an argument, a discussion, a heated debate, a chat, a tête-à-tête, or giving someone a "telling off." A well-known distinction in conversation types is between formal and informal communication. Formal communication involves assigning certain roles to people and prescribing *a priori* the types of turns that people are allowed to take in a conversation. For example, at a board meeting, it is decided who is allowed to speak, who speaks when, who manages the turn-taking, and what the participants are allowed to talk about.

In contrast, informal communication is the chat that goes on when people socialize. It also commonly happens when people bump into each other and talk briefly. This can occur in corridors, at the coffee machine, when waiting in line, and walking down the street. Informal conversations include talking about impersonal things like the weather (a favorite) and the price of living, or more personal things, like how someone is getting on with a new roommate. It also provides an opportunity to pass on gossip, such as who is going out to dinner with whom. In office settings, such chance conversations have been found to serve a number of functions, **including coordinating group work, transmitting knowledge about office culture,** establishing trust, and general team building (Kraut et al, 1990). It is also the case that people who are in physical proximity, such as those whose offices or desks are close to one another, engage much more frequently in these kinds of informal chats than those who are in different corridors or buildings. Most companies and organizations are well aware of this and often try to design their office space so that people who need to work closely together are placed close to one another in the same physical space.

4.2.2 Designing collaborative technologies to support conversation

As we have seen, "talk" and the way it is managed is integral to coordinating social activities. One of the challenges confronting designers is to consider how the different kinds of communication can be facilitated and supported in settings where there may be obstacles preventing it from happening "naturally." A central concern has been to develop systems that allow people to communicate with each other when they are in ***physically different locations*** and thus not able to communicate in the usual face to face manner. In particular, a key issue has been to determine how to allow people to carry on communicating as if they were in the same place, even though they are geographically separated—sometimes many thousands of miles apart.

Email, videoconferencing, videophones, computer conferencing, chatrooms and messaging are well-known examples of some of the collaborative technologies that have been developed to enable this to happen. Other less familiar systems are collaborative virtual environments (CVEs) and media spaces. CVEs are virtual worlds where people meet and chat. These can be 3D graphical worlds where users explore rooms and other spaces by teleporting themselves around in the guise of avatars (See Figure 4.1 on Color Plate 5), or text and graphical "spaces" (often called MUDs and MOOs) where users communicate with each other via some

form of messaging. Media spaces are distributed systems comprising audio, video, and computer systems that "extend the world of desks, chairs, walls and ceilings" (Harrison et al., 1997), enabling people distributed over space and time to communicate and interact with one another as if they were physically present. The various collaborative technologies have been designed to support different kinds of communication, from informal to formal and from one-to-one to many-to-many conversations. Collectively, such technologies are often referred to as computer-mediated communication (CMC).

ACTIVITY 4.2 Do you think it is better to develop technologies that will allow people to talk at a distance as if they were face to face, or to develop technologies that will support new ways of conversing?

Comment

On the one hand, it seems a good idea to develop technologies supporting people communicating at a distance that emulate the way they hold conversations in face to face situations. After all, this means of communicating is so well established and second nature to people. Phones and videoconferencing have been developed to essentially support face to face conversations. It is important to note, however, that conversations held in this way are not the same as when face to face. People have adapted the way they hold conversations to fit in with the constraints of the respective technologies. As noted earlier, they tend to shout more when misunderstood over the phone. They also tend to speak more loudly when talking on the phone, since they can't monitor how well the person can hear them at the other end of the phone. Likewise, people tend to project themselves more when videoconferencing. Turn-taking appears to be much more explicit, and greetings and farewells more ritualized.

On the other hand, it is interesting to look at how the new communication technologies have been extending the way people talk and socialize. For example, SMS text messaging has provided people with quite different ways of having a conversation at a distance. People (especially teenagers) have evolved a new form of fragmentary conversation (called "texting") that they continue over long periods. The conversation comprises short phrases that are typed in, using the key pad, commenting on what each is doing or thinking, allowing the other to keep posted on current developments. These kinds of "streamlined" conversations are coordinated simply by taking turns sending and receiving messages. Online chatting has also enabled effectively hundreds and even thousands of people to take part in the same conversations, which is not possible in face to face settings.

The range of systems that support computer-mediated communication is quite diverse. A summary table of the different types is shown in Table 4.1, highlighting how they support, extend and differ from face to face communication. A conventionally accepted classification system of CMC is to categorize them in terms of either synchronous or asynchronous communication. We have also included a third category: systems that support CMC in combination with other collaborative activities, such as meetings, decision-making, learning, and collaborative authoring of documents. Although some communication technologies are not strictly speaking computer-based (e.g., phones, video-conferencing) we have included these in the classification of CMC, as most now are display-based and interacted with or controlled via an interface. (For more detailed overviews of CMC, see Dix et al. (Chapter 13,1998) and Baecker et al. (Part III and IV, 1993).

Table 4.1 Classification of computer-mediated communication (CMC) into three types: (I) Synchronous communication, (ii) Asynchronous communication and (iii) CMC combined with other activity

i. Synchronous communication

Where conversations in real time are supported by letting people talk with each other either using their voices or through typing. Both modes seek to support non-verbal communication to varying degrees.

Examples:

- Talking with voice: video phones, video conferencing (desktop or wall), media spaces.
- Talking via typing: text messaging (typing in messages using cell phones), instant messaging (real-time interaction via PCs) chatrooms, collaborative virtual environments (CVEs).

New kinds of functionality:

- CVEs allow communication to take place via a combination of graphical representations of self (in the form of avatars) with a separate chatbox or overlaying speech bubbles.
- CVEs allow people to represent themselves as virtual characters, taking on new personas (e.g., opposite gender), and expressing themselves in ways not possible in face-to-face settings.
- CVEs, MUDs and chatrooms have enabled new forms of conversation mechanisms, such as multi-turn-taking, where a number of people can contribute and keep track of a multi-streaming text-based conversation.
- Instant messaging allows users to multitask by holding numerous conversations at once.

Benefits:

- Not having to physically face people may increase shy people's confidence and self-esteem to converse more in "virtual" public.
- It allows people to keep abreast of the goings-on in an organization without having to move from their office.
- It enables users to send text and images instantly between people using instant messaging.
- In offices, instant messaging allows users to fire off quick questions and answers without the time lag of email or phone-tag.

Problems:

- Lack of adequate bandwidth has plagued video communication, resulting in poor-quality images that frequently break up, judder, have shadows, and appear as unnatural images.
- It is difficult to establish eye contact (normally an integral and subconscious part of face-to-face conversations) in CVEs, video conferencing, and videophones.
- Having the possibility of hiding behind a persona, a name, or an avatar in a chatroom gives people the opportunity to behave differently. Sometimes this can result in people becoming aggressive or intrusive.

ii. Asynchronous communication

Where communication between participants takes place remotely and at different times. It relies not on time-dependent turn-taking but on participants initiating communication and responding to others when they want or are able to do so.

Examples:

- email, bulletin boards, newsgroups, computer conferencing

New kinds of functionality:

- Attachments of different sorts (including annotations, images, music) for email and computer conferencing can be sent.
- Messages can be archived and accessed using various search facilities.

Benefits:

- Ubiquity: Can read any place, any time.
- Flexibility: Greater autonomy and control of when and how to respond, so can attend to it in own time rather than having to take a turn in a conversation at a particular cue.
- Powerful: Can send the same message to many people.
- Makes some things easier to say: Do not have to interact with person so can be easier to say things than when face to face (e.g., announcing sudden death of colleague, providing feedback on someone's performance).

(Continued)

Table 4.1 *(Continued)*

Problems:

- Flaming: When a user writes incensed angry email expressed in uninhibited language that is much stronger than normally used when interacting with the same person face to face. This includes the use of impolite statements, exclamation marks, capitalized sentences or words, swearing, and superlatives. Such "charged" communication can lead to misunderstandings and bad feelings among the recipients.
- Overload: Many people experience message overload, receiving over 30 emails or other messages a day. They find it difficult to cope and may overlook an important message while working through their ever increasing pile of email—especially if they have not read it for a few days. Various interface mechanisms have been designed to help people manage their email better, including filtering, threading, and the use of signaling to indicate the level of importance of a message (via the sender or recipient), through color coding, bold font, or exclamation marks placed beside a message.
- False expectations: An assumption has evolved that people will read their messages several times a day and reply to them there and then. However, many people have now reverted to treating email more like postal mail, replying when they have the time to do so.

iii. CMC combined with other activity

People often talk with each other while carrying out other activities. For example, designing requires people to brainstorm together in meetings, drawing on whiteboards, making notes, and using existing designs. Teaching involves talking with students as well as writing on the board and getting students to solve problems collaboratively. Various meeting- and decision-support systems have been developed to help people work or learn while talking together.

Examples:

- Customized electronic meeting rooms have been built that support people in face-to-face meetings, via the use of networked workstations, large public displays, and shared software tools, together with various techniques to help decision-making. One of the earliest systems was the University of Arizona's GroupSystem (see Figure 4.2).

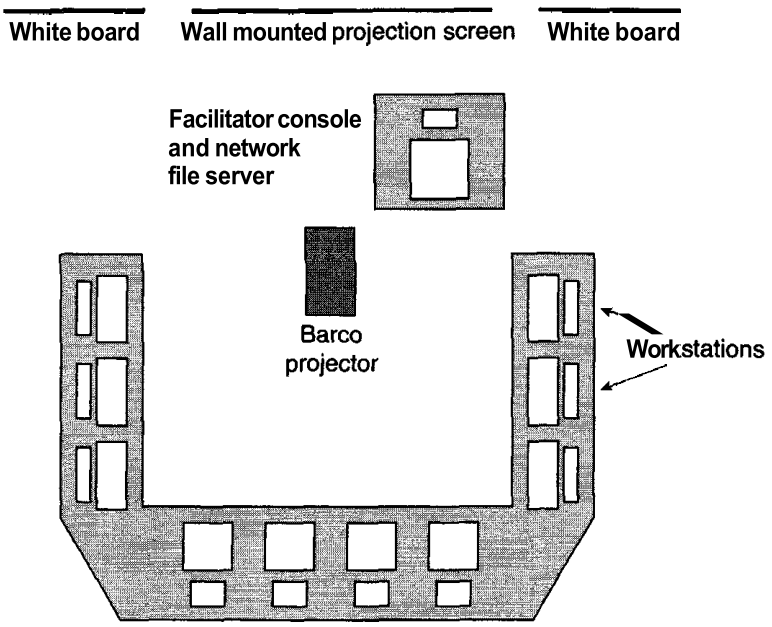


Figure 4.2 Schematic diagram of a group meeting room, showing relationship of workstation, whiteboards and video projector.

(Continued)

Table 4.1 (Continued)

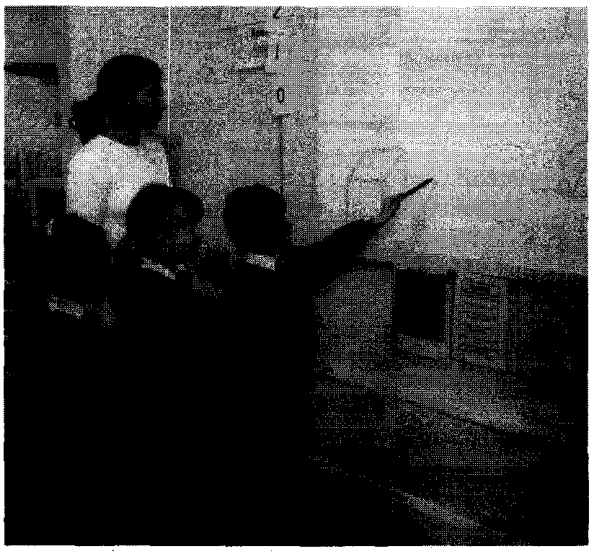


Figure 4.3 An ACTIVBoard whiteboard developed by Promethean (U.K. company) that allows children to take control of the front-of-class display. This allows them to add comments and type in queries, rather than having to raise their hands and hope the teacher sees them.

- Networked classrooms: Recently schools and universities have realized the potential of using combinations of technologies to support learning. For example, wireless communication, portable devices and interactive whiteboards are being integrated in classroom settings to allow the teacher and students to learn and communicate with one another in novel interactive ways (see Figure 4.3).
- Argumentation tools which record the design rationale and other arguments used in a discussion that lead to decisions in a design (e.g. gIBIS, Conklin and Begeman, 1989). These are mainly designed for people working in the same physical location.
- Shared authoring and drawing tools that allow people to work on the same document at the same time. This can be remotely over the web (e.g., shared authoring tools like Shredit) or on the same drawing surface in the same room using multiple mouse cursors (e.g., KidPad, Benford et al., 2000).

New kinds of functionality:

- Allows new ways of collaboratively creating and editing documents.
- Supports new forms of collaborative learning.
- Integrates different kinds of tools.

Benefits:

- Supports talking while carrying out other activities at the same time, allowing multi-tasking—which is what happens in face-to-face settings.
  Speed and efficiency: allows multiple people to be working an same document at same time.
- Greater awareness: allows users to see how one another are progressing in real time.

Problems:

- WYSIWIS (what you see is what I see): It can be difficult to see what other people are referring to when in remote locations, especially if the document is large and different users have different parts of the document on their screens.
- Floor control: Users may want to work on the same piece of text or design, potentially resulting in file conflicts. These can be overcome by developing various social and technological floor-control policies.

ACTIVITY 4.3 One of the earliest technological innovations (besides the telephone and telegraph) developed for supporting conversations at a distance was the videophone. Despite numerous attempts by the various phone companies to introduce them over the last 50 years (see Figure 4.4), they have failed each time. Why do you think this is so?

Comment

One of the biggest problems with commercial videophones is that the bandwidth is too low, resulting in poor resolution and slow refresh rate. The net effect is the display of unacceptable images: the person in the picture appears to move in sudden jerks; shadows are left behind when a speaker moves, and it is difficult to read lips or establish eye contact. There is also the social acceptability issue of whether people want to look at pocket-sized images of each other when talking. Sometimes you don't want people to see what state you are in or where you are.

Another innovation has been to develop systems that allow people to communicate and interact with each other in ways not possible in the physical world. Rather than try to imitate or facilitate face to face communication (like the above systems), designers have tried to develop new kinds of interactions. For example, ClearBoard was developed to enable facial expressions of participants to be made visible to others by using a transparent board that showed their face to the others (Ishii et al., 1993). HyperMirror was designed to provide an environment in which the participants could feel they were in the same virtual place even

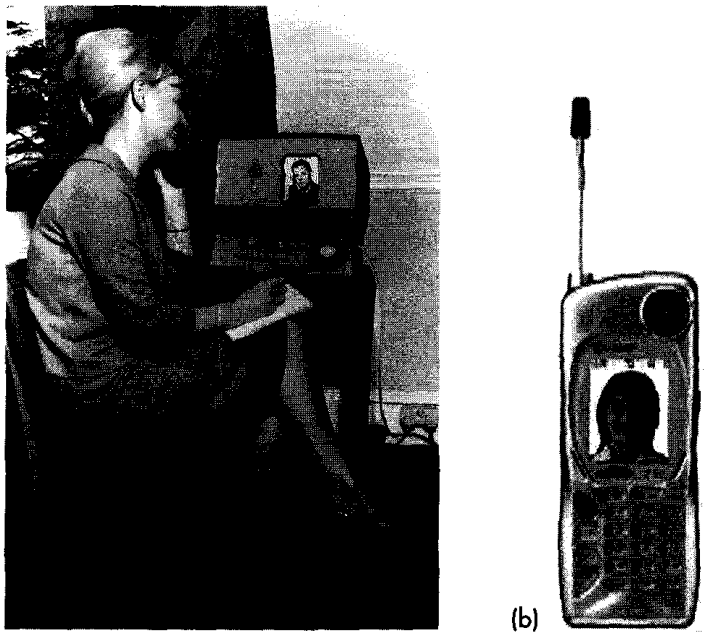


Figure 4.4 (a) One of British Telecom's early videophones and (b) a recent mobile "visualphone" developed in Japan.

BOX 4.1 Supporting Informal Chatting through Audio-Video Links

A number of researchers have tried to capitalize on the social phenomenon of informal communication and the important role it plays at work. In particular, they have been interested in finding ways of using audio-video links to *mimic* physical settings that are conducive to informal communication for people who are geographically separated. One of the first systems to be built, at Bellcore in 1989, was the *VideoWindow System* (see Figure 4.5). The goal was to design a shared space that would allow people in different locations to carry on a conversation as they would do if sitting in the same room drinking coffee together. Two lounge areas that were 50 miles apart were connected with high-bandwidth video channels and full-duplex four-channel audio. Connecting them was a 3 × 8 foot "picture-window" onto which video images were projected. The large size was meant to allow viewers to see a room of people roughly the same size as themselves. The system was designed to be active 24 hours a day, so that anyone entering one room could speak to whoever happened to be in the other room.

A study by Kraut et al. (1990) of how effective the system was showed that in general, many of the interactions that took place between the remote conversants were indeed indistinguishable from similar face-to-face interactions—with the exception that they spoke a bit louder and constantly talked about the video system. However, they also found that people who were in the same room tended to talk more with each other than with those in the video-linked room. Various usability problems were identified as contributing to this reluctance to talk with video images of other people. One of these was the tendency for people to move closer to the picture window to strike up a conversation with someone (which is what one would normally do in a face-to-face setting); this had the opposite effect to what the person intended, as it moved his or her head out of the picture and also out of microphone range, meaning he or she could not be seen or heard. Thus rather than getting nearer to the other person, this behavior had the counter-intuitive effect of removing him or her from the "picture." Moreover, there was no way for participants to know whether they were being seen and heard by the others in the other room. This inability to monitor how others are or are not "receiving" you caused numerous problems. Another problem was that the system allowed only public conversations, meaning that

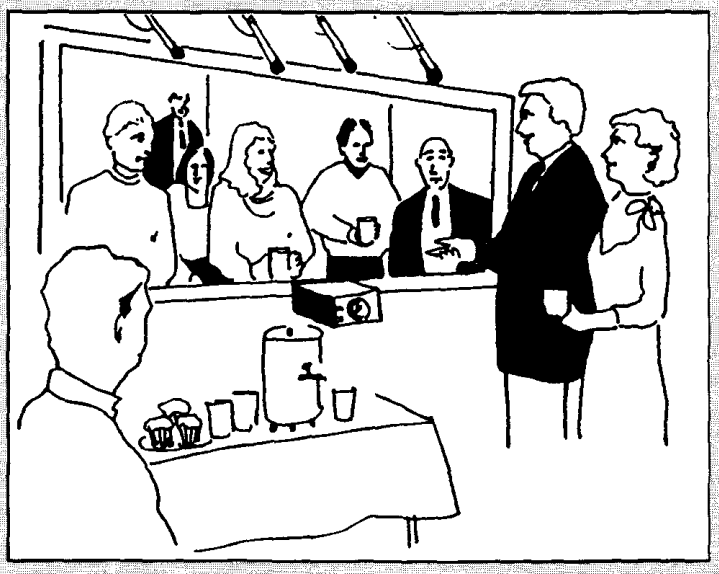


Figure 4.5 Diagram of VideoWindow system in use.

they could be heard by everyone in the rooms. Such public broadcasting contrasts with how people normally engage in informal face-to-face conversations, where they will often whisper and conspire with each other when a topic becomes more private or secret. Such private conversations clearly could not be supported by the VideoWindow system.

A number of other synchronous audio-video systems have since been developed that have tried to incorporate different kinds of conversational mechanisms to facilitate informal communication in video-linked locations. For example, Cruiser was designed to support informal communication by placing separate audio and video equipment on the desktop of each person who was connected to the system (Fish, 1989). This set-up differed from the VideoWindow system in that it enabled both public and private interactions to take place. It also provided additional functionality that allowed people to initiate conversations by typing in a cruise command followed by a question like, "I'm bored. Anyone want a chat?" or "Can someone help me?"—the aim here being explicitly to encourage people to engage in the kind of talk that they normally do when they bump into each other, but this time over the computer network. A further conversation mechanism built into Cruiser was a "glance" feature that allowed users to check whether the person they wanted to talk to was in fact available before trying to initiate a conversation.

Commercial systems are now available that support multiple connections among sites. These can be very useful for virtual centers that have multiple groups working at a number of different sites. For example, the Distributed Systems Technology Center (a research collaboration between Australian universities and industries) has been using a commercial videoconferencing system (Polycom Viewstation 128) that enables people at its main sites (Brisbane, Sydney, and Melbourne) to keep in contact via regular informal and formal meetings. Formal meetings among the sites involve all the teams (over a hundred people) getting together and reporting on their projects. Each site has a camera which it controls, projecting different images to the other sites of what is happening at the site. This can be an image of the person who is doing the talking, a person being talked about, or a person who is being funny. The images from the different sites are displayed side by side on a large screen at each site. Weekly informal project meetings are also held among small groups. Informal chatting is also supported by having a continuous virtual presence via a video wall of the kitchen of one Queensland University site in another, and vice versa (see Figure 4.6).

Figure 4.6 A commercial videoconferencing system being used to support informal chatting among researchers across sites at Queensland University. In contrast to the VideoWindow system, a window of each site is shown in the top left-hand corner of their display to let the participants monitor their own behavior.

(a) (b) (c)

Figure 4.7 Hypermirror in action, showing perception of virtual personal space. (a) A woman is in one room (indicated by arrow on screen), (b) while a man and another woman in the other room chat to each other. They move apart when they notice they are "overlapping" her and (c) virtual personal space is established.

though they were physically in different places (Morikawa and Maesako, 1998). Mirror reflections of people in different places were synthesized and projected onto a single screen, so that they appeared side by side in the same virtual space. In this way, the participants could see both themselves and others in the same seamless virtual space. Observations of people using the system showed how quickly they adapted to perceiving themselves and others in this way. For example, participants quickly became sensitized to the importance of virtual personal space, moving out of the way if they perceived they were overlapping someone else on the screen (see Figure 4.7).

4.2.3 Coordination mechanisms

Coordination takes place when a group of people act or interact together to achieve something. For example, consider what is involved in playing a game of basketball. Teams have to work out how to play with each other and to plan a set of tactics that they think will outwit the other team. For the game to proceed both teams need to follow (and sometimes contravene) the rules of the game. An incredible amount of coordination is required within a team and between the competing teams in order to play.

In general, collaborative activities require us to coordinate with each other, whether playing a team game, moving a piano, navigating a ship, working on a large software project, taking orders and serving meals in a restaurant, constructing a bridge or playing tennis. In particular, we need to figure out how to interact with one another to progress with our various activities. To help us we use a number of coordinating mechanisms. Primarily, these include:

- verbal and non-verbal communication
- schedules, rules and conventions
- shared external representations

Verbal and non-verbal communication

When people are working closely together they talk to each other, issuing commands and letting others know how they are progressing with their part. For example, when two or more people are collaborating together, as in moving a piano, they shout to each other commands like "Down a bit, left a bit, now straight forward" to coordinate their actions with each other. As in a conversation, nods, shakes, winks, glances, and hand-raising are also used in combination with such coordination "talk" to emphasize and sometimes replace it.

In formal settings, like meetings, explicit structures such as agendas, memos, and minutes are employed to coordinate the activity. Meetings are chaired, with secretaries taking minutes to record what is said and plans of actions agreed upon. Such minutes are subsequently distributed to members to remind them of what was agreed in the meeting and for those responsible to act upon what was agreed.

For time-critical and routinized collaborative activities, especially where it is difficult to hear others because of the physical conditions, gestures are frequently used (radio-controlled communication systems may also be used). Various kinds of hand signals have evolved, with their own set of standardized syntax and semantics. For example, the arm and baton movements of a conductor coordinate the different players in an orchestra, while the arm and baton movements of a ground marshal at an airport signal to a pilot how to bring the plane into its allocated gate.

ACTIVITY 4.4 How much communication is non-verbal? Watch a soap opera on the TV and turn down the volume and look at the kinds and frequency of gestures that are used. Are you able to understand what is going on? How do radio soaps compensate for not being able to use non-verbal gestures? How do people compensate when chatting online?

Comment

Soaps are good to watch for observing non-verbal behavior as they tend to be overcharged, with actors exaggerating their gestures and facial expressions to convey their emotions. It is often easy to work out what kind of scene is happening from their posture, body movement, gestures, and facial expressions. In contrast, actors on the radio use their voice a lot more, relying on intonation and surrounding sound effects to help convey emotions. When chatting online, people use emoticons and other specially evolved verbal codes.

Schedules, rules, and conventions

A common practice in organizations is to use various kinds of schedules to organize the people who are part of it. For example, consider how a university manages to coordinate the people within it with its available resources. A core task is allocating the thousands of lectures and seminars that need to be run each week with the substantially smaller number of rooms available. A schedule has to be devised

that allows students to attend the lectures and seminars for their given courses, taking into account numerous rules and constraints. These include:

- A student cannot attend more than one lecture at a given time.
- A professor cannot give more than one lecture or seminar at a given time.
- A room cannot be allocated to more than one seminar or lecture at a given time.
- Only a certain number of students can be placed in a room, depending on its size.

BOX 4.2 Shared Calendars—My Time Or Yours?

Trying to schedule meetings for different people in an organization to attend can be a nightmare. Typically, a secretary will send out email or try to call those who need to be assembled for a given meeting. Some of those people may not be there at the time they are contacted and so the secretary must wait for them to reply before being able to set up a meeting. In the meantime, the others who have replied (but have not heard back about the meeting) may start to fill up the times they said they would be free with other engagements. When the secretary eventually gets back to them with a proposed date, it is often too late. The consequence is that the secretary will have to start again, coming up with a new set of times. The more people to be scheduled, the more difficult it becomes to keep track of people's calendars in order to find mutually free time slots to arrange a meeting. All in all, it is a very time-consuming and laborious activity, one in which it would seem a computer-based scheduling tool would be very useful.

Indeed, a number of shared calendars have been developed. Some of the most recent ones have been developed as web applications, allowing individuals to use them both as *personal calendars*, reminding them of events they have to attend, and as *public calendars*, broadcasting to anyone accessing their web pages, when they are busy and when they are free.

A number of studies of the implementation of shared calendars in various organizations have found them to be successful computer-supported coordination tools. For example, a study of one system, called MeetingMaker, showed that the users found it had greatly simplified meeting scheduling and was much faster than manual scheduling (Mosier and Tammaro, 1997). The shared tool provided a range of facilities, including supporting group scheduling, resource and room scheduling, to-do lists and various permissions to permit others to schedule meetings. The system could also inform individual users of upcoming events and meetings via a pop-up dialog box.

Other shared calendars, however, have not been so successful. These are usually ones that have been designed to allow people to look at a person's schedule, and upon finding a free slot, book a meeting with them without any form of negotiation. The normal etiquette when arranging a meeting is to ask someone when they are free to meet and suggest a number of dates and times. When a meeting is simply imposed, people can feel that their privacy has been invaded, especially if they had planned to use the allocated slot to work on something else. A typical response is to simply stop using the shared calendar. The problem with one person doing this, however, is that the rest of the group cannot continue using the shared calendar as a coordinating tool, since that person is now excluded.

The more successful scheduling tools, like MeetingMaker, have overcome this sensitive privacy issue by providing users with a proxy function. This allows users to mark off parts of their calendars as "private" while giving permissions to others to read and/or write to their calendars. Providing this more flexible control allows users to decide *a priori* which times of the week they will make themselves available for meetings and which times they want to keep private to get on with their own work, without revealing to others what they are up to.

Other coordinating mechanisms that are employed by groups working together are rules and conventions. These can be formal or informal. Formal rules, like the compulsory attendance of seminars, writing monthly reports, and filling in of timesheets, enable organizations to maintain order and keep track of what its members are doing. Conventions, like keeping quiet in a library or removing meal trays after finishing eating in a cafeteria, are a form of courtesy to others.

Shared external representations

Shared external representations are commonly used to coordinate people. We have already mentioned one example, that of shared calendars that appear on user's monitors as graphical charts, email reminders, and dialog boxes. Other kinds that are commonly used include forms, checklists, and tables. These are presented on public noticeboards or as part of other shared spaces. They can also be attached to documents and folders. They function by providing external information of who is working on what, when, where, when a piece of work is supposed to be finished, and who it goes to next. For example, a shared table of who has completed the checking of files for a design project (see Figure 4.8), provides the necessary information from which other members of the group can at a glance update their model of the current progress of that project. Importantly, such external representations can be readily updated by annotating. If a project is going to take longer than planned, this can be indicated on a chart or table by extending the line representing it, allowing others to see the change when they pass by and glance up at the whiteboard.

Shared externalizations allow people to make various inferences about the changes or delays with respect to their effect on their current activities. Accordingly,

Sheet no	Gary copied in	Kate & Gary plot file created	Mark checked by Phil	Kate plot sent	Mark plot file created	Mark plot sent mylar
596S6	✓	✓				
S7	✓	✓				
S8	✓					
S9	✓					
S10	✓					

Figure 4.8 An external representation used to coordinate collaborative work in the form of a print-out table showing who has completed the checking of files and who is down to do what.

they may need to reschedule their work and annotate the shared workplan. In so doing, these kinds of coordination mechanisms are considered to be tangible, providing important representations of work and responsibility that can be changed and updated as and when needed.

4.2.4 Designing collaborative technologies to support coordination

Shared calendars, electronic schedulers, project management tools, and workflow tools that provide interactive forms of scheduling and planning are some of the main kinds of collaborative technologies that have been developed to support coordination. A specific mechanism that has been implemented is the use of conventions. For example, a shared workspace system (called POLITeam) that supported email and document sharing to allow politicians to work together at different sites introduced a range of conventions. These included how folders and files should be organized in the shared workspace. Interestingly, when the system was used in practice, it was found that the conventions were often violated (Mark, et al., 1997). For example, one convention that was set up was that users should always type in the code of a file when they were using it. In practice, very few people did this, as pointed out by an administrator: "They don't type in the right code. I must correct them. I must sort the documents into the right archive. And that's annoying".

The tendency of people not to follow conventions can be due to a number of reasons. If following conventions requires additional work that is extraneous to the users' ongoing work, they may find it gets in the way. They may also perceive the convention as an unnecessary burden and "forget" to follow it all the time. Such "productive laziness" (Rogers, 1993) is quite common. A simple analogy to everyday life is forgetting to put the top back on the toothpaste tube: it is a very simple convention to follow and yet we are all guilty sometimes (or even all the time) of not doing this. While such actions may only take a tiny bit of effort, people often don't do them because they perceive them as tedious and unnecessary. However, the consequence of not doing them can cause grief to others.

When designing coordination mechanisms it is important to consider how socially acceptable they are to people. Failure to do so can result in the users not using the system in the way intended or simply abandoning it. A key part is getting the right balance between human coordination and system coordination. Too much system control and the users will rebel. Too little control and the system breaks down. Consider the example of file locking, which is a form of concurrency control. This is used by most shared applications (e.g., shared authoring tools, file-sharing systems) to prevent users from clashing when trying to work on the same part of a shared document or file at the same time. With file locking, whenever someone is working on a file or part of it, it becomes inaccessible to others. Information about who is using the file and for how long may be made available to the other users, to show why they can't work on a particular file. When file-locking mechanisms are used in this way, however, they are often considered too rigid as a form of coordination, primarily because they don't let other users negotiate with the first user about when they can have access to the locked file.

A more flexible form of coordination is to include a social policy of floor control. Whenever a user wants to work on a shared document or file, he must initially request "the floor." If no one else is using the specified section or file at that time, then he is given the floor. That part of the document or file then becomes locked, preventing others from having access to it. If other users want access to the file, they likewise make a request for the floor. The current user is then notified and can then let the requester know how long the file will be in use. If not acceptable, the requester can try to negotiate a time for access to the file. This kind of coordination mechanism, therefore, provides more scope for negotiation between users on how to collaborate, rather than simply receiving a point-blank "permission denied" response from the system when a file is being used by someone else.

BOX 4.3 Turning Technology Inside Out: Online versus Physical Coordination Mechanisms

Many software applications now exist to support coordination, notably project management systems. From the project manager's perspective, they provide a flexible means of scheduling, distributing, and monitoring collaborative work and enabling them continuously to remind people of deadlines and milestones via the use of email and other kinds of representations. From the perspective of the individuals working in the company, they give them a means of letting others know when they are available for meetings and where they will be.

In practice, however, project management systems that rely exclusively on computer-mediated coordination mechanisms have not been found to be as effective as hoped. This tends to happen when the system is used to coordinate a large number of events or projects. People begin not to take notice of the numerous internal reminders and messages that are sent to them by the system, finding them to be too intrusive, overwhelming, or annoying. This can then lead to missing important meetings and deadlines. A work-around in some organizations has been to print out the schedules and events that have been entered into the project management online database and display them as paper-based external representations (see Whittaker and Schwartz, 1995). A study that looked at the creation and use of shared external representations in collaborative work (Bellotti and Rogers, 1997) found that in several cases, information that is represented online becomes re-represented as a physical entity because the online version often gets lost, forgotten, or overlooked. This was particularly prevalent at new media companies producing web content that needed to be updated regularly. The various groups had to be coordinated across a number of parallel-running, time-critical projects.

At one site, a project coordinator would write up on a physical whiteboard every morning the main projects, schedules, and deadlines relevant for that day extracted from the online project management software. When asked why she laboriously wrote down by hand information that could be readily accessed by everyone over the computer network, she replied that, owing to the multiplication of projects and people working on them, it had become very difficult to keep track of everything that was going on. Moreover, people had become de-sensitized to the many email reminders that the software application provided, so they often forgot their significance immediately after having acknowledged them. Consequently, everyone (including herself) needed to be reminded of what was urgent and what needed dealing with that day. Placing this critical information on a physical whiteboard in a prominent public place that was clearly distinct from the continuous stream of other online information and messages provided a more effective public reminder of what was urgent and needed doing that day. In essence, the company had resorted to "*turning the technology inside out.*"

ACTIVITY 4.5 Why are whiteboards so useful for coordinating projects? How might electronic whiteboards be designed to extend this practice?

Comment

Physical whiteboards are very good as coordinating tools as they display information that is external and public, making it highly visible for everyone to see. Furthermore, the information can be easily annotated to show up-to-date modifications to a schedule. Whiteboards also have a gravitational force, drawing people to them. They provide a meeting place for people to discuss and catch up with latest developments.

Electronic whiteboards have the added advantage that important information can be animated to make it stand out. Important information can also be displayed on multiple displays throughout a building and can be extracted from existing databases and software, thereby making the project coordinator's work much easier. The boards could also be used to support on-the-fly meetings in which individuals could use electronic pens to sketch out ideas-that could then be stored electronically. In such settings they could also be interacted with via wireless handheld computers, allowing information to be "scraped" off or "squirted onto the whiteboard.

4.2.5 Awareness mechanisms

Awareness involves knowing who is around, what is happening, and who is talking with whom (Dourish and Bly, 1992). For example, when we are at a party, we move around the physical space, observing what is going on and who is talking to whom, eavesdropping on others' conversations and passing on gossip to others. A specific kind of awareness is peripheral awareness. This refers to a person's ability to maintain and constantly update a sense of what is going on in the physical and social context, through keeping an eye on what is happening in the periphery of their vision. This might include noting whether people are in a good or bad mood by the way they are talking, how fast the drink and food is being consumed, who has entered or left the room, how long someone has been absent, and whether the lonely guy in the corner is finally talking to someone—all while we are having a conversation with someone else. The combination of direct observations and peripheral monitoring keeps people informed and updated of what is happening in the world.

Similar ways of becoming aware and keeping aware take place in other contexts, such as a place of study or work. Importantly, this requires fathoming when is an appropriate time to interact with others to get and pass information on. Seeing a professor slam the office door signals to students that this is definitely not a good time to ask for an extension on an assignment deadline. Conversely, seeing teachers with beaming faces, chatting openly to other students suggests they are in a good mood and therefore this would be a good time to ask them if it would be all right to miss next week's seminar because of an important family engagement. The knowledge that someone is amenable or not rapidly spreads through a company, school, or other institution. People are very eager to pass on both good and bad news to others and will go out of their way to gossip, loitering in corridors, hanging around at the photocopier and coffee machine "spreading the word."

Figure 4.9 An external representation used to signal to others a person's availability.

In addition to monitoring the behaviors of others, people will organize their work and physical environment to enable it to be successfully monitored by others. This ranges from the use of subtle cues to more blatant ones. An example of a subtle cue is when someone leaves their dorm or office door slightly ajar to indicate that they can be approached. A more blatant one is the complete closing of their door together with a "do not disturb" notice prominently on it, signaling to everyone that under no circumstances should they be disturbed (see Figure **4.9).**

Overhearing and overseeing

People who work closely together also develop various strategies for coordinating their work, based on an up-to-date awareness of what the others are doing. This is especially so for interdependent tasks, where the outcome of one person's activity is needed for others to be able to carry out their tasks. For example, when putting on a show, the performers will constantly monitor what one another is doing in order to coordinate their performance efficiently.

The metaphorical expression "closely-knit teams" exemplifies this way of collaborating. People become highly skilled in reading and tracking what others are doing and the information they are attending to. A well-known study of this phenomenon is described by Christian Heath and Paul Luff (1992), who looked at how two controllers worked together in a control room in the London Underground. An overriding observation was that the actions of one controller were tied very closely to what the other was doing. One of the controllers was responsible for the movement of trains on the line (controller **A),** while the other was responsible for providing information to passengers about the current service (controller B). In many instances, it was found that controller B overheard what controller **A** was doing and saying, and acted accordingly—even though controller A had not said anything explicitly to him. For example, on overhearing controller A discussing a problem with a train driver over the in-cab intercom system, controller B inferred from the ensuing conversation that there was going to be a disruption to the service

and so started announcing this to the passengers on the platform before controller A had even finished talking with the train driver. At other times, the two controllers keep a lookout for each other, monitoring the environment for actions and events which they might have not noticed but may be important for them to know about so that they can act appropriately.

ACTIVITY 4.6 What do you think happens when one person of a closely knit team does not see or hear something or misunderstands what has been said, while the others in the group assume they have seen, heard, or understood what has been said?

Comment In such circumstances, the person is likely to carry on as normal. In some cases this will result in inappropriate behavior. Repair mechanisms will then need to be set in motion. The knowledgeable participants may notice that the other person has not acted in the manner expected. They may then use one of a number of subtle repair mechanisms, say coughing or glancing at something that needs attending to. If this doesn't work, they may then resort to explicitly stating aloud what had previously been signaled implicitly. Conversely, the unaware participant may wonder why the event hasn't happened and, likewise, look over at the other people, cough to get their attention or explicitly ask them a question. The kind of repair mechanism employed at a given moment will depend on a number of factors, including the relationship among the participants (e.g., whether one is more senior than the others—this determines who can ask what), perceived fault or responsibility for the breakdown and the severity of the outcome of not acting there and then on the new information.

4.2.6 Designing collaborative technologies to support awareness

The various observations about awareness have led system developers to consider how best to provide awareness information for people who need to work together but who are not in the same physical space. Various technologies have been employed along with the design of specific applications to convey information about what people are doing and the progress of their ongoing work. As mentioned previously, audio-video links have been developed to enable remote colleagues to keep in touch with one another. Some of these systems have also been developed to provide awareness information about remote partners, allowing them to find out what one another is doing. One of the earliest systems was Portholes, developed at Xerox PARC research labs (Dourish and Bly, 1992). The system presented regularly-updated digitized video images of people in their offices from a number of different locations (in the US and UK). These were shown in a matrix display on people's workstations. Clicking on one of the images had the effect of bringing up a dialog box providing further information about that individual (e.g., name, phone number) together with a set of lightweight action buttons (e.g., email the person, listen to a pre-recorded audio snippet). The system provided changing images of people throughout the day and night in their offices, letting others see at a glance whether they were in their offices, what they were working on, and who was around (see Figure 4.10). Informal evaluation of the

Figure 4.10 A screen dump of Portholes, showing low resolution monochrome images from offices in the US and UK PARC sites. (Permission from Xerox Research Centre, Europe)

set-up suggested that having access to such information led to a shared sense of community.

The emphasis in the design of these early awareness systems was largely on supporting peripheral monitoring, allowing people to see each other and their progress. Dourish and Bellotti (1992) refer to this as shared feedback. More recent distributed awareness systems provide a different kind of awareness information. Rather than place the onus on participants to find out about each other, they have been designed to allow users to notify each other about specific kinds of events. Thus, there is less emphasis on monitoring and being monitored and more on explicitly letting others know about things. Notification mechanisms are also used to provide information about the status of shared objects and the progress of collaborative tasks.

Hence, there has been a shift towards supporting a collective "stream of consciousness" that people can attend to when they want to, and likewise provide information for when they want to. An example of a distributed awareness system is Elvin, developed at the University of Queensland (Segall and Arnold, 1997), which provides a range of client services. A highly successful client is Tickertape, which is a lightweight instant messaging system, showing small color-coded messages that scroll from right to left across the screen (Fitzpatrick et al., 1999). It has been most useful as a "chat" and local organizing tool, allowing people in different locations to effortlessly send brief messages and requests to the public tickertape display (see Figure 4.11). It has been used for a range of functions, including organizing shared

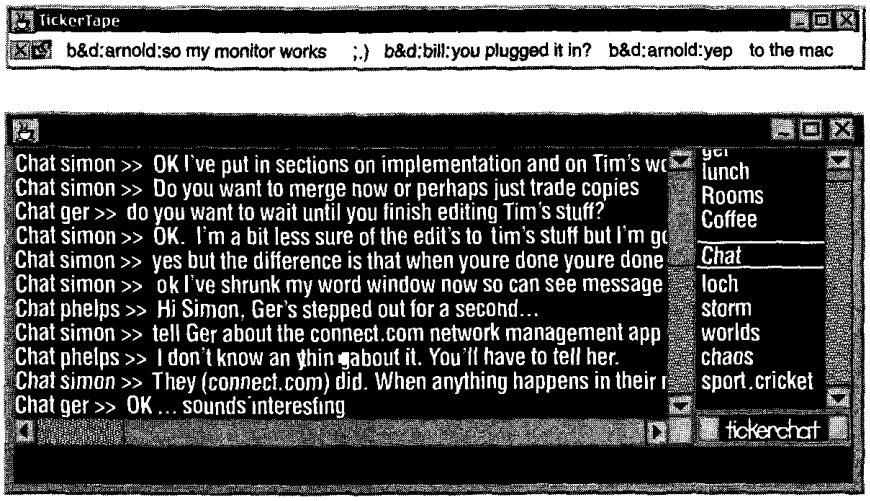


Figure 4.11 The Tickertape and Tickerchat interface for ELVIN awareness service.

events (e.g. lunch dates), making announcements, and as an "always-on" communication tool for people working together on projects but who are not physically co-located. It is also often used as a means of mediating help between people. For example, when I was visiting the University of Queensland, I asked for help over Tickertape. Within minutes, I was inundated with replies from people logged onto the system who did not even know me. At the time, I was having problems working out the key mappings between the PC that I was using in Australia and a Unix editor I couldn't find a way of quitting from on a remote machine in the UK. The suggestions that appeared on Tickertape quickly led to a discussion among the participants, and within five minutes someone had come over to my desk and sorted the problem out for me!

In addition to presenting awareness information as streaming text messages, more abstract forms of representation have been used. For example, a communication tool called Babble, developed at IBM (Erickson et al., 1999), provides a dynamic visualization of the participants in an ongoing chat-like conversation. A large 2D circle is depicted with colored marbles on each user's monitor. Marbles inside the circle convey those individuals active in the current conversation. Marbles outside the circle convey users involved in other conversations. The more active a participant is in the conversation, the more the corresponding marble is moved towards the center of the circle. Conversely, the less engaged a person is in the ongoing conversation, the more the marble moves towards the periphery of the circle (see Figure 4.12).

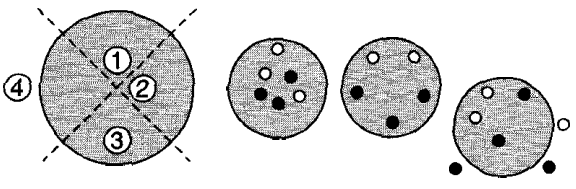


Figure 4.12 The Babble interface, with dynamic visualization of participants in ongoing conversation.

4.3 Ethnographic studies of collaboration and communication

One of the main approaches to informing the design of collaborative technologies that takes into account social concerns is carrying out an ethnographic study (a type of field study). Observations of the setting, be it home, work, school, public place, or other setting, are made, examining the current work and other collaborative practices people engage in. The way existing technologies and everyday artifacts are used is also analyzed. The outcome of such studies can be very illuminating, revealing how people currently manage in their work and everyday environments. They also provide a basis from which to consider how such existing settings might be improved or enhanced through the introduction of new technologies, and can also expose problematic assumptions about how collaborative technologies will or should be used in a setting (for more on ***how to use*** ethnography to inform design, see Chapter 9; ***how to do*** ethnography is covered in Chapter 12).

Many studies have analyzed in detail how people carry out their work in different settings (Plowman et al., 1995). The findings of these studies are used both to inform the design of a ***specific*** system, intended for a particular workplace, and more generally, to provide input into the design of new technologies. They can also highlight problems with existing system design methods. For example, an early study by Lucy Suchman (1983) looked at the way existing office technologies were being designed in relation to how people actually worked. She observed what really happened in a number of offices and found that there was a big mismatch between the way work was actually accomplished and the way people were supposed to work using the office technology provided. She argued that designers would be much better positioned to develop systems that could match the way people behave and use technology, if they began by considering the ***actual details*** of work practice.

In her later, much-cited study of how pairs of users interacted with an interactive help system—intended as a facility for using with a photocopier—Suchman (1987) again stressed the point that the design of interactive systems would greatly benefit from analyses that focused on the ***unique details*** of the user's particular situation—rather than being based on preconceived models of how people ought to (and will) follow instructions and procedures. Her detailed analysis of how the help system was unable to help users in many situations, highlighted the inadequacy of basing the design of an interactive system purely on an abstract user model.

Since Suchman's seminal work, a large number of ethnographic studies have examined how work gets done in a range of companies (e.g., fashion, design, multimedia, newspapers) and local government. Other settings have also recently come under scrutiny to see how technologies are used and what people do at home, in public places, in schools, and even cyberspace. Here, the objective has been to understand better the social aspects of each setting and then to come up with implications for the design of future technologies that will support and extend these. For more on the way user studies can inform future technologies, see the interview at the end of this chapter with Abigail Sellen.

4.4 Conceptual frameworks

A number of conceptual frameworks of the "social" have been adapted from other disciplines, like sociology and anthropology. As with the conceptual frameworks derived from cognitive approaches, the aim has been to provide analytic frameworks and concepts that are more amenable to design concerns. Below, we briefly describe two well known approaches, that have quite distinct origins and ways of informing interaction design. These are:

- Language/action framework
- Distributed cognition

The first describes how a model of the way people communicate was used to inform the design of a collaborative technology. The second describes a theory that is used primarily to analyze how people carry out their work, using a variety of technologies.

4.4.1 The language/action framework

The basic premise of the language/action framework is that people act through language (Winograd and Flores, 1986). It was developed to inform the design of systems to help people work more effectively through improving the way they communicate with one another. It is based on various theories of how people use language in their everyday activities, most notably speech act theory.

Speech act theory is concerned with the functions utterances have in conversations (Austin, 1962; Searle, 1969). A common function is a request that is asked indirectly (known as an indirect speech act). For example, when someone says, "It's hot in here" they may really be asking if it would be OK to open the window because they need some fresh air. Speech acts range from formalized statements (e.g., I hereby declare you man and wife) to everyday utterances (e.g., how about dinner?).

There are five categories of speech acts:

- Assertives—commit the speaker to something being the case
- Commissives--committhe speaker to some future action
- Declarations—pronouncesomething has happened
- Directives—get the listener to do something
- Expressives-express a state of affairs, such as apologizingor praising someone

Each utterance can vary in its force. For example, a command to do something has quite a different force from a polite comment about the state of affairs.

The language/action approach was developed further into a framework called conversations for action (CfA). Essentially, this framework describes the sequence of actions that can follow from a speaker making a request of someone else. It depicts a conversation as a kind of "dance" (see Figure 4.13) involving a series of steps that are seen as following the various speech acts. Different dance steps ensue depending on the speech acts followed. The most straightforward kind of dance involves progressing from state 1 through to state 5 of the conversation,

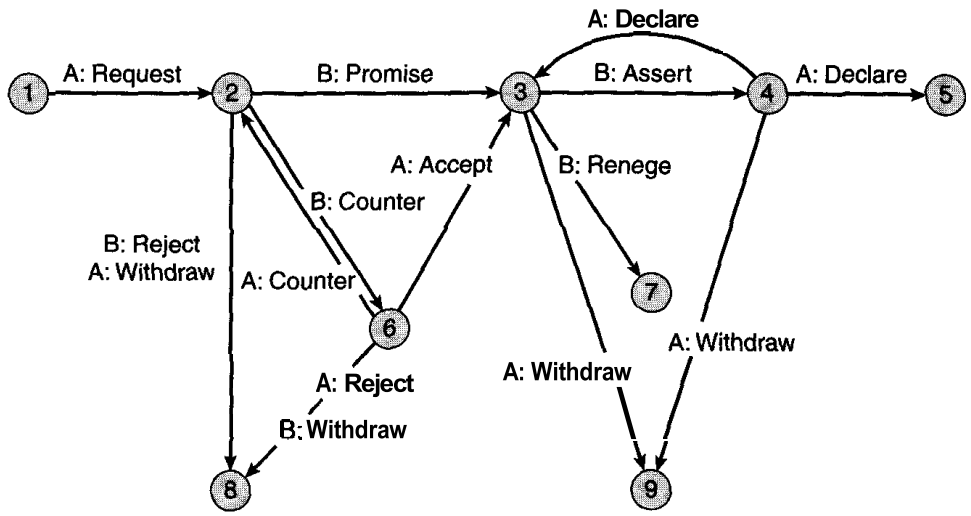


Figure 4.13 Conversation for action (CfA) diagram (from Winograd and Flores, 1986, p. 65).

in a linear order. For example, A (state 1) may request B to do homework (state 2), B may promise to do it after she has watched a TV program (state 3), B may then report back to A that the homework is done (state 4) and A, having looked at it, declares that this is the case (state 5). In reality, conversation dances tend to be more complex. For example, A may look at the homework and see that it is very shoddy and request that B complete it properly. The conversation is thus moved back a step. B may promise to do the homework but may in fact not do it at all, thereby canceling their promise (state 7), or A may say that B doesn't need to do it any more (state 9). B may also suggest an alternative, like cooking dinner (moving to state 6).

The CfA framework was used as the basis of a conceptual model for a commercial software product called the Coordinator. The goal was to develop a system to facilitate communication in a variety of work settings, like sales, finance, general management, and planning. The Coordinator was designed to enable electronic messages to be sent between people in the form of explicit speech acts. When sending someone a request, say "Could you get the report to me", the sender was also required to select the menu option "request." This was placed in the subject header of the message, thereby explicitly specifying the nature of the speech act. Other speech-act options included offer, promise, inform, and question (see Figure 4.14). The system also asked the user to fill in the dates by which the request should be completed. Another user receiving such a message had the option of responding with another labeled speech act. These included:

- acknowledge
- promise
- counter-offer
- decline
- free form

Table A:	Menu items for initiating a new conversation.
Request	Sender wants receiver to do something.
Offer	Sender offers to do something, pending acceptance.
Promise	Sender promises to do something (request is implicit).
What if	Opens a joint exploration of a space of possibilities.
Inform	Sender provides information.
Question	A request for information.
Note	A simple exchange of messages (as in ordinary E-mail).

Figure 4.14 Menu items for initiating a conversation.

Thus, the Coordinator was designed to provide a straightforward conversational structure, allowing users to make clear the status of their work and, likewise, to be clear about the status of others' work in terms of various commitments. To reiterate, a core rationale for developing this system was to try to improve people's ability to communicate more effectively. Earlier research had shown how communication could be improved if participants were able to distinguish among the kinds of commitments people make in conversation and also the time scales for achieving them. These findings suggested to Winograd and Flores that they might achieve their goal by designing a communication system that enabled users to develop a better awareness of the value of using "speech acts." Users would do this by being explicit about their intentions in their email messages to one another.

Normally, the application of a theory backed up with empirical research is regarded as a fairly innocuous and systematic way of informing system design. However, in this instance it opened up a very large can of worms. Much of the research community at the time was incensed by the assumptions made by Winograd and Flores in applying speech act theory to the design of the Coordinator System. Many heated debates ensued, often politically charged. A major concern was the extent to which the system *prescribed* how people should communicate. It was pointed out that asking users to specify *explicitly* the nature of their implicit speech acts was contrary to what they normally do in conversations. Forcing people to communicate in such an artificial way was regarded as highly undesirable. While some people may be very blatant about what they want doing, when they want it done by, and what they are prepared to do, most people tend to use more subtle and indirect forms of communication to advance their collaborations with others. The problem that Winograd and Flores came up against was people's resistance to radically change their way of communicating.

Indeed, many of the people who tried using the Coordinator System in their work organizations either abandoned it or resorted to using only the free-form message facility, which had no explicit demands associated with it. In these con-

texts, the system failed because it was asking too much of the users to change the way they communicated and worked. However, it should be noted that the Coordinator was successful in other kinds of organizations, namely those that are highly structured and need a highly structured system to support them. In particular, the most successful use of the Coordinator and its successors has been in organizations, like large manufacturing divisions of companies, where there is a great need for considerable management of orders and where previous support has been mainly in the form of a hodgepodge of paper forms and inflexible task-specific data processing applications (Winograd, 1994).

4.4.2 Distributed cognition

In the previous chapter we described how traditional approaches to modeling cognition have focussed on what goes on inside one person's head. We also mentioned that there has been considerable dissatisfaction with this approach, as it ignores how people interact with one another and their use of artifacts and external representations in their everyday and working activities. To redress this situation, Ed Hutchins and his colleagues developed the distributed cognition approach as a new paradigm for conceptualizing human work activities (e.g., Hutchins, 1995) (see Figure 4.15).

The distributed cognition approach describes what happens in a cognitive system. Typically, this involves explaining the interactions among people, the artifacts

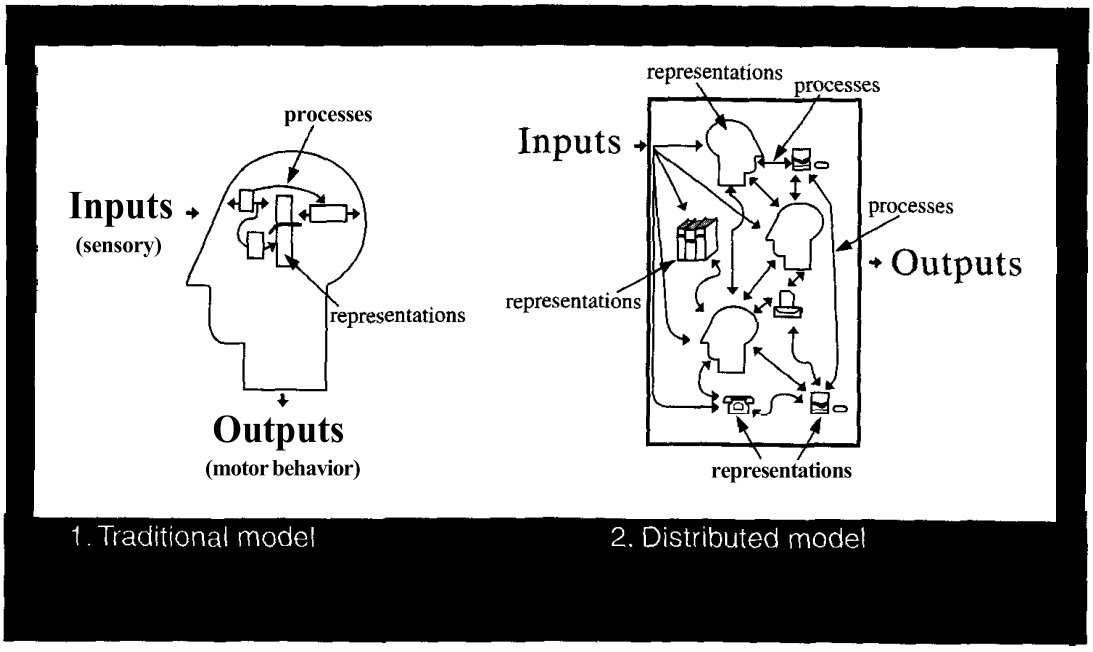
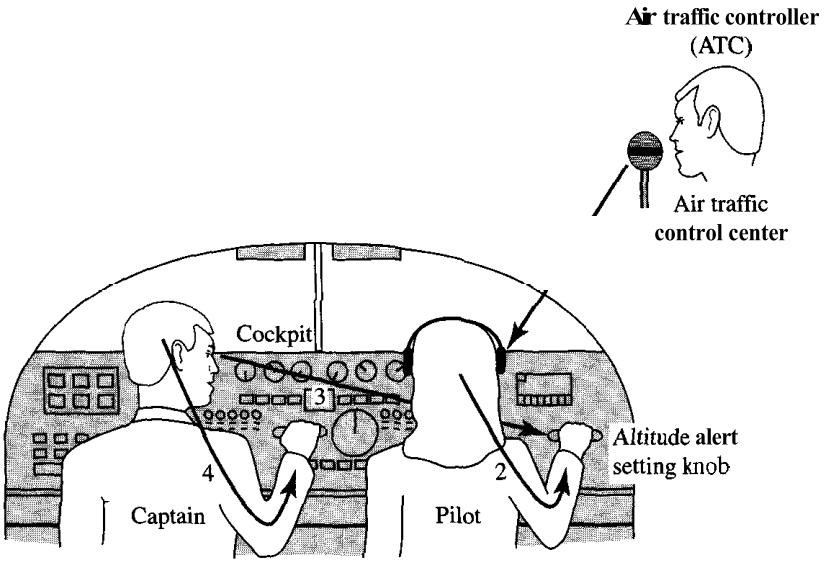


Figure 4.15 Comparison of traditional and distributed cognition approaches.



Propagation of representational states:
1 ATC gives clearance to pilot to fly to higher altitude (verbal)
2 Pilot changes altitude meter (mental and physical)
3 Captain observes pilot (visual)
4 Captain flies to higher altitude (mental and physical)

Figure 4.16 A cognitive system in which information is propagated through different media.

they use, and the environment they are working in. An example of a cognitive system is an airline cockpit, where a top-level goal is to fly the plane. This involves:

- the pilot, co-pilot and air traffic controller ***interacting*** with one another
- the pilot and co-pilot ***interacting*** with the instruments in the cockpit
- the pilot and co-pilot ***interacting*** with the environment in which the plane is flying (e.g., sky, runway).

A primary objective of the distributed cognition approach is to describe these interactions in terms of how information is propagated through different media. By this is meant how information is represented and re-represented as it moves across individuals and through the array of artifacts that are used (e.g., maps, instrument readings, scribbles, spoken word) during activities. These transformations of information are referred to as changes in ***representational state.***

This way of describing and analyzing a cognitive activity contrasts with other cognitive approaches (e.g., the information processing model) in that it focuses not on what is happening inside the heads of ***each*** individual but on what is happening across individuals and artifacts. For example, in the cognitive system of the cockpit, a number of people and artifacts are involved in the activity of "flying to a higher altitude." The air traffic controller initially tells the co-pilot when it is safe to fly to a higher altitude. The co-pilot then alerts the pilot, who is flying the plane, by moving a knob on the instrument panel in front of them, indicating that it is now safe to fly (see Figure 4.16). Hence, the information concerning this activity is transformed

through different media (over the radio, through the co-pilot, and via a change in the position of an instrument).

A distributed cognition analysis typically involves examining:

- the distributed problem solving that takes place (including the way people work together to solve a problem)
- the role of verbal and non-verbal behavior (including what is said, what is implied by glances, winks, etc., and what is not said)
- the various coordinating mechanisms that are used (e.g., rules, procedures)
- the various communicative pathways that take place as a collaborative activity progresses
- how knowledge is shared and accessed

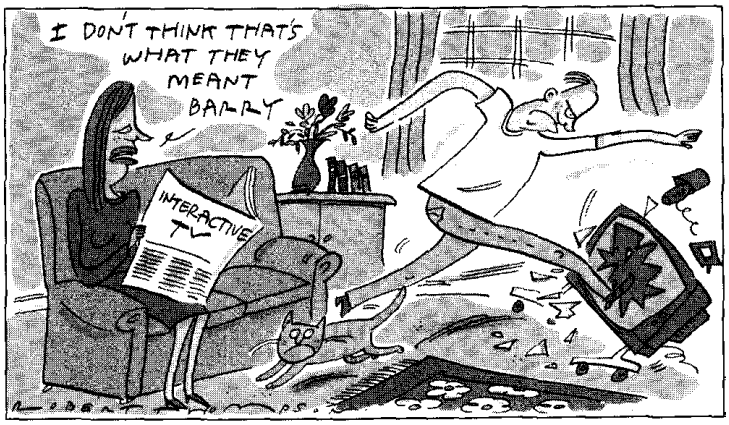
In addition, an important part of a distributed cognition analysis is to identify the problems, breakdowns, and concomitant problem-solving processes that emerge to deal with them. The analysis can be used to predict what would happen to the way information is propagated through a cognitive system, using a different arrangement of technologies and artifacts and what the consequences of this would be for the current work setting. This is especially useful when designing and evaluating new collaborative technologies.

DILEMMA Who Should Have Control?

A core design dilemma facing those involved in developing collaborative technologies is how much control to implement and how much to leave to the users to configure for themselves. Should coordinating mechanisms like rules, procedures, and conventions be designed as part of the system architecture, or should the system be designed to be more free and open, allowing all users to do the same things? For example, when designing a file-sharing system is it more desirable to allow all users free access to all files, or is it preferable to implement some kind of social protocol that allows different user privileges and permissions? Likewise, when designing shared applications such as shared workspaces and collaborative authoring tools, to what extent should system-mediated control mechanisms be implemented that prescribe (and make quite clear) to users how they should share and collaborate? What happens when it is left up to the users to decide upon their own social protocols on how they should coordinate and collaborate with one another? Does it result in anarchy, or do they succeed in creating a shared environment that supports a harmonious way of working?

When working through these design concerns, it is important to consider what happens if too much or too little control is implemented in the collaborative technology. If there is too much "social engineering," then it could result in the users refusing to use it in the way intended. For example, the Coordinator system was found to be unusable in a number of organizations because it required people to radically change their way of communicating in ways considered unacceptable. Similarly, many of the conventions implemented in the POLITeam workspace system (e.g., always typing in the code of a file when using it) were not followed because they required the users to do extra work that was seen by them as being tedious and unnecessary.

Conversely, if not enough consideration is given to the way control is handled, the resulting system could end up being unusable and unacceptable. For example, some of the early shared calendar systems that had a free-for-all policy (anyone could look at anyone else's calendar and arrange a meeting in a free slot) were found to be too much of an invasion of people's privacy.



There are several other well known conceptual frameworks that are used to analyze how people collaborate and communicate, including activity theory, ethnomethodology,situated action and common ground theory.

Assignment

The aim of this design activity is for you to analyze the design of a collaborative virtual environment (CVE) with respect to how it is designed to support collaboration and communication.

Visit an existing CVE (many are freely downloadable) such as V-Chat (vchat.microsoft.com), one of the many Worlds Away environments (www.worlds.net), or the Palace (www.communities.com). Try to work out how they have been designed to take into account the following:

(a) ***General social issues***

What is the purpose of the CVE?

What kinds of conversation mechanisms are supported?

What kinds of coordination mechanisms are provided?

What kinds of social protocols and conventions are used?

What kinds of awareness information is provided?

Does the mode of communication and interaction seem natural or awkward?

(b) ***Specific interaction design issues***

What form of interaction and communication is supported (e.g., textlaudiolvideo)?

What other visualizations are included? What information do they convey?

How do users switch between different modes of interaction (e.g., exploring and chatting)? Is the switch seamless?

Are there any social phenomena that occur specific to the context of the CVE that wouldn't happen in face to face settings (e.g., flaming)?

(c) ***Design issues***

What other features might you include in the CVE to improve communication and collaboration?

Summary

In this chapter we have looked at some core aspects of sociality, namely communication and collaboration. We examined the main social mechanisms that people use in different settings in order to collaborate. A number of collaborative technologies have been designed to support and extend these mechanisms. We looked at representative examples of these, highlighting core interaction design concerns. A particular concern is social acceptability that is critical for the success or failure of technologies intended to be used by groups of people working or communicating together. We also discussed how ethnographic studies and theoretical frameworks can play a valuable role when designing new technologies for work and other settings.

Key points

- Social aspects are the actions and interactions that people engage in at home, work, school, and in public.
- The three main kinds of social mechanism used to coordinate and facilitate social aspects are conversation, coordination, and awareness.
- Talk and the way it is managed is integral to coordinating social activities.
- Many kinds of computer-mediated communication systems have been developed to enable people to communicate with one another when in physically different locations.
- External representations, rules, conventions, verbal and non-verbal communication are all used to coordinate activities among people.
- It is important to take into account the social protocols people use in face to face collaboration when designing collaborative technologies.
- Keeping aware of what others are doing and letting others know what you are doing are important aspects of collaborative working and socializing.
- Ethnographic studies and conceptual frameworks play an important role in understanding the social issues to be taken into account in designing collaborative systems.
- Getting the right level of control between users and system is critical when designing collaborative systems.

Further reading

DIX, A., FINLAY, J., ABOWD, G., AND BEALE, R. (1998) ***Human-Computer Interaction.*** Upper Saddle River, NJ: Prentice Hall. This textbook provides a comprehensive overview of groupware systems and the field of CSCW in Chapters 13 and 14.

ENGESTROM, Y AND MIDDLETON, D. (1996) (eds.) ***Cognition and Communication at Work.*** Cambridge: Cambridge University Press. A good collection of classic ethnographic studies that examine the relationship between different theoretical perspectives and field studies of work practices.

PREECE, J. (2000) ***Online Communities: Designing Usability, Supporting Sociability.*** New York: John Wiley and Sons. This book combines usability and sociability issues to do with designing online communities.

BAECKER, R. M., GRUDIN, J., BUXTON, W. A. S., AND GREENBERG, S. (eds.) (1995) ***Readings in Human-Computer Interaction: Toward the Year 2000,*** (second edition) San Francisco, Ca.: Morgan Kaufmann, 1995.

BAECKER, R. M. (ed.) (1993) ***Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration,*** San Mateo, Ca.: Morgan Kaufmann. These two collections of readings include a number of representative papers from the field of CSCW, ranging from social to system architecture issues.

MUNRO, A.J., HOOK, K. AND BENYON, D. (eds.) (1999) ***Social Navigation of Information Space.*** New York: Springer Verlag. Provides a number of illuminating papers that explore how people navigate information spaces in real and virtual worlds and how people interact with one another in them.

INTERVIEW with Abigail Sellen

Abigail Sellen is a senior researcher at Hewlett Packard Labs in Bristol, UK. Her work involves carrying out user studies to inform the development of future products, including appliances and web-based services. She has a background in cognitive science and human factors engineering, having obtained her doctorate at the University of California, San Diego. Prior to this Abigail worked at Xerox Research Labs in Cambridge, UK, and Apple Computer Inc. She has also worked as an academic researcher at the Computer Systems Research Institute at the University of Toronto, Canada and the Applied Psychology Unit in Cambridge, UK. She has written widely on the social and cognitive aspects of paper use, video conferencing, input devices, human memory, and human error, all with an eye to the design of new technologies.

YR: Could you tell me what you do at Hewlett Packard Research Labs?

AS: Sure, I've been at HP Labs for a number of years now as a member of its User Studies and Design Group. This is a smallish group consisting of five social scientists and three designers. Our work can best be described as doing three things: we do projects that are group-led around particular themes, like for example, how people use digital music or how people capture documents using scanning technology. We do consulting work for development teams at HP, and thirdly, we do a little bit of our own individual work, like writing papers and books, and giving talks.

YR: Right. Could you tell me about user studies, what they are and why you consider them important?

AS: OK. User studies essentially involve looking at how people behave either in their natural habitats or in the laboratory, both with old technologies and with new ones. I think there are many different questions that these kinds of studies can help you answer. Let me name a few. One question is: who is going to be the potential user for a particular device or service that you are thinking of developing? A second question—which I think is key—is, what is the potential value of a particular product for a user? Once we know this, we can then ask, for a particular situation or task, what features do we want to deliver and how best should we deliver those features? This includes, for example, what would the interface look like? Finally, I think user studies are important to understand how users' lives may change and how they will be affected by introducing a new technology. This has to take into account the social, physical, and technological context into which it will be introduced.

YR: So it sounds like you have a set of general questions you have in mind when you do a user study. Could you now describe how you would do a user study and what kinds of things you would be looking for?

AS: Well, I think there are two different classes of user studies and both are quite different in the ways you go about them. There are evaluation studies, where we take a concept, a prototype or even a developed technology and look at how it is used and then try to modify or improve it based on what we find. The second class of user studies is more about discovering what people's unmet needs may be. This means trying to develop new concepts and ideas for things that people may never have thought of before. This is difficult because you can't necessarily just ask people what they would like or what they would use. Instead, you have to make inferences from studying people in different situations and try to understand from this what they might need or value.

YR: In the book we mention the importance of taking into account social aspects, such as awareness of others, how people communicate with each other and so on. Do you think these issues are important when you are doing these two kinds of user studies?

AS: Well, yes, and in particular I think social aspects really are playing to that second class of user study I mentioned where you are trying to discover what people's unmet needs or requirements may be. Here you are trying to get rich descriptions about what people do in the context of their everyday lives—whether this is in their working lives, their home lives, or lives on the move. I'd say getting the social aspects understood is often very important in trying to understand what value new products and services might

bring to people's day-to-day activities, and also how they would fit into those existing activities.

YR: And what about cognitive aspects, such as how people carry out their tasks, what they remember, what they are bad at remembering? Is that also important to look into when you are doing these kinds of studies?

AS: Yes, if you think about evaluation studies, then cognitive aspects are extremely important. Looking at cognitive aspects can help you understand the nature of the user interaction, in particular what processes are going on in their heads. This includes issues like learning how users perceive a device and how they form a mental model of how something works. Cognitive issues are especially important to consider when we want to contrast one device with another or think about new and better ways in which we might design an interface.

YR: I wonder if you could describe to me briefly one of your recent studies where you have looked at cognitive and social aspects.

AS: How about a recent study we did to do with building devices for reading digital documents? When we first set out on this study, before we could begin to think about how to build such devices, we had to begin by asking, "What do we mean by reading?" It turned out there was not a lot written about the different ways people read in their day-to-day lives. So the first thing we did was a very broad study looking at how people read in work situations. The technique we used here was a combination of asking people to fill out a diary about their reading activities during the course of a day and interviewing them at the end of each day. The interviews were based around what was written in the diaries, which turned out to be a good way of unpacking more details about what people had been doing.

That initial study allowed us to categorize all the different ways people were reading. What we found out is that actually you can't talk about reading in a generic sense but that it falls into at least 10 different categories. For example, sometimes people skim read, sometimes they read for the purpose of writing something, and sometimes they read very reflectively and deeply, marking up their documents as they go. What quickly emerged from this first study was that if you're designing a device for reading it might look very different depending on the kind of reading the users are doing. So, for example, if they're reading by themselves, the screen size and viewing angle may not be as important as if they're reading with others. If they're skim reading, the ability to quickly flick through pages is important. And if they're reading and writing, then this points to the need for a pen-based interface. All of these issues become important design considerations.

This study then led to the development of some design concepts and ideas for new kinds of reading devices. At this stage we involved designers to develop different "props" to get feedback and reactions from potential users. A prop could be anything from a quick sketch to an animation to a styrofoam 3D mockup. Once you have this initial design work, you can then begin to develop working prototypes and test them with realistic tasks in both laboratory and natural settings. Some of this work we have already completed, but the project has had an impact on several different research and development efforts.

YR: Would you say that user studies are going to become an increasingly important part of the interaction design process, especially as new technologies like ubiquitous computing and handheld devices come into being—and where no one really knows what applications to develop?

AS: Yes. I think the main contribution of user studies, say, 15 years ago was in the area of evaluation and usability testing. I think that role is changing now in that user studies researchers are not only those who evaluate devices and interfaces but also those who develop new concepts. Also, another important development is a change in the way the research is carried out. More and more I am finding that teams are drawing together people from other disciplines, such as sociologists, marketing people, designers, and people from business and technology development.

YR: So they are essentially working as a multidisciplinary team. Finally, what is it like to work in a large organization like HP, with so many different departments?

AS: One thing about working for a large organization is that you get a lot of variety in what you can do. You can pick and choose to some extent and, depending on the organization, don't have to be tied to a particular product. If, on the other hand, you work

for a smaller organization such as a start-up company, inevitably there is lots of pressure to get things out the door quickly. Things are often very focused. Whether large or small, however, I think one of the hardest things I have found in working for corporate research is learning to work with the development teams. They put huge pressures on you because they have huge pressures on them. You really have to work at effectively incorporating user studies findings into the development process. This can be incredibly challenging, but it's also satisfying to have an impact on real products.