

Videoconferencing across cultures – a conceptual framework for floor control issues

SCHAHRAM DUSTDAR

University of Art at Linz, A-4010 Linz, Austria

GERT JAN HOFSTEDE

Wageningen University, Wageningen, The Netherlands

This paper discusses critical issues in cross-cultural communication and collaboration using desktop videoconferencing tools. Our first objective is to propose a conceptual framework for predicting which issues will be important for communication in cross-cultural desktop videoconferencing. Using this framework, we suggest necessary functionality regarding floor control policies in videoconferences. We then suggest some actual software mechanisms that could implement these policies to videoconferencing software builders. Finally, we present some conduct guidelines for people who will be involved as participants in cross-cultural desktop videoconferences.

Introduction

When people communicate using computer-mediated information technologies (ITs), such as videoconferencing systems, the effectiveness and style of their communication is affected. In videoconferencing, there are fewer clues between the participants about who is who and about who can take the floor than in face-to-face meetings. This can be expected to lead to a number of changes in the participants' behaviour. For instance, a participant wanting to impress authority on the others does not have the means to stand up, pace to and fro or gesture menacingly in a meeting where only a head is seen and where the frame rate does not allow rapid movements. A participant who is eager to speak up but needs a superior to give him the floor cannot sit and fidget to signal this non-verbally. A participant wishing to give support or criticism cannot hum or gesture to indicate this. Perhaps head shaking can replace humming or gesturing in this last example. More generally, new ways have to be found to express actions in videoconferences. The sort of issues we are specifically interested in can be summarized as 'floor control'. Floor control has to do with who is allowed to speak in a meeting and what mechanisms the participants use to give and take the floor. The above can be expected to be contingent upon national culture. Verbal and non-verbal behaviour in interpersonal communication varies quite widely depending upon a person's cultural background (Gudykunst and Ting-Toomey, 1988). This is obviously the case when first-world people meet non-first-world people, but it also holds among countries that might at first sight not appear to be so different. For

instance, consider the following quote from a Dutchman who had worked in the USA for a long time: 'In the Dutch situation, meetings were places where problems were discussed and common solutions were sought; they served for making decisions. In the US situation as he had known it, meetings were opportunities for participants to assert themselves: to show how good they were. Decisions were made by individuals elsewhere' (Hofstede, 1991, p. 95). If differences in the role of meetings exist across countries, they will also exist if these meetings are videoconferences. Generally, we expect computer-mediated communication to change the relationship between the verbal and non-verbal aspects of communication relative to a face-to-face situation. This relationship is very intricate in face-to-face conversation. Posture, gestures, distance, eye movements and intonation all modify what is being said and in fact overrule the meaning of the verbal part. 'I hate you' can mean 'I love you' depending on the non-verbal setting. In e-mail communication, no non-verbal aspect is present. This has been shown in the literature to lead potentially to misunderstanding and frustration (Markus, 1994). Smileys, for example, are an attempt to overcome this limitation. The main thrust of an e-mail message, however, is left to the verbal content. This puts high demands on the author's and the reader's literacy. It can be expected that people who already know each other will be much better able to understand each other's e-mail messages than people who do not. For instance, for strangers wishing to communicate, the urgency of a request via e-mail, this can be very awkward. In our personal experience we have noted that adopting formal versus colloquial syntax can be used to

indicate hostile versus conciliatory feelings. Incidentally, the lack of prominence of non-verbal communication in e-mails can sometimes be an advantage too: some people may not like each other's presence but be able to handle necessary communication effectively through e-mail. In addition, people who do like each other may take more time doing things when in face-to-face conversation than when exchanging e-mail. To their employers they may 'waste time' simply because they enjoy being together. So, to conclude, e-mail contact does not favour non-verbal communication. The non-verbal aspect of an e-mail message has to be either left out or incorporated painstakingly into the verbal part. Videoconferences are in between face-to-face and e-mail contact. To begin with, this is true in a trivial sense. E-mail conversation is essentially one to one and so is most of what is called videoconferencing today. The analogy to a real meeting would be a multipoint (many-person) videoconference, as shown in Figure 1 and used on the Internet's Mbone overlay network. Videoconferences are also in between face-to-face and e-mail contact as far as the integration of verbal and non-verbal clues is included. Just as in the e-mail context, people will wish to find ways to deal with the non-verbal aspects of communication. For the participants, it is important not to get frustrated or spoil relationships. For those who pay for the videoconference, it is important to get value for their money, i.e. an effective meeting. In this paper we focus on the use of desktop videoconferencing systems. Desktop videoconferencing systems differ from room-based systems in the way people communicate and collaborate, since the mediating technology is the workstation they use for daily work. In particular, we made studies using Mbone – the Multicast Backbone (Casner, 1993) – videoconferencing tools on the Internet. Our first objective is to propose a conceptual framework for predicting which issues will be important for communication in cross-cultural desktop videoconferencing. This framework can serve as a source of inspiration to researchers of cross-cultural issues in computer-mediated communication. Using this framework, we suggest necessary functionality regarding floor control policies in videoconferences. We then suggest some actual software mechanisms that could implement these policies to videoconferencing software builders. Finally, we present some conduct guidelines for people who will be involved as participants in cross-cultural desktop videoconferences.

Prior research on videoconferencing, national culture and floor control

This article's title contains three key notions: videoconferencing, culture and floor control. All have been



Figure 1 Multipoint videoconference on the Internet

researched before in the information systems (IS) world and we shall present some relevant research on all three subjects in this section.

Videoconferencing

Videoconferencing has its roots in the development of computer-based, shared work spaces. The notion of group collaboration using computers was introduced by Bush (1945) using Memex, a group hypertext system. Engelbart (1968) implemented NLS/ AUGMENT, one of the first systems using computers for synchronous and asynchronous group interaction. Johansen (1988) predicted the growth of interest in groupware technologies. Since then, there has been considerable development in collaborative systems. There is literature on tools for multiuser text editing (Ellis *et al.*, 1991), annotation systems (Cavalier *et al.*, 1991), sketching tools (Jacobson and McCanne, 1993; Ishii *et al.*, 1995) and group support systems. Mantei (1988) designed computer-equipped meeting rooms for small groups using the computers for problem solving in a face-to-face context. Electronic meeting rooms can be used for inter- or intraorganizational meetings or classroom settings. The systems discussed so far are 'room based', i.e. the participants have to assemble in one place. Another line of development was constituted by systems that allowed remote contact. Early electronic meeting systems lacked the ability to integrate multiple media types such as audio, video and textual information in one multimedia system. The merging of workstation technology and real-time computer conferencing has had a significant impact on CSCW (Computer Supported Cooperative Work) and group decision making and led to the term 'desktop conferencing'. Research on early video-

conferencing systems, such as those developed at A.T. & T. Bell Laboratories (Ahuja *et al.*, 1990), Bellcore (Root, 1988) or NEC (Watabe *et al.*, 1990), had the provision of the facilities found at face-to-face meetings with remote groups as their aim. Recent developments in multimedia systems and networking technology show that using desktop videoconferencing for collaborative work on wide area networks such as the Internet is possible (Macedonia and Brutzman, 1994). It is generally accepted that computer-supported decision making and communication results in many changes in communication patterns (Heath and Luff, 1991), greater task orientation (Niemic, 1984) and shorter meetings (Heath and Luff, 1991). Regarding the video component, Ishii *et al.* (1995) pointed out the importance of gaze awareness, the ability to monitor the direction of someone's gaze and, thus, the focus of attention. Similar results were found by Heath and Luff (1991). However, researchers have often discussed the lack of video of sufficient quality to support interpersonal communication (Egido, 1990).

National culture

The title's second keyword is culture. Now that communication technology is spanning the whole globe, national culture has surfaced as a variable in IS research of many kinds. Following Hofstede (1991), we define 'national culture' as 'the collective programming of the mind which distinguishes the members of one country from another'. That is, national culture is taught from early infancy. It does not encompass personality, nor, on the other hand, human nature. One of the theories on national culture that is widely adopted by IS researchers is Hofstede's (1980, 1991) framework of five dimensions of national culture. Hofstede's (1980, 1991) main findings were that the attitudes of people towards their jobs and employers can be classified along a number of 'dimensions of culture'. These were originally empirically derived in a comparison among IBM employees from many nationalities. They have been found to be quite useful for describing intercultural communication (Gudykunst and Ting-Toomey, 1988; Pedersen and Ivey, 1993). Erez and Earley (1993) indicated why Hofstede's (1980, 1991) work has become quite popular for experimental work. In their words 'from a purely management perspective, Hofstede's work has been of great benefit to researchers because it is *approachable*; it (the model) is sufficiently clear and parsimonious to lend itself to empirical tests' (Erez and Earley, 1993, p. x).

Hofstede's (1980, 1991) dimensions constitute the framework for our research model. They are (1) power distance, (2) individualism versus collectivism, (3) masculinity versus femininity, also termed competi-

tiveness versus cooperativeness and (4) uncertainty avoidance. A fifth dimension was found with South-East Asian cultures: (5) long-term versus short-term orientation. The first three dimensions represent three fundamental relationships between people: vertical, horizontal and gender related. The latter two can be seen as the Western and Eastern versions of a culture's orientation towards time. The fifth dimension, long-term orientation, did not figure in Hofstede's (1980, 1991) original data, because the original questionnaires, having been drawn up by a team without South-East Asian members, did not contain questions addressing this dimension. It was discovered later using a questionnaire designed by South-East Asians. The particulars can be read in Hofstede (1991). All five dimensions are orthogonal, that is a country's score on power distance has nothing to do with its score on individualism, masculinity, uncertainty avoidance or long-time orientation. To see how this framework can help explain communication during meetings, let us return to the Dutch expatriate quoted in the Introduction. The Netherlands and the USA are both countries with low power distance, high individualism and low uncertainty avoidance. However, the USA is considerably more masculinist in its national culture than The Netherlands. This difference can account for the Dutchman's comment.

Floor control

The third keyword in the paper's title is floor control. A meeting is a group setting in which only one participant can speak or 'have the floor' at any point in time. Thus, informally, 'the floor' means 'the temporary monopoly for distributing signals to the other participants'. A more precise definition of the concept of floor in videoconferencing can be found in Dommel and Garcia-Luna-Aceves (1997): 'Floors are temporary permissions granted dynamically to collaborating users in order to mitigate race conditions and guarantee mutually exclusive resource usage' (p. 23). This second definition shows that more than one floor can exist at any moment in time in an electronic meeting. One person could have the audio floor, while another one has a whiteboard floor or a minutes floor. However, there will almost always be a main speaker at any point in time and the loose definition, which captures that concept, will suffice for the present paper. How the floor is taken and lost or granted is crucial for a meeting's process and outcome. A meeting needs a social protocol, part of which is a floor control mechanism, in order to proceed in a manner that is satisfactory to the participants. For instance, a chairman may be appointed whose task it is to keep a first-in, first-out queue of people who have indicated a wish to speak by raising their hand.

Alternatively, a participant with high prestige may grant the word and the others will just wait until they are addressed by this high-status person. Or everybody may just try and get a word in, using their voice and body posture to indicate the urgency of their contribution. It is not hard to see how differences in culture can affect the floor control mechanism that is chosen. For example, participants from countries with low power distance will expect that anybody can take the floor whenever they want, unless other arrangements – such as appointing a floor-granting chairperson – have been explicitly made. The floor control mechanisms are engrained in national culture. Distributed collaborative systems such as videoconferencing systems must indicate something about the social world they represent, in particular questions such as who is on the system, what are the others currently doing and in which context they are. These social activity indicators (Ackerman and Starr, 1996) are becoming increasingly important the more people from various cultural and professional backgrounds have to collaborate. Furthermore, the lack of social activity indicators in collaborative systems slows down the process of establishing a critical mass of users. When dealing with floor control issues two areas have to be discussed: floor control policies and floor control mechanisms. Floor control policies are employed within shared work spaces to control the form and type of access. Policies describe how conference participants request the floor and how the floor is assigned and released. In the context of desktop videoconferencing systems, the simplest form of floor control would be if only one conference participant had the floor at any given time and the floor is handed off whenever requested by other participants. To obtain the floor the conference participant may either be requested to take an explicit action such as pushing one button in the user interface and being queued or to notify a conference moderator or chairperson. Floor control policies were discussed in Craighill *et al.* (1993). Floor control mechanisms are low-level means used to implement floor control policies (Reinhard *et al.*, 1994). A discussion of consistency mechanisms can be found in Ellis *et al.* (1991). Among the technical issues which need further investigation is research related to session control of videoconferencing, which manifests itself in floor control mechanisms. A loosely controlled session, as initiated by using the SD (session directory), has little to less interaction between participants. In loosely controlled sessions there is no mechanism for negotiating on parameters such as media type, media encoding, encryption keys and membership issues. Basically, we differentiate between point-to-point videoconferences and multipoint videoconferences. In point-to-point desktop videoconferences the need for floor control

policies and mechanisms is very unlikely, but as the size of desktop videoconferences and user participation grows the more issues of floor control become evident and important.

Research model and hypotheses

As stated, Hofstede's (1991) five dimensions of national culture are the basis for our research model. Two of the five usually breed most problems in organizations if people who cooperate differ along them. These are power distance and uncertainty avoidance (Hofstede, 1991). To our knowledge, specific research on the influence of national culture on communication during meetings has not been carried out, although much anecdotic evidence exists. However, Hofstede's (1991) dimensions are being used to train intercultural consultants, indicating their relevance in this context. Pedersen and Ivey (1993) introduced the concept of a 'synthetic culture', a script for an extreme manifestation of any one of the five dimensions. For instance, a synthetic culture can be 'high power distance' or 'long-term orientation'. Experiments in which meeting participants enacted extreme manifestations of one of the dimensions showed a marked effect of enacted national culture on both the process and the outcome of meetings (Hofstede, 1996). Although in their book *Culture and Interpersonal Communication* Gudykunst and Ting-Toomey (1988) gave many clues, we have so far not found a comprehensive treatment of floor control issues during meetings in a cultural perspective. What is very clear from both literature and practice is that culture clash can lead to severe communication breakdowns in any cross-culture conversation. Our interest is in seeing whether videoconferencing systems and procedures can be adapted to the particular cross-cultural setting they are used in so that the occurrence of culture-clash-related misunderstandings is minimized. Our research framework consists of five dimensions (power distance, individualism/collectivism, masculinity/femininity, uncertainty avoidance, short/ long-term orientation), so theoretically there are ten possible monodimensional synthetic cultures. In a mixed-culture meeting, the situation will be confounded, because the needs and expectations of the participants will vary. This will place even higher demands on the technical environment and on the meeting's organization.

Experiences with videoconferencing on the Internet

In the following section we summarize our findings. We discuss some issues which are crucial in point-to-

point and multipoint desktop videoconferences and give some suggestions for dealing with the problems we encountered. The videoconferencing tools on the Internet/MBone can be classified as (1) video tools, (2) audio tools and (3) shared whiteboards. These tools are not integrated, i.e. it is not possible to use an audio tool to transmit video either from point-to-point or to multipoint. The same is true for video tools and shared whiteboards. One of the first things one says after establishing a videoconference link is the question 'Can you hear me and/or can you see me?'. This circumstance makes it impracticable to establish a videoconference without using a telephone as a backup medium. We shall first give an account of a number of experiences in an informal way, ordered by type of communication medium. This is done in order to capture as much variety as we can. Then we shall try to interpret these in the light of the five culture dimensions and see whether we can find evidence that relates to our research framework. In a point-to-point telephone conversation, the listener will frequently hum or say 'yes' to indicate he or she is listening and hears what is being said. In a multipoint telephone conference, this is not so easy. The listeners tend to be silent. This may leave the speaker wondering how what they have said has been received. This is particularly disturbing for 'political' persons, who rely much on atmosphere and eye contact to assess how their messages are received. The next thing one encounters is the problem of concurrent speaking. Generally speaking, the network bandwidth is limited in most cases. Since audio and video streams are sent over the Internet traffic is a problem. Therefore audio and video data packets get lost. In the case of an audio packet loss one will feel very disturbed and the conversation will decrease in quality. It was mentioned above that collaborative systems lack social activity indicators. Due to the fact that the videoconferencing tools are not integrated it is difficult for conference participants to check which of the participants has, for example, an audio or whiteboard window only. Using the Mbone tool 'vat' on the Internet one can see the active audio conference participants – represented by their e-mail address – but in this window there is no link to the video tool. However, the audio tool shows which conference participant is speaking. There is a 'meter' below the loudspeaker symbol which moves according to the speakers' volume. Another useful feature is the 'mute' button. There are cases in which a conference participant needs some 'private' time or simply does not want to transmit audio from his or her site. As stated above, the network bandwidth is limited in most cases. Over the Internet, data packets get lost. In cases of video packet loss human perception is not as sensitive as it is with audio packet loss. If video packets get

lost the motion of the other videoconference participant(s) is 'jerky'. Nuances in facial expression will be lost and it will not be possible to infer how the others respond to what is being said. On the other hand, one can stare at a particular participant unnoticed and still infer things for example from fidgeting. A problem which occurs in multipoint videoconferences, as shown in Figure 1, is the missing link between audio and video tools. In Figure 1 there are 11 active conference participants. One can hardly tell who is currently speaking, even when one enlarges the thumbnail video of the anticipated conference participant manually by clicking on it. As can be seen in Figure 1 it is very difficult to have all video windows open, even if the user has a 21 inch monitor. One has to think of also having the necessary audio and shared whiteboard windows open concurrently. A feature which one can find in face-to-face meetings but which is missing here is the possibility of directing audio or video streams to only one or to a subgroup of the conference participants ('side chatting'). In a face-to-face meeting you can turn to your neighbour and have a short chat. Using desktop multipoint videoconferencing this notion gets lost. Figure 2 also illustrates the problem of 'conference participants' context'. As one can see from the figure, the two large video windows differ a great deal. The window on top shows a large video image of a conference participants' head and a few clues about his office context. If you communicate with this participant you do not really know if he or she is alone in their office or if maybe two other people are sitting in a position where the camera cannot see them. Figure 2 illustrates the problem. The window below shows (at least) two conference participants sharing one videoconference workstation. Communication and collaboration using one workstation with two participants can be quite disturbing. In this area we suggest that desktop videoconferencing participants have at least two video cameras installed so that the conference partners have more clues on the organizational and personal context.



Figure 2 Context of videoconference participant

Finally, the shared whiteboard heavily depends on cultural issues. The Internet/MBone whiteboard application 'wb' has some drawing and writing tools and some shared document space. wb shows a shared whiteboard participant information window which indicates which conference participant is currently writing or drawing. In some cases this might be a desired function but there are cases in which it makes a lot of sense to 'anonymize' the results by not showing which participant wrote on the whiteboard. If a contribution is anonymous, those who read it will tend to take it at face value. If it is not, it will be interpreted politically, on the basis of the reader's relationship to the author. Another issue that surfaces in whiteboarding or in simple e-mail is use of language. For instance, whether a non-native English speaker has been taught British or American English makes a great deal of difference in the style they have learnt to adopt. Misunderstandings can arise because participants unknowingly offend the receiver. For example, USA inhabitants frequently do not bother to put any header (such as 'Dear X' or even 'Hi X!') above their e-mail messages. Receivers in many countries will not be favourably struck by what they perceive as a lack of respect or manners.

Implications for floor control

Now, we shall try to draw some implications from these findings. For the sake of clarity, these will be sorted by dimension of culture, knowing this is rather artificial. After all, every person is a 'living mix' of all five dimensions and also each person has a number of personal characteristics and a personal history. The mix of all of these can lead to different behaviour than can be explained by only one dimension of culture. Low scores on power distance will lead to equal distribution of speaking time and spontaneous taking the floor by all. Given the audio quality that is reached with current technology, this will lead to packet loss. Video-conference participants from low power distance cultures will have to get used to waiting for a silence before they talk. High scores will lead to unequal distribution of speaking time, with subordinates only speaking when they are explicitly granted the floor by a superior. If the status of participants is not clear, communication will be tentative until clarity is reached. If the participants to a meeting are of similar power distance cultures, there will not be any problems, as long as they know each other's relative status. Each participant will then know their place. However, if they differ along the power distance axis, those low on power distance will tend to take the word uninvited and assume that the others will do the same. This can lead to frustration. To tackle this issue, one

could think of a feature such as a 'speaking time meter' which shows to all conference participants the amount of time spoken to the group of participants. Currently, no such thing exists, but it is technically easy to implement in a video or audio conference. A session chair could use such a device to ensure participation from all. In fact it is the sort of innovation that electronic media can bring. Individualists will all freely express their opinions, with a focus on the task at hand. They may be of diametrically opposed opinions without having any personal antipathy. They may invest considerable energy into forming and negotiating *ad hoc* coalitions. In order to do so, they will occasionally engage in side chatting in small groups. As stated above, current videoconferencing tools do not allow for this. More-over, individualists who are low on power distance will frequently bump into each other's statements. This will be awkward in electronic meetings. Collectivists of a single in-group will all express similar views, with a view to preserving group harmony. To them, expressing a different view equals an attack on the group. Collectivists from different in-groups will likely either not express any opinion at all for fear of saying the wrong thing or they might fiercely attack out-group participants who express deviant views. Achievement-oriented (masculinist) participants will not hesitate to engage in open, up-front conflict when there is a difference of opinion. Cooperation-oriented (feminist) participants will avoid open conflict, but rather they will try to resolve differences of opinion by compromising and will mediate. Speakers high on uncertainty avoidance will wish to set formal rules for the meeting, including floor control regulation. They will be emotional in their statements, possibly exaggerating. Those low on uncertainty avoidance will be calm, possibly 'dull', will like to explore divergent ideas and will be tolerant of individuals who deviate from any rules. Differences in uncertainty avoidance are known to be the cause of much aggravation in organized life. In a setting of mixed uncertainty avoidance, it is probably a good idea to lay out the rules of the game very clearly right at the outset of a meeting, so that there will at least not be much ambiguity about them. As to long-term orientation, speakers who are short-term oriented will be much concerned with issues of 'face': not insulting somebody, reciprocating compliments and so on. Long-term oriented speakers will not offend others either, unless their own objectives for the meeting are at stake. Table 1 summarizes the hypotheses ventured above.

Discussion

So far, we have been concerned mainly with the perspective of a participant who has the floor or would wish

Table 1 Characteristics of meetings

Culture	Implications for floor control	
	Low dimension of culture	High dimension of culture
Power distance	Spontaneously taking the floor	Unequal distribution of time, with superior controlling the floor
Individualism	Non-verbal signals important Deviant views lead to strong feelings and often to withdrawal	Free opinion sharing Quality of signals not so important. Interruptions will occur or could be blocked. Side chatting occurs
Masculinity	Compromise and mediation	Open up-front conflict
Uncertainty avoidance	Calm meetings and few policies	Participants want formal rules for the meeting Strong non-verbal support for utterances
Long-term orientation	Much concerned with saving face	Achieving objectives more critical, even if it means being impolite

to take it. However, in an actual meeting, most participants are just listeners most of the time. It is only after having listened that they decide whether to take the floor and what to say or write. And it is particularly for a listener that an electronic meeting differs from a face-to-face meeting. For instance, consider a meeting where the participants are unknown to each other. Somebody issues a statement. A listener will not only listen to the content of what is being said. Depending on their culture, listeners will also wish to know something about the speaker in order to be able to interpret what is being said. Our hypotheses can be summarized as follows. People high on individualism and low on the other dimensions will readily engage in communication on the basis of the utterances only, without knowing who is talking. In fact this coincides with the success of the Internet, where 'nobody knows you're a dog', in Anglo-Saxon countries. The hypotheses also propose that listeners will feel uncomfortable in settings where the clues they culturally wish to have about a speaker are not provided. There are two ways to respond to this uncertainty. Either one could adapt the communication medium to include the necessary clues (the 'social identity clues mode', or 'clues mode' for short) or one could deliberately leave them out, forcing communication on

the basis of ideas only (the 'mind mode'). In the context of videoconferencing, the clues mode would necessitate giving explicit information about status and group membership (through institutional and educational background or curriculum vitae information). Gender can be seen, as well as some limited information about social identity, through hair style, clothing, pronunciation, vocabulary and verbosity. Mind mode would lead to the use of whiteboarding techniques, so that contributions could be anonymized. In an actual electronic meeting, actual groupware tools will be used. What floor control policies and mechanisms should these possess, according to our findings and hypotheses? Table 2 provides an overview of the proposed floor control policies and mechanisms.

Groupware should, if possible, accommodate all possible mixes of culture. Therefore, all the features in Table 2 are desirable. It is up to the participants to choose how to use them. Let us imagine a multicultural videoconference with all participants from synthetic cultures who do not know each other well. There are two chairpersons: one is the regular meeting chair and the other is a technical chair who knows about the software. The meeting could start with a little 'looking around' round in which all participants tell the others where they are and show the others if their camera can be moved. This will allow the participants to do whatever their synthetic culture has taught them to do: be polite, come on strong or obtain information about the others' status or group membership. Then the technical chair could explain how the meeting will proceed and give examples of the features offered by the soft-

Table 2 Implications of culture for floor control policies

Culture	Implications for videoconferencing sessions	
	Low dimension of culture	High dimension of culture
Power distance	Wish to interrupt indicator for listener that floor holder can choose to acknowledge or disregard	Show people's responsibilities Privilege of floor granting
Individualism	Enhance video quality and provide background for participants	Allows side chatting
Masculinity		Possibilities for non-verbal displays of strength are limited
Uncertainty avoidance		Have a formal protocol
Long-term orientation		

ware. Prior consultation between the general chair and the technical chair would have set the guidelines here. Depending on the meeting's aim, the general chair could distribute the floor or 'I wish to interrupt' buttons could be used, anonymity of contributions could be an option, speaking time could be regulated and side chatting could be disabled. In addition, the technical chair could give some advice about how to enhance communication, such as getting proper lighting for one's face, speaking clearly and not moving about too vehemently when speaking. Thus prepared, the participants can start the meeting. It will be rather more orchestrated than a face-to-face meeting. Among other things, the time slot for the meeting is likely to be quite strict, so that the meeting's end-time is predetermined. The general chair will take this into account and do whatever is necessary to wind up the meeting in time.

Future research

Together with an Internet connection, the MBone desktop videoconferencing tools provide the user and an organization with the possibility of communicating and collaborating on a global scale. However, the tools need to be integrated into organizational information systems, such as work flow and groupware systems, word processing, project management software and spreadsheet applications and, most important, they need to support 'social protocols' and be culturally aware. Among the organizational issues, questions of floor control policies and the necessary support for cross-cultural communication and collaboration have to be resolved. The authors are convinced that this research area needs interdisciplinary efforts since multimedia itself is an interdisciplinary field. Our future research will consist of both controlled experiments and observations of non-controlled videoconferences. In the experiments, synthetic cultures will be used. In the observations, the participants will be asked to answer a questionnaire about their cultural values. In the experiments, both the clues mode and mind mode can be used, as far as technology permits. In the observations, the mode of working will likely be out of the researchers' control. To observe what actually happens during a conference, a variant of speech act theory, as it is being used by discourse analysts (Te Molder, 1995), is envisaged. Speech act theory links verbal utterances to the purposes they serve in a conversation, such as 'make a proposal', 'give support to' or 'object to'. As Ishii *et al.* (1995) stated, we are interacting not with computers, but through computers. This should gradually lead to a new and better understanding of cultural issues as dependent variables in desktop videoconferencing. Social protocols that surround videoconferences will have to emerge.

References

- Ackerman, M.S. and Starr, B. (1996) Social activity indicators for groupware. *IEEE Computer*, **29**(6), 37–42.
- Ahuja, S.R., Ensor, J.R. and Lucco, S.E. (1990) A comparison of application-sharing mechanisms in real-time desktop conferencing systems, in *Proceedings of the COIS*, pp. 238–48.
- Bush, V. (1945) As we may think. *Atlantic Monthly*, **176**(1), 101–8.
- Casner, S. (1993) *Frequently Asked Questions on the Multicast Backbone* (Information Sciences Institute, University of Southern California).
- Cavalier, T., Chandhok, R., Morris, J., Kaufer, D. and Neiwirth, C. (1991) A visual design for collaborative work: columns for commenting and annotation, in *Proceedings of the 24th Annual Hawaii International Conference in System Sciences*, Hawaii, pp. 729–38.
- Craighill, E., Lang, R., Skinner, K. and Fong, M. (1993) CECED: a system for informal multimedia collaboration *Proceedings of the Conference on Multimedia '93*, p. 437.
- Dommel, H.-P. and Garcia-Luna-Aceves, J.J. (1997) Floor control for multimedia conferencing and collaboration. *Multimedia Systems* **5**(1), 23–38.
- Egido, C. (1990) Teleconferencing as a technology to support cooperative work: its possibilities and limitations, in *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, Galegher J., Kraut R.E. and C. Egido, (eds.), (Erlbaum, Hillsdale), pp. 351–72.
- Ellis, C.A., Gibbs, S.J. and Rein, G.L. (1991) Groupware: some issues and experiences. *Communications of the ACM*, **34** (1), 39–58.
- Engelbart, D.C. (1968) A research center for augmenting human intellect. *Proc. FJCC*, **33**(1), 395–10.
- Erez, M. and Earley, P.C. (1993) *Culture, Self-identity, and Work* (Oxford University Press, New York).
- Gudykunst, W.B. and Ting-Toomey, S. (1988) *Culture and Interpersonal Communication*. (Sage, Newbury Park, CA).
- Heath, C. and Luff, P. (1991) Disembodied conduct: communication through video in a multimedia office environment, in *Proceedings of Computer Human Interaction '91, Conference on Human Factors in Computing Systems*, pp. 99–103.
- Hofstede, G. (1980) *Culture's Consequences: International Differences in Work-related Values*. (Sage Publications, Beverly Hills, CA).
- Hofstede, G. (1991) *Culture and Organizations*. (McGraw-Hill, London).
- Hofstede, G.J. (1996) The windmills of our minds: a workshop on culture clash in CSCW, In *Information Systems and Technology in the International Office of the Future*. Glasson, B., Vogel, D., Bots, P. and Nunamaker, J. (eds) (Chapman & Hall, London), 145–59.
- Ishii, H., Kobayashi M. and Arita, K. (1994) Iterative design of seamless collaboration media. *Communications of the ACM*, **37**, 83–97.

- Jacobson, V. and McCann, S. (1993) Vat – XII-based audio teleconferencing tool, *Unix Manual Page* (Lawrence Berkeley Laboratory, University of California, Berkeley).
- Johansen, R. (1988) *Groupware: Computer Support for Business Teams* (The Free Press: New York).
- Macedonia, M.R. and Brutzman, D.P. (1994) MBone provides audio and video across the Internet, *IEEE Computer*, 27, 30–6
- Mantei, M. (1988) Capturing the capture lab concepts: a case study in the design of a computer supported meeting environment, in *Proceedings of the ACM Computer Supported Cooperative Work*, Portland, OR, pp. 257–70.
- Markus, M.L. (1994) Finding a happy medium: explaining the negative effects of electronic mail on social life at work. *ACM Transactions on Information Systems*, 12(2), 119–49.
- Niemiec, A. (1984) CMITS: communication and craft, in *The Teleconference Resource Book: A Guide to Application and Planning*. Parker, L.A. and Olgren, C.H. (eds), Elsevier Science Publishers, Amsterdam, pp. 109–15.
- Pedersen, P.B. and Ivey, A. (1993) *Culture-centered Counseling and Interviewing Skills* (Praeger, Westport).
- Reinhard, W., Schweitzer, J. and Volsen, G. (1994) CSCW tools: concepts and architectures, *Communications of the ACM*, 27(5), 28–36.
- Root, R.W. (1988) Design of a multi-media vehicle for social browsing, in *Proceedings of Computer Supported Cooperative Work*, pp. 25–38.
- Te Molder, H. (1995) Discourse of dilemmas, an analysis of government communicators' talk. Doctoral dissertation, WAU (Wageningen University, The Netherlands).
- Watabe, K., Sakata, S., Maeno, Fukuoka, K. and Ohmori, T. (1990) Distributed multi-party desktop conferencing system: MERMAID, in *Proceedings of Computer Supported Cooperative Work*, pp. 27–38.

Biographical Note

Dr Schahram Dustdar is the head of the Center for Informatics at the University of Art at Linz in Austria. His main fields of research are organizational implications of information systems and advanced information technology. His current research interests are multimedia information systems and organizational change where he co-authored a book entitled *Multimedia Information Systems* and co-edited a book on telecooperation in organizations. He is on the editorial board of the *Journal of Multimedia Tools and Applications* (Kluwer). He is a visiting research fellow at the Department of Information Systems of the London School of Economics during the academic year 1993/94 and in spring 1998 a visiting research scientist at NTT Multimedia Communications Laboratories, Palo Alto. He is co-founder and CEO of Caramba Labs, a software company developing Java based groupware and workflow systems.

Dr. Gert Jan Hofstede is an Associate Professor of information systems at Wageningen University in The Netherlands. He has worked as an information systems professional since 1984. He has published in various areas about the applicability of formal models in decision support systems. His research interests include Data Modelling, Soft Systems Methodology and the influence of cultural factors on the Information Systems field, in particular for teams collaborating through electronic media.

Address for correspondence: Schahram Dustdar, Center for Informatics (ZID), University of Art at Linz in Austria, Hauptplatz 8, A-4010 Linz, Austria.