Video-Mediated Communication: Psychological and Communicative Implications for Advice on Good Practice.

Christopher Fullwood

Thesis submitted for the degree of Doctor of Philosophy in the Department of Psychology, Faculty of Human Sciences, University of Stirling, June, 2003.

Abstract

This thesis investigates whether certain practices improve the use of video-mediated communication; specifically video-mediated gazing (the act of looking directly into the camera) and face-to-face familiarisation prior to video-mediated meetings. This is done through comparisons of conditions where such practices are employed and control conditions. The successful adoption of these practices is assessed using a multi-level approach: investigating the communicative process, participant perceptions and task outcome. Participant perceptions are directed towards assessing the media, assessing other participants using the media, perceptions of task performance and communicative success, and perceptions of social co-presence. In cases where task outcome is assessed, an objective measurement of performance is taken. Communicative process is assessed through investigating participants use of gazing behaviour and verbal aspects of process: for example turn length, dialogue length and the number of interruptions. Verbal aspects of process are also measured using Conversational Games analysis, where the functions of participants' utterances are assessed.

The results show that participants who gaze at the camera are perceived more favourably. Accompanying speech with video-mediated gazing also results in improved recall of information. Face-to-face familiarisation alters participant perceptions of others using the media and feelings of social co-presence. It is concluded that for certain applications (specifically social tasks) and with an appropriate level of training (specifically with the use of video-mediated gazing) the use of such strategies benefits video-mediated communication.

Contents

List of Appendices

List of Tables

List of figures

Acknowledgements

I would like to thank my supervisor Gwyneth Doherty-Sneddon for all her invaluable help and advice throughout my research and with the final draft of this thesis.

I would also like to thank Bob Lavery for this technical assistance and advice. Thanks also to the Open University for allowing me to use their students in my research. I extend special thanks to the members of Psychology staff at Wolverhampton University for their patience and assistance during the final year of my work.

In addition I'd like to thank my Mother and Father for all their support, advice and faith in me during the past four years. I'd also like to thank Thomas Davies, Ross Fullwood, Les Fullwood, Paul Davies and Alan Morrow for their friendship, support and encouragement.

This thesis is dedicated to my twin brother Ben, for being my biggest fan and my biggest critic.

Declaration

I declare that this work is my own and was completed under the normal supervisory conditions.

Introduction

Research on Video-Mediated Communication

Human communication is a major area of interest in the field of psychology. It has been argued that successful communication is an integration of linguistic and non-verbal processes. Furthermore, human communication traditionally takes place in a co-present, face-to-face context. Technological developments in the 20th century however have introduced new ways for humans to communicate. The telephone for example allows humans to communicate from geographically dispersed locations. The recent development of video-mediated technologies has also introduced a new means of remote communication. It has been suggested that the ability to relay visual images in addition to an audio-channel to communication offers the possibility of being able to more closely replicate face-to-face interaction. The manner in which video-mediated communication (V.M.C) affects our ability to communicate successfully is therefore a major academic question.

Many researchers have focused on what makes video-mediated communication different from face-to-face interaction. There is now a huge literature on the use of video-mediated technologies in applied settings. Much of this research has focused on how well humans adapt remote communication technologies to real world applications. Experimental research has also compared V.M.C with face-to-face and audio-only communication on a number of different types of tasks, for example negotiation and problem solving. Furthermore, various measures of communicative success have been investigated, for example task outcome, communication process measures and participant perceptions. Once more, a common theme in such research is the identification of the ways in which

V.M.C differs from face-to-face and audio-only communication. Far less work has been done on developing practices to improve user use of video-mediated technologies.

A number of ingenious systems have been developed which allow V.M.C to replicate many face-to-face processes. For example eye-tracking systems have been developed to facilitate gaze awareness. A problem with many of these systems however is that they are not always readily available to the user. Therefore, the aim of this thesis is to facilitate user use of the technology that is currently widely available.

This thesis seeks to investigate the effects of two specific practices, which aim to improve the use of existing technology. The effects of face-to-face familiarisation and video-mediated gazing (looking directly into the camera) will be investigated.

One major problem associated with the use of video-mediated technologies is the loss of social co-presence. Essentially, because V.M.C does not take place in a co-present environment, users often report feelings of isolation. This is partly due to the attenuation of visual cues and partly due to physical remoteness. It is therefore expected that reducing psychological distance through face-to-face familiarisation and allowing individuals to use gazing behaviour (and therefore projecting their images more effectively) will result in a number of communicative and psychological benefits. Such research will also contribute to our understanding of the psychology of social relationships and the role that visual signals play in them.

This thesis reports analyses using a multi-level approach. Participant perceptions concerning the successful adoption of video-mediated gazing and face-to-face familiarisation are assessed. These perceptions reflect participant evaluations of the media

and others using the media. They also reflect perceptions of communicative success and feelings of social co-presence. It is expected that participant evaluations will be higher when familiarisation takes place face-to-face and when participants adopt video-mediated gazing behaviours.

This thesis also reports analyses of the verbal and non-verbal channels. Verbal aspects of process are investigated through making global linguistic assessments, for example turn length, dialogue length, and the number of interruptions made by users. Such analyses will indicate whether communication is more efficient and successful. It is expected that face-to-face familiarisation and video-mediated gazing behaviours will result in a more efficient use of language. Verbal aspects of process are also measured using Conversational Games analysis, where the functions of participants' utterances are assessed. In particular, the number of times participants check their partner has understood a message, and the number of times participants check their understanding of a partner's message. This will help to identify how well participants monitor understanding and attention in conversation. It is expected that participants will be better able to monitor understanding and attention when face-to-face familiarisation occurs and when participants employ video-mediated gazing behaviours. This assumption can be justified in terms of an increase in perceived levels of social co-presence.

Communicative process will also be assessed through an investigation of participant use of gazing behaviour. Specifically, the number of gazes to the camera, the amount of time spent gazing at the camera, and the proportion of time spent gazing at the camera. This analysis will give useful insights in to whether participants adopt video-mediated gazing behaviour successfully. Furthermore, gazing behaviour plays a number of crucial roles in

communication. Therefore, an analysis of gazing behaviour will attempt to identify whether gaze can be used in V.M.C to fulfil the same roles as in face-to-face communication.

In cases where task outcome is assessed, an objective measurement of performance is taken. Measures of task performance will help to identify whether video-mediated gazing and face-to-face familiarisation aid users to complete tasks more successfully. A number of different types of tasks will therefore be assessed, including negotiation and problem solving. It is expected that face-to-face familiarisation and video-mediated gazing behaviours will help participants to complete tasks more successfully. If these expectations are met this illustrates that the use of video-mediated technologies can be improved with the use of simple practices that do not require technological alterations or advancements.

Conclusion

In conclusion, there are a number of problems associated with use of video-mediated technologies. One problem in particular concerns a loss of social co-presence, accredited to the attenuation of visual cues and physical remoteness. This thesis addresses these issues by investigating the effects of face-to-face familiarisation and video-mediated gazing. This is achieved through a multi-level analysis of participant perceptions, task performance, and communicative process (both verbal and non-verbal).

Chapter 1 of this thesis reviews literature in the area of human communication. The roles which non-verbal signals, for example gaze and gesture, play in human communication are discussed. The roles that verbal information play in human interaction are also discussed.

The aim of this chapter is to identify whether human communication is more satisfactory and successful when it takes place in a co-present environment and with access to both visual and non-verbal cues.

Chapter 2 covers a review of the literature in the area of video-mediated communication. An introduction to V.M.C, covering background information and the effects of differing qualities and set-ups of systems takes place. Field studies and experimental studies are also reviewed. A review of such studies helps to identify whether video-mediated communication is as successful as face-to-face interactions on a number of different measures, for example task outcome, user satisfaction, and communicative process. The purpose of this chapter is to ascertain whether access to visual information in a remote conferencing environment results in similar benefits as access to visual information in face-to-face interactions.

Chapter 3 reports analysis of questionnaire data and an expert analysis of video-mediated communication. The purpose of this chapter is to identify whether users of current technology view V.M.C as a viable alternative to face-to-face communication. Perceptions from users of the technology and experts in the field of video-mediated communication will also be used to help validate the focus of this research. User perceptions will help to identify the problems that exist in the use of video-mediated communication and methods that can be used to improve the use of the technology.

Chapter 4 reports findings from investigations into the effects of face-to-face familiarisation. Participant evaluations of the media and others using the media are made, comparing face-to-face familiarisation and control conditions. The effects of face-to-face

familiarisation on task performance are also assessed. Comparisons are made between face-to-face familiarisation on negotiation and problem-solving tasks. Finally, an analysis of conversational process takes place, ranging from global linguistic assessments (for example turn length, number of turns, and number of interruptions) to a more detailed coding of utterance function using the Conversational Games analysis.

Chapter 5 reports findings from investigations into the effects of video-mediated gazing. Once more participant evaluations of the media and others using the media are key. In particular, participant evaluations of other participants employing video-mediated gazing behaviours. This is compared to assessments of participant performance when no gazing behaviour is employed. Task performance is assessed in terms of recall of information. Finally, an analysis of gazing behaviour is reported. In particular, the number of gazes made to the camera, the amount of time spent gazing at the camera, and the proportion of speech spent gazing at the camera.

This thesis is therefore an examination of the effects of video-mediated gazing and face-to-face familiarisation on video-mediated communication.

The basic questions addressed in this thesis are:

1) How does video-mediated communication differ from audio-only and face-to-face communication?

2) What types of communicative difficulties arise in the use of video-mediated technologies?

3) Can certain practices be employed to improve user use of current technology?

4) Can face-to-face familiarisation reduce psychological distance and in turn improve communication via video-mediated technologies?

5) Can video-mediated gazing reduce psychological distance and allow participants to project their images more effectively, and in turn improve communication via video-mediated technologies?

Chapter 1: Literature Review:

Research on Human Communication

1.1: Introduction: Chapter and thesis aims

This thesis will consider whether video-mediated communication is as successful a medium of communication as face-to-face interaction. The purpose of the thesis is to consider the psychological and communicative effects of using video-mediated technologies, with an aim to determine whether certain practices can improve the use of existing technology.

In order to achieve this aim it is necessary to evaluate the functions of verbal and non-verbal communication as well as definitions of 'successful' communication. Successful communication will be considered in relation to performance measures (e.g. task performance and user satisfaction), process measures (e.g. the surface structure of conversation and non-verbal responses) and the effects of different tasks on communication. Experimental evidence will also be reviewed for the differences between co-present face-to-face interaction and remote audio-only communication. This will help identify whether human communication is more satisfactory and successful when it takes place in a co-present environment and with access to both visual and non-verbal cues. Psychological and communicative differences between face-to-face co-present interaction and audio-only communication will be explained using theories such as the 'non-verbal hypothesis' (Whittaker, 1995; 1996), Argyle and Dean's (1965) 'equilibrium theory,' and 'media richness theory,' (based on the work of Daft and Lengel, 1986).

1.2: What is communication? A definition

Cherry (1966) noted that the word communication derives from the Latin 'Communico,' meaning 'to share.' Cherry therefore postulates that communication is primarily concerned with the sharing of information. Parsons and Hughes (1970) describe communication as "the process of imparting ideas and feelings" (pg. 1). "Communication can have many purposes – to entertain, to inform, to persuade – but it is basically concerned with producing a change in the attitudes or behaviour of the receiver" (pg. 1). All of the following terms form a regular part of the communication situation.

The 'transmitter' is the individual who attempts to convey something to someone else. Communication begins when the transmitter formulates a message (Parsons & Hughes, 1970). When communicating verbally the transmitter is also referred to as the speaker. The 'message' is the material that the transmitter wishes to communicate (Parsons & Hughes, 1970). The 'medium' is the mode of communication; the channel through which the message is communicated (Parsons & Hughes, 1970). Finally, the 'receiver' is the person who receives and understands the message (Parsons & Hughes, 1970). When communicating verbally the receiver is also known as the listener.

Researchers such as Parson and Hughes (1970), Grice (1969) and Harding (1983) allude to the intentional nature of human communication. According to Grice (1969), the transmitter voluntarily sends a message with the explicit understanding that the receiver is also capable of voluntary action. The transmitter also understands that the receiver is aware that there is an intention to the transmitter's message. Harding (1983) suggests that communication takes place when the receiver reacts to a message (i.e., the message has an effect), at least one of the conversational participants views the situation as

communication, and there is an intentional element to the communication. Communication therefore takes place when the transmitter intentionally imparts a message to the receiver and this message is understood and reacted to. The message can be conveyed through any number of mediums: for example face-to-face, telephone, or video-mediated communication.

Human communication takes place on a number of different levels. Humans can communicate both verbally and non-verbally. Many researchers agree that communication is an integration of non-verbal and linguistic processes (for example, Beattie, 1980; Clark and Brennan, 1991; Goldin-Meadow, Wein and Chang, 1992; McNeil, 1985; Weiner, Shilkret and Devoe, 1980). According to Doherty-Sneddon et al this would indicate that although communication lacking visual signals (audio-only communication) is possible, it is likely to be different, or perhaps in some cases even inferior (Doherty-Sneddon, Anderson, O'Malley, Langton, Garrod & Bruce, 1997). In face-to-face communication there are two different types of visible information. The first is information about the behaviours of other conversational participants and the set of non-verbal communicative actions that they perform (such as hand gestures, eye gaze and facial expressions). The second type of visible information concerns the visible environment that conversational participants share, including access to shared objects, shared events, and information regarding the movements and activities of other people (Whittaker and O'Conaill, 1997).

So, communication is about the sharing of information or meaning. This occurs through multiple communication channels and involves speech and a variety of visual communication cues. The following sections will address the importance of these cues in human communication and the functions that they perform. Human communication

(verbal and non-verbal) can be categorised by two separate functions: those that co-ordinate content and those that co-ordinate process (Whittaker and O'Conaill, 1997). Content refers to the actual subject matter of the conversation, and therefore co-ordination of content is concerned with how participants arrive at, and maintain, common understanding in conversation (Clark and Brennan, 1991). Process refers to the actual mechanisms of the conversation itself, and therefore process co-ordination involves "the set of procedures by which participants agree to begin and end entire conversations and the rules that allow participants to switch roles between speaking and listening" (Whittaker and O'Conaill, 1997, pg. 25).

What must be considered, however, is that although emotional information can be considered content (for example yawning for effect to signal that 'I am bored'), if there is no intended communicative function for emotional information (for example yawning because you are tired), then such examples cannot be considered content. Although such information can be interesting and may be reacted to, according to Parson and Hughes' (1970), Grice's (1969) and Harding's (1983) definitions of communication cannot be considered communication due to a lack of intentionality. It should also be noted that although there is a distinction between content and process co-ordination, they are not necessarily mutually exclusive. Normally, good process feeds in to good content. Further to this divide, the importance of visual and auditory cues will be related to the needs of human interaction, as outlined by Daly-Jones, Monk and Watts (1998). These are: to make contact, to allocate turns during conversation, to monitor understanding and attention, and to support deixis.

In the following section a review will take place concerning the functions of oral communication. A distinction here should be made between oral and verbal

communication. Oral communication includes verbal information (i.e. language), and non-verbal vocalisations. Although non-verbal vocalisations can be considered non-verbal behaviour, a review of such behaviour will take place in section 1.3 on oral communication. The reason for this distinction concerns the fact that this thesis concentrates on the differences between information that can be transmitted with access to visual signals and information that can be transmitted without access to visual signals: non-verbal vocalisations do not require a visual channel in order to be transmitted. This will be followed by a review of the functions of non-verbal communication (in section 1.4). This will include gaze, gesture, facial expressions, posture and information concerning the environment: all of which require a visual channel in order to be transmitted. The aim of these sections is to discover whether visual information adds value to communication over and above oral communication. Whereas non-verbal and verbal processes are intricately linked, for the purpose of the following sections they will be considered separately.

1.3: Oral Communication

1.3.1 Introduction

Human language has evolved as a mouth-to-ear system of communication. Taylor (1976) indicates that such a method of communication is advantageous for a number of reasons. Oral communication requires no line of sight for it to take place; it can be used day and night; it requires little energy; and it is flexible (many different sounds can be combined in a vast number of ways). Typically, verbal information has been considered more important than non-verbal information with non-verbal behaviour viewed as a rudimentary predecessor or redundant associate to verbal information (Weiner, Shilkret, & Devoe, 1980). The validity of this view will be discussed in further detail later in the chapter. The use of language is what separates us from non-human primates. Non-human primates also depend on visual communication more than the oral-auditory channel. In fact, it has been estimated that only five percent of non-human primate social messages are expressed in the vocal channel (Altman, 1967). Further to this, animal communication has been noted to be limited to directives, commands and expressions of emotion (Taylor, 1976), and attempts to teach animals human language have been met with limited success (for example Hayes, 1951). Humans on the other hand use words as symbolic representations to stand for events and objects in the real world.

The number of phonemes (speech sounds) that a human being is capable of producing is far greater than those used in a given human language. Although humans are capable of producing several hundred phonemes, each language is typically built from approximately 20 to 50 phonemes (Parry, 1970). There are a number of exceptions however, and Hawaiian, for example, only uses 12 speech sounds, whereas Abkhaz uses 70 distinct sounds (Taylor, 1976). Languages also differ in how these sounds are combined to form

'morphemes,' "the minimal unit of grammatical analysis" (Taylor, 1976, pg. 51) and the basis stems of words. A word can be a representation for persistent or recurrent elements, for example an object or person. Words are also symbols for an entire class of rather different objects, for example 'father' can stand for all male parents. Words can also be applied to events, ideas and processes (Parry, 1970). Lyons (1968) emphasises that words tend to be internally stable. In other words the order of phonemes tends to be consistent, for example 'book-s' is acceptable, whereas 's-book' is not. Lyons also indicates that words are positionally mobile, in that they can come in a number of different positions in any sentence (for example 'that *ball* is red' and 'that is a red *ball*'). In all human languages, words are combined together to form sentences, and further to this all languages have universal rules concerning grammar, syntax, semantics, and phonology (Chomsky, 1957). Verbal behaviour can be further categorised as utterances, which may consist of the verbal equivalent of a sentence. The smallest utterances used by humans consist of grunts, for example the use of verbal agreements such as 'mhm-hmm,' or 'uh-huh.' (Argyle, 1969).

Utterances can be classified in many different ways. For example, an utterance can be classified in relation to its intended function (described more fully in relation to speech acts later in the chapter). Bales (1950), for example, found that verbal utterances could be classified into 12 distinct categories. These include agreement, disagreement, asking for opinion, asking for suggestion, giving suggestion and giving opinion. A further distinction can be made concerning whether the utterance is about something external to the conversational partners or whether the utterance is directly related to conversational participants or the interaction that is taking place between them (Lennard and Bernstein, 1960). Different classes of utterance may have different effects on the interactors, for

example, a question should lead to an answer (whether it be right or wrong) or an order may lead to an action. Such sequences form what has been called 'adjacency pairs' (Schegloff & Sacks, 1973). Utterances may be further categorised by the topic of conversation. Topics can be personal, impersonal, abstract or remote, for example (Argyle, 1969). While vocal communication can be considered on many levels, the primary focus of this thesis is at the level of communicative function.

1.3.2 Functions of oral communication

Oral communication serves many functions. Oral communication can be used to co-ordinate conversational content (for example monitoring understanding, feedback cues and interpersonal information cues) and co-ordinate conversational processes (for example making contact and interaction management) (Whittaker and O'Conaill, 1997). All these functions are discussed in the following sections.

1.3.2A Making contact and availability cues reference

Communication is a joint activity that requires both co-ordination of process and content (Clark & Wilkes-Gibbs, 1986). The first necessity of human interaction is to identify a conversational partner and then establish contact with him/her. Linguistic behaviour is particularly useful in initiating an interaction with another person, for example through explicit vocal greetings such as "hello" or specific questions such as "how are you doing?" The recipient is not a dormant partner in this exchange, however, and may indicate a willingness to be contacted by providing a vocal acknowledgement, or indeed may signal vocally that they do not want to be contacted (Daly-Jones et al, 1998). Schegloff (1968) has coined this co-operative process a 'summons-answers sequence.' In other words, the process through which the caller attempts to make contact with a recipient, who in turn

signals whether or not they are available for communication. This process however, does not necessarily have to take place face-to-face, and indeed Schegloff (1968) points out that a telephone ring and the subsequent answering share many similar characteristics. As will be noted in section 1.4 on non-verbal communication, a number of researchers consider visual cues to be particularly important for making contact with others (for example, Heath 1984; Suchman and Wynn, 1984; Schegloff, 1968; Frolich, 1995).

1.3.2B Turn-taking and interaction management

Once an exchange has been set-up it also has to be managed. One important aspect of this is the negotiation of conversational turns between speakers. In order to attain a high level of interactivity, switches between speaker turns should be smooth and should not disrupt the overall flow of conversation. In order to achieve this speakers can use a number of devices (for example intonational, syntactic and pragmatic) in order to signal that they are about to finish their turn. The fact that listeners are able to predict when speakers are about to finish means that pauses between speaker switches are usually very small. Jaffe and Feldstein (1970), for example, found that pauses usually varied between 620 msec and 770 milliseconds. In many cases pauses do not occur at all, with an actual, but typically not perceptual overlap in speech being evident. Turn-taking is also regulated in the linguistic domain through the use of overt cues. In order to select the next speaker, individuals can aim direct questions at others for example.

Several aspects of speech contribute to the co-ordination of turns within a conversation. Duncan and Fiske (1977), note that transitions from one speaker to the next are co-ordinated by the completion of a grammatical clause, the rise or fall in voice pitch at the end of an utterance, any drawl or expansion of the final syllable and the use of stereotyped

expressions such as 'you know.' These behaviours have been referred to as 'regulators' (Ekman and Friesen, 1969) of talk and help to structure turns and guide the flow of conversation. Other aspects of spoken delivery for example restarts, hesitations and pauses in speech can also serve in aiding the transition from one speaker to another. An example of this comes from Schegloff (1987), who described 'recycled turn beginnings.' A 'recycled turn beginning' occurs when a conversational participant pauses and then restarts immediately. This device is used to keep hold of the conversational floor through claiming another turn and holding onto the floor even if this turn is interrupted by other conversational participants.

During conversation speakers can also formulate messages that define the possible responses of the listener. Thus a question begs an answer. For example, asking someone directly "what is your opinion on this?" requests another individual to take a turn, but limits the response to the topic indicated by the speaker. This particular resource identifies clearly the potential points of entry into the conversation, and in turn provides a means by which to judge the how appropriate the response is (Sacks, Schegloff, & Jefferson, 1974).

1.3.2C Feedback cues: Monitoring understanding and attention

The communication process is more complex than the expression of predefined sentences at suitable moments. Rather, communication is a dynamic process in which the composition and delivery of our utterances are linked to the responses and actions of those with whom we are conversing. In other words the manner in which we communicate is affected by the responses, both verbal and non-verbal, of those we are talking to (Daly-Jones et al, 1998). According to researchers such as Clark and Schaefer (1989), Goodwin (1979) and Garfinkel (1967) speech is not predefined but rather is formed to a large degree

during conversation. The unpredictable and dynamic nature of human communication requires both the speaker and the listener to be able to judge understanding and attention effectively. There is little point in conveying a message if the listener has not listened or has misunderstood the content of the message.

Even with visual indicators of attention (discussed in section 1.4), conversational participants may still find it necessary to volunteer explicit verbal confirmations of their understanding of a previous utterance. The use of such formulations indicates that shared knowledge can be clarified during conversation in the auditory domain, and indeed this is very useful (Heritage and Watson, 1979). Another auditory resource that can be used to assess a listener's level of understanding comes from the manner in which conversation is organised into 'adjacency pairs' (Schegloff, 1968). As well as being useful for interaction management (for example, setting out the parameters of the conversation itself), a request for information may also be viewed as understood if an appropriate answer is given to the question. If the question "where did you go last night?" is followed by the response "green and blue" then this will serve as a clear indication that the message was not understood or received by the listener (either that or the listener does not speak English very well). Listeners can acknowledge speakers through the use of 'continuance expressions.' Such behaviours include phrases such as 'go on' (Schegloff, 1991), assessments, for example 'o.k,' and 'yes' (Clark and Schaefer, 1989) and backchannel cues, for example 'mm-hm' (Yngve, 1970). These behaviours provide a good indication of the listener's level of attention and understanding. They may also indicate levels of participant satisfaction with the conversation itself (Daly-Jones et al, 1998).

A key aspect of human interaction is therefore listener feedback. The speaker delivers utterances incrementally, while the listener concurrently signals to the speaker that the utterances have been understood. According to Clark and Schaefer (1980), this positive concurrent feedback informs the speaker that he or she can reply and build upon the listener's understanding. In cases where this feedback is absent or even delayed, the speaker's ability to formulate efficient messages is reduced (Krauss and Bricker, 1967). Without feedback, whether verbal or non-verbal, speakers are unable to assume the message has been understood and may, therefore, attempt to clarify or reiterate points, sometimes without need, to ensure that the listener has not misconstrued what has been said. Absence or delay of feedback can therefore encourage the speaker to take longer turns and explain points in more depth than perhaps is necessary (Krauss and Bricker, 1967). In normal face-to-face interaction, the flow of the speaker is not interrupted by backchannels because the audio channel is two-way, although backchannels may occur at the same time as the other speaker's speech, but will not be perceived as interruptions. The communication process is also interactive in the sense that breakdowns in understanding can be resolved by the listener. For example, the listener can interrupt the speaker and request clarification or register disagreement (Clark and Shaefer, 1980).

Duncan and Niederhe (1974) provide a useful distinction between vocal backchannels and visual backchannels. They state that visual backchannels include head nods and headshakes, but may also comprise of a number of facial expressions such as smiles and frowns. Vocal or auditory backchannels include guttural acknowledgements (such as 'mm-hm'), sentence completions, requests for clarifications and brief restatements. Brief restatements constitute what Jefferson (1972) has called 'side-sequences,' where the content of what has already been stated is reiterated for further clarification, so that the

conversation can proceed as planned. This type of backchannel is beneficial to both the listener and the speaker. Kraut, Lewis and Swezey (1982) conducted a study in which a speaker related a story to two listeners. One listener was allowed to provide vocal feedback and influence the speaker's description, whereas the other listener had to remain silent. Results from the study indicate that active participants recall more of the storyline than the silent participants. Kraut et al suggest that feedback allowed the active participants to have sections reiterated that they didn't understand or didn't hear properly, which in turn lead to a greater understanding of the story. The silent participants on the other hand, could not have information that they had not understood or not heard reiterated, unless it was the same material that the active participant wanted to clarify.

A problem for speakers in co-ordinating content is to determine whether their utterance had the anticipated effect. In other words did the listener draw the correct set of conclusions from what has been uttered? The speaker does not usually have direct access to the listener's knowledge and beliefs (Walker, 1993), so feedback in conversation is crucial for the maintenance of mutual understanding. Speakers provide listeners with frequent opportunities to offer feedback about what was just said (O'Conaill, Whittaker and Wilbur, 1993) in order that they can show (Clark and Schaefer, 1989) or clarify their level of understanding (Walker and Whittaker, 1990). These feedback processes take place on a moment-by-moment basis in conversation, so that misunderstandings can be quickly identified and rectified. In other words, common ground is updated accurately with the process of 'grounding.' In conversation this means that conversational partners attempt to establish mutual understanding to the extent that the receiver has understood what the transmitter meant to a satisfactory criterion for present purposes (Clark and Brennan, 1991).

1.3.2D Verbal indicators of interpersonal information

A final aspect of content co-ordination concerns the affective state and interpersonal attitudes of the participant. This includes social information about the participants' feelings, emotions, and attitudes towards the other conversants and to what is being discussed. Participants generally do not make this information explicit, so it usually has to be inferred (Whittaker and O'Conaill, 1997). Access to affective information is important: it can change the outcome of conversations in situations where emotion plays a critical role, for example negotiation (Short, William and Christie 1976).

Non-verbal vocalisations may be used to portray emotional information. In a review of 39 studies, Scherer (1986) found that pitch, intensity, intonation and vocalisation rate are used to signal emotions. Joy and elation, for example, are signalled through raised pitch, variability of pitch, range of pitch in an utterance, pitch intensity, speech rate, and gentle contours. Depression is signalled through lowered pitch, intonation, range of pitch, pitch intensity, speech rate, and reduced energy at high pitches. Argyle (1988) argues that the main dimension for emotion in vocalisations is arousal, which is communicated by loudness, pitch and speed. Joy, anger, confidence and fear all score high on these parameters. In order to distinguish between these emotions it is argued that pitch contour is important. Frick (1985), for example, noted that happiness is signalled by gentle contours, anger by sudden increases in pitch contour, surprise in a rising pitch, and sarcasm in a fall on a stressed syllable. Argyle (1988) also argues that the voice is a 'leakier' channel than the face, in that facial expressions are more easily controlled than vocalisations. Hence the voice is more likely to reveal true feelings. Further to this,

individuals can signal feelings and emotions verbally, for example 'I am sad.' Non-verbal indicators of emotion are discussed in section 1.4.

1.3.2E Supporting deixis

Deixis can be defined as "the use of a word or expression whose referent can only be understood by taking into account the non-linguistic context in which the word is used" (Sutherland, 1995, pg.118). According to Daly-Jones et al (1998) people can achieve higher levels of clarity in conversation through reference to shared artefacts. Terms such as 'that,' 'this,' 'he,' and 'she' can be used as a form of linguistic pointing (Jarvella and Klein, 1982), aiding to make reference to objects, events, people (referents) more efficient. According to McCarthy and Monk (1994), the ability to refer to shared artefacts in everyday communications allows for higher levels of efficiency in conversation. As will be noted in section 4.1, however, non-verbal communication also plays an important role in supporting deixis.

<u>1.4: Non-verbal communication</u>

1.4.1 Introduction

Non-verbal communication includes information from the eyes (gaze), face (facial expression), hands and arms (gestures), and the movements and orientation of the body (posture) (Argyle, 1988). The purpose of this section is to evaluate the functions of such non-verbal behaviours in face-to-face communication with the intention of finding whether they are important for human interaction and add value to communication over and above the information transmitted in speech. Traditionally non-verbal behaviour has been seen as portraying primarily emotional information, however as will be discussed in the following section, non-verbal communication also contributes a significant amount of propositional

(real content) information to communication. It is important to note here that the correspondence between visible information and communicative function is not direct: the same type of visible information can support multiple communicative functions, and likewise different types of visible information may support one communicative function. For example gaze supports multiple communicative functions. Gaze can co-ordinate reference, give feedback of understanding the utterance, and aid turn-taking. Similarly, one communication function may be mediated by multiple different types of visible behaviour. Thus gaze, gesture and posture can all support turn-taking (Beattie, 1981 and Kendon, 1967).

As gazing behaviour is of particular relevance to this thesis, the communicative functions of gaze will be discussed separately. Although other forms of non-verbal behaviour (gesture, facial expressions, and posture) are also important, the experimental chapters of this thesis do not focus on such non-verbal behaviours, and therefore they will be discussed concurrently.

1.4.2 Gazing behaviour

One type of non-verbal behaviour that has been studied in depth with respect to its importance in the communication process is gaze. Gazing behaviour is both a signal and channel: a channel for the transmitter (gazer), and a signal for the receiver. The gazer can perceive the expressions of others, for example visual information from the face. The act and manner of looking can be interpreted by the receiver and can have meaning as signals, for example the amount of interest in another person (Argyle, 1988). Studies of gazing behaviour have noted that participants spend relatively long periods of time looking at others' faces. The speaker for example spends approximately 40% of time looking at their

conversational partner's face. The listener, on the other hand, will typically spend 75% of time looking at their conversational partner's face (Argyle, 1988). Mutual gaze occurs less frequently, and typically occurs 30% of the conversational time, however recent studies have reported levels below 5% (Anderson et al, 1997). The exact amount of gazing behaviour is influenced by a number of factors, including the relationship between the participants, and the communication task in hand. Gazing behaviour changes significantly in the presence of relevant objects, for example a map during a map reading exercise. In this situation amounts of gaze directed at others can be as low as 3% to 7% of conversational time (Argyle, 1988).

Gaze has a number of communicative functions. Gazing behaviour can be used to co-ordinate conversational content (for example monitoring understanding, feedback cues and interpersonal information cues) and co-ordinate conversational processes (for example making contact and interaction management) (Whittaker and O'Conaill, 1997). All these functions are discussed in the following sections.

1.4.2A Making contact and availability cues reference

Many communications are unplanned and therefore require participants to establish precisely when other people are available for communication and when it is opportune to initiate interactions. As discussed verbal expressions provide valuable mechanisms for making contact, but so too do visual cues. Indeed, if available, visual cues are typically the first source of information used in determining whether or not to initiate an interaction. Gazing behaviour is particularly important for making contact with others. Gaze can also be used as a channel to assess availability, even in remote communication technologies that support access to visual cues (Fish, Kraut, Root & Rice, 1993). From the 'caller's'

(the individual attempting to start up a conversation with another) perspective it is essential to locate the intended recipient and in turn assess their availability for interaction. This is based on the awareness of others' movements and activities (Heath and Luff, 1991). For example, should a person be currently engaged in some form of activity then it is up to the caller to make a decision concerning when it would be appropriate to interject based on information gained from observing the individual (Frolich, 1995). Changes in gazing behaviour could then be used to initiate interactions. Actively gazing in the direction of the intended recipient of conversation is likely to result in a response for example (Kendon, 1967; Heath and Luff, 1992). The recipient is not a silent partner in this exchange, however, and may indicate that he/she is unwilling to be contacted, through avoiding eye contact (Daly-Jones et al, 1998). Indeed, avoidance of eye contact is interpreted in a number of socially relevant ways (many of which will be discussed later in the section).

In addition, other information about physical appearance is conveyed visibly and can supply inferences about other participants' gender, age and possibly disposition. Sightings of others can lead one to fall into spontaneous conversation, such as in public areas like coffee areas (Kendon and Ferber, 1973). In the workplace, merely seeing a co-worker may also remind one of an issue that needs to be discussed, so that the sighting serves as a reminder that a conversation needs to take place (Kraut, Fish, Root & Chalfonte, 1993). Not only does visible information allow us to know if a person is available, but also influences the termination and make-up of conversation. For example, the arrival of a third party can prompt an interaction to end or change the nature of an interaction. It is often the case that this third party indicates a desire to interrupt or join the conversation by 'hovering,' in other words waiting for the current conversation to a reach a point where they can break in (Whittaker, Frolich & Daly-Jones, 1994).

1.4.2B Turn-taking and interaction management

Gazing behaviour therefore facilitates the starting of conversations. In addition to this macro level of task management, some researchers have proposed that gaze serves to help regulate turn-taking in human interactions. Kendon (1967) noted the importance of gaze in facilitating the turn-taking process. He observed how his participants tended to look away more from people when speaking themselves, but when arriving at the end of a conversational turn they would begin to gaze more at the person they were speaking to. Kendon interpreted these results to indicate that gaze plays a crucial role in conversational turn-taking. Kendon's findings have been supported by other researchers, for example Duncan and Niederehe (1974), Levine and Sutton-Smith (1973) and Goodwin (1981). Kendon's work, however, has been criticised on a number of levels, including the small sample from which he collected his data. More importantly, however, research has shown that turn-taking can occur without the use of visual cues.

Beattie (1981) suggested that non-verbal cues, including gaze, are not crucial in regulating the exchange of speaker turns. In attempting to identify non-verbal signals that were regularly used to facilitate the turn-taking process, he did note a correlation between a subject's use of turn yielding cues and the probability of that subject taking up a conversational turn. Beattie, however, proposed that syntactic and paralinguistic signals were also highly correlated with a subject taking up a conversational turn, and therefore visual signals are not essential in facilitating turn-taking. Furthermore, we manage to take turns on the phone even when visual signals are not available. However, as described in the next chapter, visual signals can influence the process of turn-taking, resulting in a different communication style.

1.4.2C Feedback cues: Monitoring understanding and attention

Eye gaze has been noted as a good indication of attention and understanding. Gazing behaviour is an obvious means for a recipient to signal to a speaker that they are actually attending to what is being said and, in group settings, as a way of indicating to whom a speaker's remarks are addressed. Eye gaze has a number of co-ordinating functions, which have a number of cognitive and social effects. Gaze may be co-ordinated between conversational participants to achieve mutual gaze or joint attention. Mutual gaze occurs when two participants are looking at each other simultaneously. Joint attention occurs when participants are jointly orientated to a common part of their shared visible environment and at the same time aware that their partner/partners are also looking at it. This helps to facilitate joint attention as the speaker can make judgements concerning which objects are important to the conversational partner, which in turn allows for greater flexibility in referring to objects (Clark and Marshall, 1981). People tend to be quite skilled at determining where others are looking. Research conducted by Rutter (1984) indicates that people can estimate where others are looking from their head and eye position, with great accuracy and after little practice. Gale and Monk (2000) showed that participants estimated with a low degree of error where another individual was looking on a flat stimulus between them. Knowing what object someone is looking at is an important aspect of conversation, as it lets us know what someone is attending to. This potential conversational resource has become known as 'gaze awareness.' Monk and Gale (2002) make a distinction here between 'full gaze awareness,' which is the ability of someone to estimate which object an individual is currently looking at, and 'partial gaze awareness,' which is the ability to gauge the general direction someone is looking in (i.e., up, down, left or right). Gaze awareness may be particularly useful in supporting deixis. Daly-Jones

et al (1998), for example noted that reference to shared artefacts (which could be achieved through gaze awareness) results in higher levels of clarity in conversation.

Gazing behaviour can act as a channel. Listeners can derive important non-verbal information about a speaker, for example his or her facial expressions, posture or use of gestures. In this way, the speaker's visible behaviour helps clarify the content of what is being said. According to researchers such as Beattie (1981) and Clark and Brennan (1991) speakers interpret listeners' attentional behaviour as feedback for determining how well their message is being understood and co-ordinate their content accordingly (Whittaker and O'Conaill, 1997). So if a listener indicates a need for feedback or information through their patterns of gaze, for example through increased levels of gazing (Burton, McGregor, and Berry, 1979) or facial expressions which indicate puzzlement, the speaker may want to clarify what has already been said, or direct questions to the listener in order to decipher why this lack of understanding has come about.

1.4.2D Gaze as an indicator of interpersonal information

Gaze is also an indicator of interpersonal attitude or affect. Speakers tend to gaze at a listener's face more when they are being persuasive (Mehrabrian & Williams, 1969), deceptive (Burns & Kintz, 1976), ingratiating (Pelligrini, Hicks & Gordon, 1970), or dominant (Exline, 1971; Ellsworth, 1975). In addition to this patterns of gaze directly affect evaluations of people. People who look at their conversational partner infrequently may be judged as being 'defensive' or 'evasive.' On the other hand, those who look more frequently at their conversational partner may be judged as being 'friendly,' 'mature' and 'sincere' (Kleck and Nuessle, 1968).

So gaze is an important signal in providing listeners with information and influences such evaluations as credibility, attentiveness, competence and dominance. Many studies in this field indicate that gaze is used by people as a sign of liking and attraction. For example, observers tend to agree that men and women share a higher level of sexual interest when accompanied by a higher level of mutual gaze (Thayer and Schiff, 1974). Moderate levels of gaze appear to be preferred over constant or low levels of gaze, as high levels appear to be uncomfortable and intimidating and low levels are associated with shy, weak and socially inadequate people. This is indicated in a study by Abele (1981) who revealed that observers favour people engaged in social interaction when they employ moderate levels of gaze as opposed to low levels of gaze. Gaze is seen also to enhance attraction and liking when people are introduced in to social situations. Exline and Eldridge (1967), for example, noted that confederates who gazed frequently received higher evaluations from participants than confederates who gazed infrequently. With regards to levels of attentiveness, judgements are also made about people based on their gazing behaviour. Kleinke, Staneski and Berger (1975) found that in interview scenarios interviewees were rated as being more attentive by the interviewer when higher levels of gaze were employed. Furthermore, interviewees gave briefer responses when the interviewer gazed less in their direction, showing the importance of gaze as a feedback cue to the speaker.

Levels of gazing also seem to correlate with ratings of intelligence or job suitability. Participants have been shown to rate people in photographs as more deserving of a job or a higher salary when they looked straight ahead as opposed to looking in a downward direction (Tankard, 1970). A study by Hemsley and Doob (1978) has indicated that gaze aversion reflects beliefs concerning a person's credibility. For example an attorney is more likely to judge a witness as credible if he/she does not avert gaze during questioning.

This is also evident in connection with customs inspection, as travellers are more likely to be searched if they avoid eye contact when being questioned (Kraut and Poe, 1980). Perceptions of what different levels of gaze mean are not always valid, for example the belief that 'looking shifty' indicates that an individual is not credible. Zuckerman & Driver (1985) indicated that highly motivated liars could overcompensate for these usual perceptions of credibility, through exhibiting high levels of gazing

Gaze can also serve as an indicator of dominance. Thayer (1969) manipulated gaze between two men in a live encounter. Men in both situations were evaluated as being more dominant when they gazed frequently as opposed to infrequently. In mixed gender role-play situations Lochman and Allen (1981) reported that women participants also rated men who gazed infrequently as being silent, passive and inattentive. On the other hand men did not appear to be influenced by female participants' lack of gazing. This finding could be explained in evolutionary terms, as it is evident that men are usually regarded as dominant and women respectively submissive. Therefore it is not surprising that men failed to pick up on women's submissive behaviour, as this would have been considered normal behaviour. On the other hand, women picked up on men's submissive behaviour, as this would have essentially been going against usual expectations. Gaze therefore communicates a number of communicative messages to the receiver and can influence decisions concerning liking and attraction, dominance and attentiveness.

1.4.2E Cognitive effects of gaze

In addition to its social impact gazing behaviour has a number of cognitive effects. Beattie's (1981) 'arousal hypothesis' suggests that excessive levels of inappropriate gazing result in high levels of physiological arousal, which may interfere with cognitive

processing. Such interference may have an effect on an individual's ability to perform cognitive tasks, for example memory tasks. This theory is also borne out by Glenberg, Schroeder and Robertson's (1998) cognitive load hypothesis. Glenberg et al found that people averted their gaze when attempting to answer questions that were deemed moderately difficult. The authors go on to propose that such behaviour is beneficial as it allows the individual to disengage from environmental stimuli (for example, the other individual's face), and therefore enhancing performance directed by non-distracting stimuli. Experimental evidence for the effects of gazing on task performance comes from Doherty-Sneddon et al (1997). Doherty-Sneddon et al found that the use of video-mediated equipment designed to mimic natural eye contact ('videotunnels,' Smith, O'Shea, O'Malley, Scanlon & Taylor, 1991) did not result in improved task performance over audio-only and video-mediated communication without eye contact. In fact, the dialogues in the video-tunnel condition (eye contact) were significantly longer (11% more turns and 10% more words), and therefore less efficient. Doherty-Sneddon et al go on to suggest that the participants using the videotunnels actually over-gazed, in other words video-mediated communication with eye contact encouraged participants to overuse the visual channel. Therefore, the benefits of having access to visual signals was counterbalanced by users becoming distracted by their partner's face and so using significantly more speech to achieve a comparable level of task success.

Many researchers have indicated that levels of gaze can result in changes in a person's physiological arousal. Direct gaze from an experimenter has been show to result in increased galvanic skin responses (Nichols and Champness, 1971; Strom and Buck, 1979). Kleinke and Pohlen (1971) noted increases in heart rate by college men when stared at by another male during a competitive game. Furthermore increases in arousal can effect

cognitive performance. The effects of arousal on cognitive performance can be explained using the Yerkes and Dodson law (1908). This law predicts a U-shaped relationship between arousal and performance. A moderate amount of arousal is required to provide the motivation to learn. Too much or too little arousal will work against the learner. Too much arousal has a hyperactive effect, and in cases where intellectually challenging (cognitive) tasks are required, lower levels of arousal are needed.

The effects of gazing behaviour on arousal also tie into another communicative function of gaze: to emphasise verbal information. Whereas research indicates that excessive amounts of gazing can interfere with cognitive processing, experimental evidence also suggests that gazing behaviour can improve memory for verbal information. For example, research has indicated that students remembered more instructions from a teacher who gazed at them more frequently (Fry & Smith, 1975). Similarly, primary school students remembered more of a story when their teacher gazed more frequently (Otteson & Otteson, 1980). These findings can be interpreted in relation to increased physiological arousal as a direct result of gazing behaviour. Whereas high levels of gazing may be detrimental, moderate levels of gazing will result in moderate levels of arousal, which are required to provide the motivation to learn (Yerkes and Dodson, 1908). It may also be the case, however, that in such instances gazing behaviour serves as a non-verbal indicator to important information that requires attention, much in the same way that eyebrow movements are used to reinforce important aspects of speech (Ekman, 1979; Whittaker and O'Conaill, 1997).

1.4.3 Gesture, facial expressions and posture

Other non-verbal behaviours that supplement verbal communication include posture, gesture and facial expressions. Posture can be defined as the information supplied by the

inclination and orientation of a conversational participant's body, in particular their trunk and upper body, although the positions of both the arms and the legs are also important (Whittaker and O'Conaill, 1997).

A gesture is best defined as "voluntarily bodily actions, by hands, head, or other parts of the body, which are intended to communicate" (Argyle, 1988, pg.188). Different classes of gesture have been identified. 'Emblems' are often used at the same time as speech (Argyle, 1988) but can also have conventionalised meanings and can substitute for phrases or spoken words, for example in the use of symbolic gestures such as the "thumbs up" or "V for victory" (Efron, 1982), and can therefore stand alone without speech. Another class of gesture are referred to as 'illustrators.' Illustrators are more directly related to speech than emblems and serve to clarify or repeat what has been said. An example of an illustrator is the 'iconic' gesture, which express propositions of a dynamic or spatial aspect. McNeill and Levy (1982) referred to iconic gestures as ones that "in form and manner of execution exhibit a meaning relevant to the simultaneously expressed linguistic meaning" (McNeill, 1985, pg. 354). An example of an iconic gesture would include the phrase "it moved like this" where the word "this" is illustrated by a gesture that infers the type of movement. Another type illustrator is known as a "beat," but has also been referred to as "batons" and "speech primacy movements." "Beats" allow speakers to "emphasise, evaluate, or 'comment on' the information they are conveying verbally" (Whittaker and O'Conaill, 1997, pg. 30) and therefore also serve to supplement verbal information.

Gesture may also be used to communicate more abstract characteristics of the space the speaker is talking about, such as the relative positions of the objects it contains, and their relative orientations (McNeil, 1992). McNeill (1985) noted that humans not only use hand

gestures to communicate concrete images (for instance actions or attributes of people, objects or events), but also to portray abstract concepts (such as mathematical concepts of factors or quotients). Gestures can stand alone, or be used to supplement verbal information. Evans and Rubin (1979) conducted an experiment in which 5 to 10 year olds were taught to play a simple board game, and then instructed to explain the rules to an adult. Evans and Rubin noticed that the children's explanations were accompanied by gestures, and approximately half of the gestures used supplied information that was not expressed verbally. This finding is particularly important with regards to communication that takes place with a restriction of visual cues, for example telephone conversations. During audio-only communication such information would not be transmitted and therefore would need to be expressed verbally. This may have a number of communicative effects: the information may not be expressed adequately, or it may take longer periods of time to express the information verbally. So gestures can be used to communicate a number of different messages. Furthermore, gestures can stand alone or be used to supplement verbal information.

Facial expressions are particularly important for the expression of emotion, attitudes towards other people and play a crucial role in social interactions (Argyle, 1988). Facial expressions also serve conversational functions, for example eyebrow movements can be used to emphasise speech (Ekman, 1979; Whittaker and O'Conaill, 1997). Facial expressions, especially lip and teeth movements are also important in deciphering speech. The shape of the lip, teeth and tongue supply the listener with useful added information concerning phonemes produced by the speaker. Unintelligible speech can therefore be interpreted with the aid of such visual information (Whittaker and O'Conaill, 1997). According to Summerfield (1992) when lip shape information is available, listeners can

interpret an additional 4-6db of noise and achieve the same level of intelligibility. The effect that visual information has on speech perception has been demonstrated clearly in what is known as the "McGurk effect." Using dubbed videos McGurk and MacDonald (1976) noted that if the lip movement indicates the verbalisation of the phoneme "ga" and the voice pronounces the phoneme "ba" this conflicting information is combined in a way that is interpretable and therefore the sound "da" is heard. As well as having a number of other communicative functions (discussed later in the section), access to facial expressions in communication (especially lip and teeth movements) allow verbal information to be interpreted more successfully.

Gesture, facial expression and posture support numerous other communicative functions. They can be used to co-ordinate conversational content (for example monitoring understanding, feedback cues and interpersonal information cues) and co-ordinate conversational processes (for example making contact and interaction management) (Whittaker and O'Conaill, 1997). All these functions are discussed in the following sections.

1.4.3A Making contact and availability cues reference

Making contact, or greeting someone, can be viewed as a ritual characterised by a series of symbolic acts, the function of which are to initiate an encounter (Goffman, 1971; Schiffrin, 1974). Body contact, posture, facial expressions and gesture all play important roles in making contact. In a naturalistic observation study, Greenbaum and Rosenfeld (1980) noted that the majority of greetings included some sort of touch (for example, a handshake, an embrace, or a kiss). Posture may be used to initiate interactions, for example through changes in body orientation (Kraut and Higgins, 1984). Orientating one's body position

towards another initiates interaction, however, this form of non-verbal communication has stronger implications when accompanied by actively gazing in the direction of the intended recipient of conversation and will more than often result in a response (Kendon, 1967; Heath and Luff, 1992). Body position and orientation may also be used by the speaker to include or exclude people from the conversation, for example by orientating one's body away from someone we signal that we do not wish to include them in the conversation (Goodwin, 1981).

Kendon and Ferber (1973) noted that greetings tended to follow a ritualised sequence of events. Upon sighting another person a gesture (for example a wave) or a smile may be made by one or both parties. This sighting leads to an approach, during which eye gaze is averted and grooming (for example brushing hair with hand) may occur. In the final approach there is mutual gaze and smiling, and the individual may present their hand (for a handshake) or open their arms (for an embrace). ' This is followed by stereotyped utterances (for example 'hello' or 'how are you?') and some form of bodily contact (for example a handshake or an embrace). Once contact is made, less stereotyped conversation ensues (for example, enquiring after recent activities or purpose of visit). Therefore, both verbal and non-verbal processes play important roles in making contact with others.

1.4.3B Turn-taking and interaction management

Gesture can serve to co-ordinate turn taking transitions, and hence serve to co-ordinate the communication process. The continuation of a transmitter's hand movement acts as a cue that the current speaker wishes to hold the conversational floor (Duncan and Fiske, 1977). Similarly, the termination of gesture acts as a signal that the transmitter is ready to hand over the conversational floor, and is therefore referred to as a 'turn-yielding' cue (Kendon, 1967a). It is not only the speaker who can use gesture to regulate turns in conversation

however, the listener may also use gestures to signal that they want to say something, for example raising a hand in the air signals a desire to interject (Goodwin, 1981). Facial expressions are also linked to speech. The eyebrows, for example act as an important conversational signal, portraying a lack of understanding and disbelief (Ekman, 1979). Furthermore, facial emblems may be used as questioning signals. In this way conversational participants can register a turn non-verbally and use facial expressions to help define the parameters of the conversation (i.e, a facial emblem used as questioning behaviour will indicate to the speaker that the listener wishes to have a point clarified).

1.4.3C Feedback cues: Monitoring understanding and attention

Facial expressions are also a good indicator of the listeners' level of understanding. There are two methods by which the face provides feedback. Using head nods listeners can provide the speaker with concurrent feedback about what has just been said (Whittaker and O'Conaill, 1997). By modifying the normal frequency and duration of head nods, Birdwhistell (1970) found that the speakers' ability to communicate was disrupted. Listeners' facial expressions also reveal interest, puzzlement, or disbelief about what they are being told (Ekman and Friesen, 1975).

Research indicates that there are differences between non-verbal and explicit verbal feedback with regards to how readily the information is picked up, specifically with younger individuals. Glucksberg and Krauss (1967) investigated the effects of listener feedback with different age groups (kindergarten, first, second, third, and fifth grade children and adults). The researchers found that all age groups offered verbal responses to listener feedback (explicit requests for additional information), therefore indicating that the feedback had been understood. The younger children, however, were more likely to

merely repeat the original message, whereas the older children and adults were more likely to rephrase or offer new information. In this case all groups readily picked up explicit requests for reformulations. In contrast, Peterson, Danner and Flavell (1972) found that listener feedback through facial expressions was ineffective for eliciting reformulations in 4 and 7-year old participants. In this case, the non-verbal requests for information were either ignored or not picked up. In general, research indicates that younger children have difficulty picking up non-verbal feedback, however adults do use the visual channel for this feedback function. Even so, adults still find it much easier to ignore facial expressions than explicit verbal requests (Peterson et al, 1972).

1.4.3D Interpersonal information

Gestures supply information about a person's knowledge of a task in addition to their verbal communication of knowledge, and therefore provide interpersonal information. Church and Goldin-Meadow (1986) looked at children's reasoning about a series of Piagetian conservation tasks. All of the children gestured spontaneously, and their gestures often portrayed specific aspects of the task they were completing. For instance, in a task that concerned the conservation of liquid, a c-shaped hand was indicative of the dish's width, whereas a fist hand arced from the glass to the dish indicated that water had been transferred from one receptacle to the other. Church and Goldin-Meadow noticed that the children's gestures sometimes matched the information conveyed in their speech, and sometimes the gestures did not convey the same information as the speech accompanying those gestures. They also noted that there was a differing degree to which the children made speech-gesture mismatches, with some children making none at all, and others making six out of a possible six. The researchers noted that the children who produced more gesture-speech mismatches were more likely to benefit from instruction

than those who made fewer mismatches. This evidence seems to support the theory that gestures can provide some information about a child's knowledge of a task, and thus gives an insight into his or her receptivity for instruction on that task.

Gestures are not the only form of non-verbal signal that can give us an insight into a person's cognitive state. For example, Machida (1986) demonstrated that first graders exhibit less direct eye contact with the speaker, more head tilting, excessive hand movements, and agitated body movements in response to verbal material that was difficult to understand, as opposed to more elementary material.

We can infer the affective and cognitive state of a person from the expressions portrayed through the eyes, mouth and eyebrows. Like gaze and gesture, facial expressions also play an active role in co-ordinating content: they provide listener feedback and serve as general indicators of emotional state (Whittaker and O'Conaill, 1997). Ekman and Friesen (1975) noted six distinct facial expressions across many different cultures, although there could be cultural differences. These are happiness, surprise, disgust, sadness, anger and fear. Therefore faces are important cues as to the emotional and attitudinal states of speakers and listeners. We might therefore predict that access to such cues will be important when the emotional/social demands of a communication task are high. This will be discussed in further detail in this chapter and the following chapter.

Posture is less dynamic than other visible behaviours with variations occurring less frequently. Posture can signal the degree of interest or engagement of a conversational participant. Listeners who are interested in what the speaker has to say are more likely to lean forward. Speakers can also signal interest by leaning forward and drawing back their

legs. In contrast, boredom is signalled by head lowering, and turning the head to one side, supporting the head on one hand and leaning back and stretching out one's legs (Bull, 1978). Posture can therefore provide feedback to the speaker about how the message is being received. Bull (1983) has also noted that mirror imaging of body orientation may signal similarity in views and rapport. This may be important in emotionally laden exercises such as negotiation.

1.4.3E Supporting deixis

According to McNeil and Levy (1982), verbal deictic terms (for example 'he,' 'she,' 'it') are often accompanied by physical pointing and gesture to express a speaker's meaning (for example 'he was late this morning' accompanied by pointing at appropriate person). According to Daly-Jones et al (1998) this type of behaviour indicates that deixis is most effective when the 'talk' can be clearly separated from what is talked about. Pointing gestures can also be used to manipulate or direct a person's attention, for example through accompanying pointing with speech, such as 'look at that' or by pointing alone (Goodwin, 1981), and therefore can be used to achieve reference.

1.5: Sharing a visible environment: Information about shared events, objects and people

In face-to-face conversation the fact that conversational participants have access to a shared physical environment means that other types of visible information (other than non-verbal communication) are also available: for example information about physical objects, events and other people (Whittaker and O'Conaill, 1997). For the purpose of co-ordinating content, sharing the same physical environment as another will enable one to make inferences about the set of objects and events that others in the same environment are likely to know about and want to talk about (Clark and Marshall, 1981). Therefore,

listeners can use information about the shared environment to reduce ambiguity, by for example making judgements about what incoming messages are likely to be about, based on this shared information. Likewise speakers can make inferences about what their audiences might expect them to converse or know about based on this shared physical information. Finally, people can make judgements about an individual's availability based upon visual information. In other words people can make inferences concerning a person's relationship with the physical environment, for example are they reading a book, or busy at a workstation? This availability information helps the process of initiating and terminating conversations. (Whittaker and O'Conaill, 1997).

Information about the visible environment often interacts in important ways with verbal and visible behaviours. For example, participants gesture at, orient themselves toward, and manipulate aspects of their environment (Whittaker and O'Conaill, 1997). According to Nardi and colleagues, (1993) having access to a shared environment is crucial when it comes to joint manipulation or modification of complex objects. In essence the participants need to be in an environment containing these objects to help co-ordinate content (Nardi, Schwartz, Kuchinsky, Leichner, Whittaker & Sclabassi 1993). Changes to the environment can be mutually observed as they happen. People can exchange or mutually orient to objects in their shared environment. This is particularly evident with regards to the use of documents in workplace interactions. Nardi et al (1996) have noted that documents relating to ongoing communications can be left on desktops, and are particularly useful as they serve as reminders about tasks that would normally not be completed in one sitting. Nardi et al (1996) have also noted similar 'context holding' functions in disconnected interactions around other types of shared objects (Whittaker and O'Conaill, 1997). Having access to a shared space would seem to be beneficial then. It

should be noted, however, that such a shared space is not available in remote communications. It is for this reason that video-mediated systems are often accompanied by virtual shared spaces (for example a shared whiteboard). The effects of such virtual shared spaces will be discussed in further detail in the next chapter.

Shared artefacts may be constructed during conversation, for example a to-do list written on a whiteboard. Alternatively, artefacts may be imported from some other source, for example participants may refer to a printed document during conversation. McCarthy and Monk (1994) argue that the ability to point to a shared artefact accounts for the extreme efficiency of many utterances in everyday communication. Shared visual 'data' has obvious communicative benefits, especially considering that access to shared objects allows for new communication to take place. A study by Tang (1991) demonstrated this role of physical artefacts. In this study, Tang analysed design meetings in order to identify the function of various activities. Tang (1991) noted that activities such as listing, drawing and gesturing occurred frequently in conversation. Such activities served to express ideas and to mediate interaction. Tang also noticed that participants pointed at specific objects on a drawing surface in order to clarify a particular comment. Whittaker, Geelhoed and Robinson (1993) have also shown similar functions for shared electronic workspaces. In this view Hutchins and Klausen (1991) also suggest that conversational participants in collaborative groups can use objects to facilitate 'distributed cognition.' In other words, participants can externalise their thoughts, through the use of visual aids for example, so that group members can attain a better understanding of the message that the individual is attempting to get across.

1.6: Summary: Verbal and non-verbal communication

Communication is a dynamic process that relies on visual and auditory information as well as a co-present environment. There is a wealth of evidence to suggest that visual signals are important in human communication. Clearly, visual signals are not entirely necessary, otherwise we could not communicate over the telephone and people without access to visual signals, for example those who are visually impaired, would be unable to communicate successfully. The resounding view, however, is that visual information facilitates communication, allowing it to be a more successful and a more satisfactory process. Furthermore, evidence from the literature would indicate that a restriction of non-verbal information would result in less satisfactory and less successful communication. Non-verbal communication supplies emotional and propositional information, and is particularly important in providing feedback to a conversational partner. Non-verbal cues are also used to help achieve reference and supply interpersonal information. Furthermore, non-verbal communication plays an active role in interaction management, for example turn-taking. Given the variety of functions served by non-verbal communication, one would predict a range of effects of removing or restricting access to them (for example communication via the telephone or video-mediated communication) that would depend on the communication task in hand. This will be described in further detail in this chapter and the next chapter. Verbal behaviours have been noted to play important roles in turn-taking and providing feedback. Verbal behaviour also supplies emotional and propositional information, independently and in conjunction with non-verbal signals. It is therefore predicted that any restricted form of communication, for example when access to either visual or verbal cues is restricted, will result in less successful communication. In order to test this prediction a review will take place measuring the success of communication that takes place with a restriction of cues.

1.7: What is successful communication?

Now that we have an understanding of the channels of communication, it is important to define successful and unsuccessful communication. Successful communication can be defined in many ways. O'Connell, Kowal, and Kaltenbacher (1990) offered one definition of successful communication. They argued that the criterion for success in any conversation is the accomplishment of the intention of the communication entertained by two or more interlocutors.

There are many ways we can measure fulfilment: task performance (how well is the end goal of the interaction carried out?), satisfaction levels (for example, preference ratings), the surface structure of conversation, and non-verbal responses. The different ways of measuring success reflect different levels of communication and many different levels of effect of restricting access to visual cues. Measures of task performance may include time to completion, overall success on a particular task, or the establishment of mutual understanding. This approach has been taken by researchers such as Chapanis, Ochsman, Parrish & Weeks (1972), Davies (1971) and Olson, Olson and Meader (1995), who have examined the differences in task performance on problem solving tasks between face-to-face communication and audio-only communication. Researchers such as Olson et al (1995), Sellen (1995), and Abbott, Dallat, Livingston and Robinson (1993) have investigated user preferences, assessing the quality of interactions subjectively through the use of qualitative methods such as questionnaires and interviews. Researchers such as Cook and Lalljee (1972) and Rosenfield (1987) have looked at the surface structure of communication. Such measures have included: the number of pauses in conversation, the number of interruptions (over-lapping speech), and the number and length of turns in conversation. Changes in non-verbal style have been assessed by researchers such as

Graham and Argyle (1975), and has specifically focused on levels of gaze, gesture and facial expressions. These different measures of success will be discussed in further detail later in this chapter with regards to experimental research concerning the differences between face-to-face and audio-only communication. These same measures will also be discussed in further detail later in the thesis with regards to the differences in communication between video-mediated communication and other communication media (audio-only and video-mediated communication).

The type of task is crucial when it comes to measuring successful communication. The characteristics of a task may necessitate different measures of outcome. For example, tasks with a more social component (for example conflict resolution, bargaining or negotiation) are likely to be deemed more successful if participants' evaluations are positive (for example, Fish et al 1993; Short et al., 1976). On the other hand, comparisons made between media using problem-solving tasks have resulted in few differences (for example, Chapanis, 1975; Ochsman and Chapanis, 1974; Short, Williams, & Christie, 1976; Williams, 1977), unless there is a consideration of subjective feelings and attitudes of the participants. Therefore a more holistic approach to measuring success is needed. Monk et al (1996) have argued that approaches that integrate many methods are particularly informative when comparing different communication media. This holistic approach was undertaken by Olson, Olson, Carter and Storrosten (1992), and Strauss and McGrath (1994). Such research will be discussed in further detail in the next chapter.

1.8: Experimental and theoretical evidence for the differences between face-to-face and audio only communication

1.8.1 Introduction

As can be seen both auditory and visual information are important aspects of human communication. Another area of research that leads to the same conclusion is that comparing communication media that differ in access to visual cues. One extreme comparison is that between face-to-face and audio-only interaction. If non-verbal communication were important to human communication then one would predict that communication that is audio-only would be less successful than face-to-face communication. In order to test this hypothesis, experimental data for differences in task performance and the communicative process between audio only and face-to-face communication will be considered. The empirical findings of these studies can be related to a number of theoretical frameworks: namely Argyle and Dean's (1965) 'equilibrium theory,' the 'non-verbal hypothesis,' and 'media richness theory,' (based on the work of Daft and Lengel, 1986).

1.8.2 Experimental evidence

Many psychologists suggest that visual signals are not important in communication. Rime (1982) for example conducted experiments in which subjects had to complete tasks either facing another person or deprived of visual information. He noted that the differences in communication for many variables, including non-verbal behaviours such as head movements and eyebrow movements were not striking. Rime suggested that most studies in this field have overestimated the significance of visual signals. However, as discussed below, a number of researchers also indicate that non-verbal behaviours add value to communication.

An abundance of studies have investigated the function and importance of gestures in human communication. Many studies imply that gestures allow humans to communicate

more successfully. If participants find it more difficult to communicate or are less successful at completing a task when they are denied the ability to gesture, we can assume that gestures play a crucial role in the communication process. This hypothesis was tested by Graham and Argyle (1975), who studied the effects of task performance when subjects were denied the ability to gesture. In this study, the information senders were allocated the task of describing geometric figures with the intention of summarising them so a recipient could make an accurate representation. Half of the subjects were permitted to use gestures, whereas the other half were prevented from using gestures through having their hands tied behind their backs. Successful communication in this experiment was measured through the accuracy of the drawings made of the figures by the information receivers. Graham and Argyle noted that subjects who had access to gestures produced better likenesses of the geometric figures, and therefore it is suggested that gestures make communication more successful. What must be considered here, however, is that in order to prevent gesturing from taking place the experimenters tied half of the subjects' hands behind their backs. Although it is clear that gesturing is important, preventing someone from using hand and arm movements may have more detrimental effects than merely denying participants the ability to gesture. All forms of communication are intricately linked and therefore being physically unable to gesture may have knock on effects for verbal communication. Indeed McNeill (1985) proposes an integral link between speech and gesture and that these serve communicative processes. We are all aware that even when talking on the telephone we use gestures. Perhaps it is the case that for some descriptions it is necessary to gesture even when gestures cannot be seen. An improved experimental method could have had the gesturing out of view of the information receivers, for example a head-only view, compared with a full-body view.

Other studies have been conducted which compare task performance between face-to-face and audio-only interactions. A study conducted by Boyle, Anderson and Newlands (1994) used the collaborative problem solving 'Map Task' (designed by Brown, Anderson, Yule and Shillcock, 1984) in order to test the effectiveness of visual signals in communication. In this task subjects were asked to describe an existing route to another subject who must plot this route on his or her copy of the map. It may be noted that although both maps have most landmarks in common there are also differing markers on each map. Half of the participants were prevented from conversing face-to-face and had a screen placed in between them, whereas the other half conversed face-to-face. To measure task success the experimenters made comparisons between how well the routes had been completed in each condition. In face-to-face dialogues speakers used 28% fewer turns and 20% fewer words than in the audio-only condition, but still, however, achieved equally good levels of task performance. There were also fewer interruptions in the face-to-face condition (8.7%) compared to the audio-only condition (12%), and therefore communication was also smoother. This study seems to indicate that visual signals can supplement verbally communicated information and can assist in the turn-taking process. Ellis and Beattie conducted a similar study in 1986. In comparing audio-only versus face-to-face communication, they also found differences in features of the dialogue associated with turn-taking, for example pauses, overlapping speech, interruptions and back channel responses. The results here are in agreement with Kendon (1967), in that visual cues have an impact on the turn-taking process. So while people manage to interact in audio-only conditions, their style of communication changes.

Doherty-Sneddon et al (1997) report further analysis of the Boyle et al corpus using the Conversational Games Analysis dialogue coding scheme. This coding scheme was

developed by Kowtko, Isard and Doherty-Sneddon (1991), and is a spoken discourse coding system. Within this system, there are two levels of analysis: referred to as Moves and Games. Individual Moves are grouped into single dialogue units called Games. A Game can be defined simply by the goal that they serve within the interaction itself. For example, an INSTRUCT Game conveys a request for instruction or action (e.g. "move up"). In an analysis of face-to-face dialogues compared with audio-only interactions Doherty-Sneddon et al (1997) found that face-to-face conversants check less often that their listener understands them (ALIGN) or that they have understood their partner (CHECK). In audio-only conditions, these games occurred 50% (ALIGN) and 28% (CHECK) more often. Where visual signals are not available speakers do more verbal checking, whereas in face-to-face conversations non-verbal signals may be substituted. Support for this data comes from pilot analyses of gaze behaviour. From a detailed analysis of a sample of 8 face-to-face dialogues, they found that gaze often occurred in face-to-face interactions at the same points in dialogue structure as additional verbal checking occurred in audio-only dialogues (Doherty-Sneddon et al, 1997). So, at certain points in a dialogue speakers may be unsure their listener has understood. In face-to-face communication participants can glance at their listener, whereas in audio-only communication this is not possible, so verbal checking is necessary. This indicates that visual cues are important in grounding mutual knowledge.

Research by Anderson et al (1997) suggests that speakers adjust their speech to allow for the loss of visual cues. Tapes of words extracted from the Boyle et al (1994) corpus were played back to new subjects who were more accurate in recognising isolated words from unseen dialogues than face-to-face conversation. This evidence seems to suggest that people alter their speech when they know that someone cannot see them in order that their

speech is more intelligible. Although non-verbal information may be important, evidence from Anderson et al (1997) would suggest that people can compensate when such cues are restricted. Furthermore, the type of communication task is likely to influence the particular impact that visual cues have. For example tasks that involve more visual descriptions may be facilitated by visual information (for example the description of geometric figures or Brown et al's Map Task).

Chapanis, Ochsman, Parrish and Weeks (1972) compared task performance between four different mediums: (1) typewriting (experienced and inexperienced) (2) handwriting (3) voice and (4) face-to-face communication. Tasks were formulated in such a manner that solutions required the efforts of both individuals working together as a team. This was done by deliberately structuring the problems so that each member of a team received complementary information folios. One member of the team, the seeker, was given a problem for which he had to find the solution. His/her information folio consisted of certain parts of the problem. Another member of the team, the source, had the folio with the remainder of the information needed to solve the problem. Therefore, while neither person could solve the problem by himself, the two of them had all the information necessary to do so. The tasks used in this experiment were referred to as the geographic orientation problem and the equipment assembly problem (Chapanis et, 1972).

The results from this research are as follows. The two oral modes (communication-rich and voice) differed significantly from the two hard-copy modes (handwriting and typewriting). Neither the difference between the communication-rich and the voice mode, nor that between handwriting and typewriting modes was statistically significant. Chapanis et al (1972) also found no difference between problems and between problems

and modes. The average subject in the communication rich and voice modes were able to reach a solution in about 16 minutes of active communication. By contrast the average experienced typist needed 32 minutes and the inexperienced typist 43 minutes of active communication to solve the same problem.

Although task performance was the same for voice-only and face-to-face, the interaction between the participants was markedly different. In the map style task, face-to-face participants spent a significantly longer time than voice mode cross-searching. In other words the seeker spent more time in the face-to-face mode referring to the source's list of names, and the source referring to seeker's map and index of street. Also joint searching took place, so large amounts of time were spent referring to the maps. In the voice condition there were increased amounts of time spent in description, explanation and checking, however these participants still completed the task in the same amount time. Therefore, the style of communication differed between the groups, once more indicating that visual signals are important in grounding mutual knowledge and allow the participants to offer feedback visually.

A study by Morley and Stephenson (1969) looked at an experimental situation concerned with the role-playing of an industrial wage dispute, in which one side was considered to have a weaker stance than the other. Results showed that the weaker side were more successful when completing the negotiation over face-to-face as opposed to audio-only. This was predicted because during face-to-face interaction the interaction was more focused on interpersonal considerations than on the objective facts of the debate. Weakness was based upon the number of arguments available to each side. However this does not take into account overwhelming arguments. Evidence from Morley and

Stephenson (1969) and Chapanis et al (1972) therefore indicates that, as well as affecting communication in problem-solving tasks, visual cues also have an impact on socially orientated tasks. Further evidence for this conclusion comes from Williams (1977) and Whittaker (1995) who indicate that a view of the face is important for bargaining.

1.8.3 The non-verbal hypothesis

Much of the research already discussed contributes to the 'non-verbal hypothesis' (Whittaker, 1995; 1996), in that visual signals (for example gaze, gesture, facial expressions and posture) provide information that is absent from audio-only communication. Therefore communication that has visual and audio elements will be more successful than audio-only communication (Whittaker and O'Conaill, 1997). One method to achieve this is to compare audio-only communication and face-to-face interaction, or to compare natural face-to-face communication with communication that has a restriction of visual cues.

As described above evidence from comparisons of face-to-face and audio-only communication are mixed, however research seems to indicate that although access to visual information isn't crucial for the successful completion of a task, having visible behaviour available alters the communication process (in some cases making it more satisfactory). For example, speakers check less often that their listeners have understood them, indicating that understanding can be signalled through the use of gestures alone (for example head nods), therefore making it unnecessary to signal understanding verbally. However, it may simply be the case that speakers and listeners feel more confident. There is also evidence to suggest that conversation is smoother, with fewer turns and less interruptions.

So, visual cues influence some task outcomes and even when they do not, they still influence surface structure and communication style measures, and therefore visual signals clearly matter. Removing visual signals, even when participants remain co-present, results in significant differences in communication. However, while one characteristic of real-life audio-only interaction (for example telephone conversations) is that participants cannot see each another, another is that they are not co-present. Are effects of remote audio-only communication due solely to the loss of visual cues, or is there a residual effect caused by lack of co-presence?

1.8.4 Argyle and Dean's 'equilibrium theory'

One theory that can help to explain differences between co-present interaction and interaction that takes place at a distance is equilibrium theory. According to Argyle (1988), intimacy is expressed by factors such as proximity, gaze levels, the degree of self-disclosure and a number of other cues for liking and intimacy. Further to this, a study by Argyle and Dean (1965) discovered that in everyday interactions individuals seek a comfortable intimacy distance. This equilibrium is maintained by adjusting the signals of intimacy, for example people will change orientation (leaning forward or leaning back), and will inevitably feel uncomfortable if they cannot. Argyle and Dean go on to propose that this is a result of a balancing of forces to approach and withdraw. Essentially we seek approval in social situations, but fear rejection from others. Their research also indicates that such forces reduce as we get further away from someone, however avoidance forces decline faster than approach forces. For example, if one person is fond of another individual, regardless of whether they find them attractive or enjoy their company, the approach forces will be stronger and the avoidance forces weaker, therefore this will result

in closer proximity (especially if the feelings are mutual). Argyle and Dean also theorise that if a person comes too close then this will arouse stronger avoidance forces than approach forces and hence the individual will back away. Equilibrium distance is signalled by all the channels of non-verbal communication. So, for example, if someone reduces proximity, another person may reduce eye contact in order that a more comfortable intimacy distance is formed. Intimacy may also be increased by other factors, for example personal topics of conversation. Exline, Gray and Schuette (1965) found that when personal topics of conversation were used, individuals compensated for higher levels of intimacy by reducing eye contact.

Further evidence for this theory comes from Peery and Crane (1980), who looked at 4 year old children's reactions to the proximity of an experimenter. The results indicate that when the experimenter approached the child withdrew, however when the experimenter withdrew the child approached in order to maintain a comfortable distance. The results also indicate that there were more withdrawals at a starting distance of 1 foot and more approaches at a starting distance of 3 feet. Patterson (1976) extends the theory by indicating that a move towards greater intimacy leads to an increase in arousal. If this move is interpreted in a pleasant way there will be reciprocity, for example smiles and head nods, however, if this move is interpreted as unpleasant, there will be a withdrawal, hence maintaining equilibrium. Increased distance also results in lower levels of embarrassment and higher levels of eye contact. According to Argyle (1988) reciprocity is common in the early stages of meetings between strangers, but this is only up to a point. The precise point at which an individual compensates will depend entirely on that individual's preferences, state of mind and even the social norms of the situation.

This theory has implications for communication that takes place at a distance. An increase in distance between individuals will result in fewer ways to express intimacy and may therefore have a number of communicative effects, for example increased eye contact and increased likeliness to discuss personal topics of conversation. In other words, other channels will be used to compensate in order to maintain equilibrium and to feel more comfortable. Argyle and Dean (1965) indicate that distributed equilibrium can be compensated for with the use of other signals of intimacy, for example higher levels of gazing. Studies have supported this finding, for example Doherty-Sneddon and McAuley's (2000) child witness research. In this study, the researchers found that when giving evidence concerning a neutral event, children were more resistant to misleading information in the video condition in comparison the face-to-face condition. The researchers go on to explain this finding by stating that these effects are due to a decrease in social co-presence, which ultimately resulted in decreased feelings of intimidation and increased feelings of confidence.

Equilibrium theory is therefore supported by theories of social presence and media richness. Research has indicated that as bandwidth of communication media increases, so do the feelings of 'social presence' (Gale, 1989). When non-verbal cues are removed from communication the result is a lack of social presence (Rice, 1993). Participants are depersonalised and therefore are seen to be less like people and more like objects (Williams, 1977). In turn, this depersonalisation has effects on social and communicative processes. Sproull and Kiesler (1986), for example, note that depersonalisation can result in participants becoming more self centred, and Siegal, Dubovsky, Kiesler and Mcguire (1985) noted that depersonalisation could even lead to antisocial behaviour. Williams (1977) also noted that lack of non-verbal cues could lead to decisions based more on facts

than personalities. Research on co-presence and equilibrium theory therefore suggests that communication that takes place at a distance will result in differences in communication. As has been noted, this could produce benefits as well as drawbacks.

1.8.5 Media richness

Another theory which may be used to explain the differences between face-to-face and audio-only communication is the 'media richness' theory. 'Media richness' theory argues that performance improves as individuals use 'richer' media (Dennis and Kinney, 1998). 'Media richness' is based upon the organizational research of Daft and Lengel (1986), who argued that task performance could be improved through the matching of media characteristics to the information processing tasks that were to be conducted. Daft and Lengel (1986) proposed that equivocality resolution and uncertainty reduction are the two motivating factors influencing information processing within organisations. Equivocality can be defined as the existence of multiple and conflicting views of the same problem or issue. The resolution of equivocality requires these differences to be resolved with the exchange of personal interpretations, and the consequent development of mutual understanding. Uncertainty can be defined as the gulf between information required to perform a task and the information that is currently available to the individual performing the task. Uncertainty reduction involves the acquisition of information to fill this gulf.

Different communication media provide varying capabilities for communicating data. Media that have a high capacity for multiple communication cues and rapid feedback can be seen as richer media. These types of media are best suited for conveying information that resolves equivocality. Lean media on the other hand, are best suited for conveying information that reduces uncertainty. Media richness should be equivalent to the

requirements of the message in order for effective and efficient communication to occur. The multiplicity of cues, the immediacy of feedback, language variety, and the personal focus of the medium all influence 'richness' (Dennis and Kinney, 1998). For example, visual cues are available in face-to-face communication and not in audio-only communication. This implies that face-to-face communication is 'richer,' or has more available cues than audio-only communication. Face-to-face communication should therefore be better equipped to resolve differences, aid in the exchange of subjective interpretations, and aid the development of a shared understanding (Dennis and Kinney, 1998).

Dennis and Kinney (1998) conducted one study in particular, which tested this 'media richness' hypothesis. To test the theory Dennis and Kinney asserted that 'richer' media would result in more successful communication on equivocal tasks as users would be able to communicate more quickly and would be able to understand ambiguous messages more readily. Equivocal tasks allow for multiple or conflicting interpretations from the given information, and therefore presenting a challenge for participants to arrive at one mutual meaning of information. It was also hypothesised that 'leaner' media would result in better performance for less equivocal tasks. To test this hypothesis a task was constructed in which subjects were asked to rank in order (from best to worst) five students seeking attempting to gain admission to the university's business school. Each participant was presented with incomplete information so that the pooling of resources was required in order to reach an informed decision. The lower equivocality task involved a set of four questions, including a mathematical, geometric, physics and logical reasoning question, each of which provided a clear framework for problem resolution. Results indicate that richer media do not necessarily result in improved performance on equivocal tasks. There

were no significant effects for decision quality, consensus and satisfaction of communication. Dennis and Kinney did, however, find a significant correlation between media richness and decision time. It would seem from this study that the new 'lean' media may be just rich enough to enable users to successfully communicate for moderately equivocal tasks (Dennis and Kinney, 1998). These findings are similar to O'Conaill et al's (1993), in that higher quality V.M.C systems are closer to face-to-face communication than lower quality systems (discussed further in the next chapter).

Dennis and Kinney go on to state that "media switching" may be the best method for optimum performance on tasks. In other words choosing which media to employ is task dependent. For example, the researchers indicate that the best media for distributing information may not necessarily be the best for coming together on a decision, and in many cases both are required to perform a task. Another example of this is also given: for person orientated tasks that require some degree of social contact, a medium's complicity of cues or its social presence may be important. However, the researchers also indicate that the use of computer mediated communication for many social tasks, for example dating, would suggest that many individuals can successfully adapt their communication behaviour in order to communicate very private information and even build relationships through media with very low social presence. It should also be noted that many factors beyond media richness and level of social presence, such as personality variables and shared histories would affect task performance.

Valacich et al (1994), also found evidence to support Dennis and Kinney's findings that there were no differences in the decision making quality (i.e., accuracy of decisions) in equivocal tasks when comparing varied media. Valacich et al, however, did find that the

time taken to complete a task was significantly quicker for video (over telephone, face-to-face and computer) on equivocal tasks, whereas telephone was fastest for unequivocal tasks. These results somewhat go in line with media richness theory. So differences in efficiency resulted. One would expect the richest media to be best (face-to-face), however the second 'richest' has been most successful. One would also expect the leanest medium to be best for unequivocal tasks, however the second leanest medium (telephone) has come out best when we would expect the computer system to be worst.

This theory may go some way to explaining the differences that were found by Doherty-Sneddon et al (1997). Doherty-Sneddon et al (1997) found that audio-only conditions resulted in participants having to check more often that the listener understood them (ALIGN) and that they have understood their partner (CHECK). Increased instances of such games may be explained through media richness, as it would seem that these participants had lower levels of shared understanding (or at least perceived lower levels) than those taking part in face-to-face conditions. Another example of 'richness' would include a comparison of immediacy of feedback between face-to-face dialogue and e-mail systems. Feedback would obviously be much quicker when conversing face-to-face as opposed to sending and receiving e-mails and therefore face-to-face interaction is 'richer' and should be more successful.

1.9: Chapter summary

Although humans can adapt reasonably well to communication media that restrict visual cues and take place at a distance, it is still the case that such communication is in many ways less efficient and results in a number of drawbacks, or at least differences. Research suggests that communication differs (arguably more successful in certain measures) when

it takes place in a co-present environment, with access to both visual and verbal information, in comparison to remote communication with access to audio signals only. Although few differences have been noted for overall task success/completion, research has suggested that communication is smoother, more efficient, and more satisfactory (from a participant's point of view) when it takes place co-presently with access to visual and audio signals. Differences between face-to-face co-present interaction and audio-only communication can be explained using the media richness theory, Argyle and Dean's (1965) equilibrium theory and the non-verbal hypothesis. Considering there are a number of communicative drawbacks associated with audio-only communication, are there similar drawbacks associated with video-mediated communication? Video-mediated communication allows users to have access to visual signals, however communication still takes place at a distance and therefore is not in a co-present environment. The next chapter will discuss the evidence for whether video-mediated communication is as successful a medium of communication as face-to-face interaction.

Chapter 2: Literature Review: Research on Video-Mediated Communication

2.1: Introduction and chapter aims

For decades the telephone has granted individuals the ability to communicate from geographically dispersed locations. Recent developments however have resulted in a technologically and financially feasible method for displaying visual information (including a visual image of the interlocutor(s) and in some cases shared visual documents), in addition to an audio channel for communication. Video-mediated communication (V.M.C) therefore offers an alternative method for meeting and conversing with others in remote locations, and presents the possibility of being able to duplicate some of the advantages associated with sharing a co-present environment (Sellen, 1997). Video-mediated communication is regarded by many as a valuable tool for applications such as remote collaboration, conferencing, and distance learning (Finn, 1997). According to Campbell (1998), videoconferencing systems duplicate the experience of face-to-face meetings as closely as possible without the burden of travel. Unlike audioconferencing, video-mediated communication allows participants to see the reactions of their colleagues. According to the 'non-verbal hypothesis' (Whittaker, 1995; 1996), picking up such non-verbal cues should allow users to respond with greater sensitivity and comprehension. For example, one should be able to notice if another participant is smiling or frowning (Campbell, 1998). Furthermore, Isaacs and Tang (1993) indicate that adding a visual channel to remote communication produces many benefits for users. These include allowing listeners to non-verbally signal their understanding or lack thereof to the speaker; speakers can augment verbal material with gestures; responding to speakers can be accomplished in a purely visual sense (for example smiling or nodding); attitudes can be conveyed through posture and expressions; and participants can better interpret the significance of conversational pauses. The validity of these views will be discussed in further detail later in the chapter.

As discussed in the previous chapter, access to visual information (including certain non-verbal behaviours and images of the environment) is beneficial for communication. Face-to-face communication is smoother, more efficient and more satisfactory than audio-only communication, although task outcome tends to be unaffected by access to visual information (with the exception of tasks that require a more social dimension, for example negotiation). The purpose of this chapter is to ascertain whether access to visual information in a remote conferencing environment is also beneficial. In other words, considering the advantages that face-to-face communication has over audio-only communication, is the use of video-mediated communication also associated with such benefits? One may assume that through adding a visual information channel to an audio channel, remote participants are introduced to a communication situation that is closer to a face-to-face co-present conversational context (Sellen, 1997). The validity of this assumption may be questioned through research, which aims to test the effectiveness of video-mediated communication. This can be achieved in two different ways: through comparisons of V.M.C with audio-only communication and comparisons of V.M.C with face-to-face communication. In comparing video to audio, it is noted how and when video information enhances speech-only communication, and in comparing video with face-to-face communication we may note how effectively V.M.C mimics face-to-face conversation (Whittaker and O'Conaill, 1997).

A review of both experimental studies and field studies will take place in this chapter. Experimental studies offer a systematic and controlled way to make comparisons between different media. Field studies give useful insights as to how users adapt to the technology, and whether V.M.C can be used successfully for real-world applications. Furthermore, the

effectiveness of V.M.C can be deliberated with the use of a number of different measures. There are many ways we can measure fulfilment: task performance (how well is the end goal of the interaction carried out), user perceptions (for example, preference ratings), the surface structure of conversation, and non-verbal responses. Differences between media on these measures will be explained using the 'non-verbal hypothesis' (Whittaker, 1995; 1996), Argyle and Dean's (1965) 'equilibrium theory,' and 'media richness theory,' (based on the work of Daft and Lengel, 1986). Furthermore, evidence from such research will be used in order to advise users on good practice when working with video-mediated technologies. Before a discussion can take place regarding differences between media, there will be a consideration of the differences within video-mediated communication and in turn how this effects communication. In other words, considering that there are a number of different systems available (all with varying degrees of quality), can V.M.C be considered a homogenous entity, or do differing systems result in different levels of communicative success?

2.2: Effects of system quality and set-up

2.2.1 Introduction

Two distinct types of video communication are becoming popular: videoconferencing supported by specialist conferencing centres and desktop videoconferencing (Anderson, Smallwood, MacDonald, Mullin, Fleming & O'Malley, 1999). With desktop videoconferencing (DVC), the user adds a camera near to his or her computer monitor and makes use of a telephone line or Local Area Network (LAN) to connect with other users. An advantage of using a personal computer to communicate in this fashion is that other functions, such as sharing files or images, can be completed simultaneously. The introduction of desktop videoconferencing systems has also resulted in cheaper equipment

that is readily available to the average person. Market research suggests that the Internet voice market will be worth $63 billion by the year 2002, compared to $741 million in 1996 (Videotalk, 1998), indicating a greater need for an understanding of the psychological and communicative effects of using such technology.

Specialist conferencing equipment comes in many different forms. For example, specially dedicated rooms can be built in company buildings, which are equipped with many monitors and cameras. "Rollabout" systems are also becoming standard. These systems include an integrated camera and monitor, which are mounted on a cart, and therefore can be moved to whichever office the system is required in. Many companies also make use of "Media Spaces," which are V.M.C systems that are specifically designed to mimic the sensation of being in a meeting room by using multiple monitors positioned around a table. This type of set-up is used in an attempt to form a higher level of co-presence between locations (Angiolillo, Blanchard, Israelski & Mane, 1997).

Telecommunication technologies are continually improving. One change in particular is that glass, as a medium of wired telecommunications is taking over from copper. Copper typically has a limited bandwidth of approximately 5,000-hertz. Considering the fact that human voices require a bandwidth of 10,000-hertz, compression techniques are necessary for communication by telephone. On the other hand, fibre optic wires can have a bandwidth of up to 2,000,000,000-hertz, meaning that the human voice, high-fidelity sound (200,000-hertz) and high-definition television (24,000,000-hertz) can be handled, making the transmission of high-definition, multi-sensory virtual reality possible (Tiffin and Rajasingham, 1995).

Although advancements in communications technology will continue, the aim of this chapter is to achieve a greater understanding of the psychological and communicative effects of using video-mediated technologies. The main aim of this thesis is to ascertain whether certain practices can improve video-mediated communication, and this is irrespective of the type of system used. In addition to the divide between desktop videoconferencing and specialist conferencing centres, individual systems vary with respect to quality and set-up. 'Media richness' is a useful concept in characterising the differences between face-to-face and audio-only communication. It is also useful in differentiating between different qualities and types of video-mediated communication. Media richness can be noted within media and in the case of videoconferencing is affected by technological factors such as the quality of the system and the system set-up. System quality can refer to many different factors that affect a participant's communication. These include technological aspects, such as absence or presence of delay and the quality of visual and audio channels. System set-up may include more individual components to the system; for example some systems are voice activated and only show the current speaker, whereas others reveal all participants on screen at one time. The following sections aim to review some of the technological factors that affect communication in videoconferencing.

2.2.2 Audio quality

Good audio quality is vital for successful videoconferencing (Angiolillo et al, 1997). Even with sophisticated equipment, good audio quality is still reliant on good room acoustics and microphone placement. Poorly placed microphones or bad acoustics can result in unwanted echoes and 'boomy,' unpleasant sounds For successful two-way conversation in which both speakers can talk at once, it is also required that sound absorbing wall coverings and other types of acoustical treatment be utilized (Angiolillo et al, 1997).

Research has indicated that high quality audio signals can have a number of psychological and communicative effects. Watson and Sasse (1996), for example noted that when the quality of an audio signal was raised, participants judged the video signal from a high definition monitor as being better. Hollier and Voelcker (1997) also noted this effect and in particular found that quality ratings were even more distinct when the image was of a person speaking.

2.2.3 Delay

One of the major technical limitations associated with video-mediated communication is delay. In other words, signals (audio and visual) from one location do not reach the other location instantaneously, but are be delayed by a period of time. Visual and audio signals may also be processed at different speeds. This results in the signals being received at different times, meaning that video and audio are not synchronised. Most vendors attempt to minimise delay between image and sound however, by introducing a similar delay in the audio channel and thus maintaining lip synchronisation. This technique is not always successful, however, as moment to moment changes in signal will result in varied lengths of delay, making it difficult to compensate. Large amounts of change from one frame to the next, for example during movement, will require more computation. A complete disregard for lip synchronisation has a negative impact on communication (Angiolillo et al, 1997), however, some researchers have argued that attempts to synchronise sound and image should be abandoned, as users are more frustrated with the audio delay than with the absence of lip synchronisation (Tangs and Isaacs, 1992).

Until recently it was assumed that due to the size of files required to represent digitised audio and video, shared networks could not be used to transmit real time audio and video.

In other words, it was believed that the delay would be far too great for desktop videoconferencing to be practical. The introduction of high speed networks, for example "SuperJanet" in the U.K, and the development of compression techniques, however means that conferencing is now becoming more widely used over the Internet (Sasse, 1996). Compression techniques reduce the number of frames per second and the amount of information in each frame of video, resulting in speedier transmission. The result of these techniques, however, is low picture definition and jerky pictures. In essence videoconferencing can be accomplished using the public telephone system, but the price to pay for cheaper communications is low quality images. Research suggests that reducing the number of frames per second (video refresh rates) may also have effects on communication. In a review of over 30 studies Barber and Laws (1994) noted that video frame rate has a number of communicative effects. Barber and Laws noted that if video is presented at 12.5 frames a second, participants' ability to perform tasks such as lip-reading or repeating back a message, is significantly poorer than if video is presented at 25 frames per second.

Bandwidth can generally be defined as "a measure of the speed of the connection" (Videotalk, 1998). With desktop videoconferencing problems with bandwidth may ensue, as the Internet itself can be unpredictable with regards to transmission performance. Heavy traffic and internal transmission problems can often result in delays that are not easily controlled, which in turn may effect a disruption in speech and video at the destination computer (Videotalk, 1998). Although Internet conferencing cannot guarantee real time audio and video, rapid development in technology will ensure improvements in bandwidth and computer power and speed. Specialist conferencing centres are more likely

to guarantee real time audio and video or at least a higher connection speed, accredited to the fact that such conferencing centres use dedicated lines.

The length of delay varies between systems. The public analog voice telephone network in the United States typically has a delay of 20 to 30 milliseconds, compared to 10 milliseconds with ISDN (Integrated Services Digital Network). Delay in V.M.C can be caused by compression techniques. All compression techniques require processing time in order to complete tasks. Delay will therefore vary from one system to another (dependent on the type of algorithm used) and within a system, dependent on the amount of information that needs to be compressed at any one time (Angiolillo et al, 1997).

Delay has been noted to have a number of psychological and communicative effects. People who experience delay often are unaware of it and experience it generally as confusion. In many cases they may attribute communicative problems actually caused by delay to the other member of the exchange (Kitawaki, Kurita & Itoh, 1991). In essence, each speaker will see themselves a being quick to respond, but will view the other speaker as being slow to respond (Kitawaki et al, 1991). As discussed in the previous chapter, co-ordinating conversation is reliant on refined timing. Research has indicated that speakers time their contributions to fall within a fraction of a second after the current speaker's turn comes to an end, in some cases causing an overlap. Even a slight delay will have detrimental effects on this process. Angiolillo et al (1997) illustrate this point quite effectively by using an example of a delay of 300 milliseconds in either direction. In this case, if person A expects a quick response to his/her previous statement, person B will hear the statement 300 milliseconds later than when the statement finished and if they respond quickly, person A will hear the response a full 600 milliseconds after he/she had finished

their original statement. The result of this will be that person A will more than likely believe that person B has not understood what has been said as the response was not as quick as anticipated. Person A may then decide to reiterate or formulate another statement, which will in turn overlap with person B's original response. According to Angiolillo et al (1997), at this point it will be unclear as to who holds the floor and a breakdown in communication will occur. In such scenarios it has been noted that people adapt by resorting to a formal manner of communication with longer turns and fewer interruptions.

Experimental evidence for the effects of delay on communication comes from O'Malley et al (1996), who compared task performance on the 'Map Task' (Brown et al, 1984) with video-mediated conditions with delay and video-mediated conditions without delay. The conditions included: an audio-only link delay condition, an audio-only with no delay condition, a high bandwidth video-mediated condition with eye contact and no delay, and a video-mediated condition with delay. Delay was accomplished using "off–the-shelf videophones" with an analog signal that had a delay of approximately 500 milliseconds. Results indicate that there was no significant difference in length, and no significant difference in task performance dependent upon access to visual signals. Poorer performance was noted however with respect to presence of delay. Participants had 36% poorer performance for the audio delay condition in comparison to the no delay in audio condition. Delay in video condition was 40% worse than telephone users with no delay and not significantly different to the audio delay condition. When completing the task with a delay in video, over 50% of a speaker's turns were interrupted. In the audio delay condition over 40% of speaker turns were interrupted, compared to only 15% in other conditions. Delay in video has an adverse effect on communication and task performance

and it is concluded that delay should be avoided in the design of remote communication systems (O'Malley et al, 1996).

2.2.4 Image size and quality

The size of the monitor/screen is also an important technical consideration. Primarily videoconferencing is used to observe orators, but it is clear that as the number of people on the screen increases, the image becomes less clear. According to Tiffin and Rajasingham (1995) videoconferencing systems are at their best when they utilize the full screen to show the person who is talking close-up, therefore taking advantage of any non-verbal signals that need to be transmitted. Small screens are advantageous in that they are portable, and they make lower resolution images look better (Angiolillo et al, 1997). On the whole, however, users show a preference for large displays up to the point where the image is life size (Inoue, Yoroizawa & Okubo, 1984). According to Tiffin and Rajasingham (1995), specialist conferencing centres are more likely to employ larger screens, and therefore the subtleties of non-verbal behaviour are more likely to be picked up. Angiolillo et al (1997) credit the preference for large screens to the need for social co-presence, with a large screen providing an instinctive feeling that the remote participants are physically present in the room. This theory is borne out by Prussog, Muhlbach and Bocker (1994), who noted that users' subjective impressions of actually being in the same room were vastly improved when a large life-like screen was used, as opposed to a smaller screen. This sentiment is furthered by De Greef and Ijsselsteijn (2001), who indicate that the addition of broadband, life-sized video communication results in significantly higher ratings of social co-presence. Angiolillo et al also consider the possibility of using cinematic or virtual reality techniques for videoconferencing, as such techniques cover a viewer's peripheral vision and therefore provide a more intense, 'virtually real' experience

(Pausch, 1993). Even though there appears to be some advantage to using larger screens, video display sophistication must have limits due to cost and bandwidth. According to Angiolillo et al (1997) a trade off is required to be practical, but also to preserve what may be the advantages of having video in the first place.

The size of the monitor is not the only important consideration for good image quality. Conferencing rooms should also be chosen with care. The physical layout of a room will determine how monitors are placed. Rooms that are poorly planned will result in less effective use of the technology: poorly planned wall space for example will limit viewing. The type of room chosen for a videoconference may also have an effect on lighting. In order for video cameras to work effectively, for example to produce minimal shadows, professional direct and indirect lighting must be available. The fact is, a usual office environment is unsatisfactory and will often result in dark and blurred images of the participants (Angiolillo et al, 1997).

2.2.5 System set-up

As well as being differences between systems with regards to 'quality' (for example speed of transmission or image quality), systems also differ with respect to how they are set up. Systems differ in the manner in which orators are observed. For example, some systems employ the use of voice-activated cameras, which automatically focus on the person talking, whereas other systems use pre-set camera positions, which can be manipulated in order to get a shot of the speaker. Further to this, differences in set-up may also include whether the monitor is used to portray an image of the conversational partner (Telepresence) or of relevant work objects (Teledata). Many systems also make use of a 'whiteboard,' which allows participants to draw and type in the same workspace

simultaneously. The effects of these differing set-ups are discussed in the following sections.

2.2.5A Experimental research on different methods for viewing conversational partners in V.M.C

A study conducted by Sellen (1995) compared three different systems; all of which had high quality video, but different approaches to observing participants. Picture-in-a-picture (PIP) displayed simultaneous pictures of all four remote participants on one colour monitor with a single speaker that broadcasted all speakers' voices. The Hydra set-up had each speaker facing three screens, with a camera mounted below each monitor and a speaker mounted below each camera. Finally the LiveWire system had a voice activated switching mode, so that the current speaker was displayed to all participants, except for the current speaker who was shown the image of the last speaker. To measure communicative success for all three systems, Sellen (1995) asked all participants to fill in questionnaires relating to their experience with video communication. A review of these questionnaires revealed that two thirds of the subjects preferred using the Hydra set-up, listing reasons such as Hydra supporting selective gaze and listening and simultaneous conversations, which are more indicative of real life conversation. One third of the subjects preferred PIP, stating that they did not have to turn their heads and also they could see an image of themselves on the monitor. So, most people preferred the system that mimicked face-to-face meetings more closely by providing a spatial layout similar to those used in co-present meetings.

2.2.5B Video data vs. video links

Another aspect of video-mediated systems that has an effect on communication relates to whether images of the interlocutors or images of relevant objects are displayed. Anderson et al (1999), for example, investigated user preferences for video links (images of people)

and video data (images of relevant objects), and also the impact of the allocation of bandwidth on different sorts of video data. The experimenters designed two simulations, (i) 'The Travel Service Simulation,' where participants plan a holiday itinerary and (ii) 'The Financial Service Simulation,' where participants choose a property and arrange an appropriate mortgage. Participants were further distinguished by what images were available to them on the monitor. In both scenarios, participants in the Telepresence (video links) condition saw a picture of the sales representative. In Teledata condition (video data) participants viewed a variety of images, for example video clips of the holiday destination, a map, an itinerary, video clips of the houses, descriptions of the houses and a map of the area.

Participants in both studies were equally satisfied with the outcomes of the tasks (97% and 98%) and with many aspects of the communication process. Subjects in both studies tended to rank Teledata high in terms of what was most useful, what the most important feature to preserve, and what was the most important feature to improve (Anderson et al, 1999). Although a view of the face is said to be important for bargaining (Short et al, 1976, Williams, 1977, Whittaker, 1995), the study showed no preference for Telepresence on the Financial Services Simulation (which required a high level of negotiation for task success). Richer media are also meant to increase social presence, however no support was found for this in the results, as there was no difference between different frame rates (Telepresence conditions were set at 25 or 12 frames per second and Teledata conditions were set at 12 or 6 frames per second). These findings add further empirical evidence to the arguments made by researchers such as Nardi et al (1993). The ability to relay video data may be a more useful aspect of multimedia communication systems than the ability to relay images of the face. So high bandwidth is not always particularly useful. The importance of

having a communicative rich medium may therefore depend on what participants are trying to use the equipment for.

In this view, Anderson et al (1999) suggest that the nature of the task to be accomplished will have an important impact on the usefulness of different types of multimedia data. For example in telemedical applications where operations or other medical procedures are relayed to remote participants for consultation and training, it is expected that a view of the patient or procedure (Teledata) is likely to be more useful than a view of the medical personnel conducting the operation or procedure. Research by Whittaker and O'Conaill (1997), for example indicates that views of participants in remote surgery allowed for four different ways of communicating to develop. Firstly, the dynamic image of the surgeon's actions allowed detailed co-ordination of interleaved physical action between the assisting nurse and the surgeon in the operating theatre. A second communicative function of the video image was that it served to disambiguate other types of surgical data that were supplied to remote consultants, such as neurophysiological monitoring data. Third, the video image served as a physical embodiment of progress through the operation. Members of the team who were involved in multiple operations at different locations and also those within the operating theatre could see the current stage of the procedure the surgeon was at. Finally, the image was used for learning and education, for example some surgeons also recorded these procedures, to use them as aids in teaching classes.

It is also clear that other tasks such as concurrent engineering or training may also be more reliant on Teledata (Egido, 1990, Nardi et al, 1993). On the other hand, purely social interactions among remote family members will benefit more from a view of the participants' faces rather than views of other images (Anderson et al, 1999). In accordance

with this hypothesis, it would appear that participants spend significantly greater amounts of time looking at relevant work objects than looking at other people (Argyle and Graham, 1977). Usual videoconferencing set-ups, however concentrate greatly on the image of the other conversant. Therefore the video image can be used to transmit real-time information about work objects as well as transmitting images of conversants, and this can then be used to co-ordinate conversational content among distributed teams, by creating a shared physical context. So, Teledata is preferable when the task involves manipulation of an object, which if done face-to-face would draw most gaze to that object. Indeed, Argyle (1988) has indicated that person gaze lowers significantly in the presence of an object of legitimate attention, and can be as low as 3% to 7% of conversational time. Compare this with face-to-face communication in the absence of an object of legitimate authority and the speaker spends approximately 40% of time looking at their partner's face, and the listener spends approximately 75% of time looking at their partner's face (Argyle, 1988).

Further evidence for the desirability of Teledata comes from Gaver, Sellen, Heath and Luff (1993). The researchers conducted a study in which they offered users a choice of different video views, in this case either of work objects or collaborators. Results indicate that most users preferred the work objects views, which were chosen significantly more often than views of collaborators. Gaver et al (1993) suggest that the use of video to provide face-to-face contact has been overemphasised compared to other ways that video can support remote collaboration. The problem with using video to support images of other participants is that a 'porthole effect' is often very noticeable, in other words users become much more focused on the individual's face then they would in a face-to-face interaction. Obviously in a face-to-face interaction there are other elements to the

environment that may be attended to, but in a videoconference, a lack of a shared mutual environment results in over viewing the only real available image, the other person's face.

An extensive research program has also been executed by Ishii, who has built a series of prototypes that use video to combine a semireflective writing surface with images of the participant's upper bodies. This enables the fusing of an image of another participant onto the work surface itself, making it possible to see both participant and object simultaneously and hence accurately track visual attention, while writing or manipulating the object (Ishii and Kobayashi, 1992). Again, a major focus of the work is that crucial collaborative information is embodied in the work object, although systematic evaluation of the benefits of adding this attentional information has not yet been conducted (Whittaker and O'Conaill, 1997). While this is true, many applications do use face-to-face views and even benefit from face access video links. The subject of this thesis is to therefore investigate how best to optimise face-to-face views.

2.2.5C Experimental evidence for effectiveness of shared whiteboard in V.M.C

Individual differences between remote communication technologies often lead to varied communicative success. One particular complementary component that has been applied in remote communication is referred to as the 'shared whiteboard.' This particular component allows remote participants to draw and type in the same workspace simultaneously, and can supplement audio communication, and audio with video communication. Gale (1989) tested whether this additional component improved communication across remote conferences. He studied three conditions: data share (whiteboard) only, data share complimenting audio communication, and data share with audio and video technologies. The participants in this study were software engineers

working in a laboratory and therefore the results would seem applicable to real work scenarios. Gale applied three tasks, which were specifically chosen to represent the various types of office work in the real world, and all required participants to negotiate in order that the task is completed successfully.

For all of the tasks communicative measures of success were determined through dimensions of social interaction and through group and individual productivity. As bandwidth of the communication media increased (from whiteboard only up to whiteboard with video and audio) feelings of 'social presence' also increased. Concerning group productivity, questionnaires were administered to all participants asking them to predict the level of productivity that would entail if the whole site and also if the whole department implemented the particular system they were using. Again, a hierarchy was evident with increasing productivity with increasing richness. Measures of task performance were also undertaken and the results would seem to suggest that the need for video is task specific. Difficult assignments that require higher levels of communication (negotiation, persuasion for example) appear to be more video dependent. Certain assignments that are easier to complete, however, would not require video channels for successful communication, and in this respect would be a waste of resources. To conclude it would seem that a shared whiteboard could be a successful complimentary element to a remote communication system and also that more bandwidth does appear to result in higher communicative success for certain measures (Gale, 1989). Furthermore, the audio and video with whiteboard condition resulted in better task performance, higher ratings of co-presence and improved perceptions of productivity than the audio with whiteboard condition.

2.2.6 Discussion of system quality and set-up

Video systems vary with respect to 'quality.' The public analog voice telephone network in the United States for example typically has a delay of approximately 20 to 30 milliseconds, ISDN (Integrated Services Digital Network) is associated with a delay of approximately 10 seconds, and Live-Net is a high quality system with no transmission lags. Videoconferencing vendors also adopt different set-ups, which have varied communicative success rates. For example, users show preference for the Hydra set-up over the PIP set-up and the voice activated switching mode. Although the quality and set-up of a system invariably have communicative effects, the aim of this thesis is to devise methods for improving communication irrespective of set-up and quality. Although video-mediated communication cannot be considered a single entity (system set-ups vary dramatically), this thesis will focus on problems associated with V.M.C as a whole, and therefore the employment of practices which will improve communication irrespective of which type of system is used. Furthermore, technology is continually improving, and therefore it is necessary to focus on problems associated with V.M.C that are not dependent upon quality. A discussion will follow which considers whether V.M.C is a better alternative to the telephone, and whether V.M.C duplicates face-to-face communication.

2.3: Experimental research: Evidence for the differences between V.M.C when compared with audio-only and face-to-face communication

2.3.1 Introduction

In order to ascertain the effectiveness of V.M.C as a communicative medium the following sections will focus on different measures of 'success,' comparing V.M.C with audio-only and face-to-face communication. Measures of success will include task performance

measures (task outcome), process measures (for example turn-taking and monitoring understanding and attention), user perceptions (preference ratings), and non-verbal responses to using the media. Monk et al (1996) have argued that approaches that integrate many methods are particularly informative when comparing different communication media. Focusing on many measures of success will therefore aid in giving a better understanding of how effective V.M.C is in comparison to audio-only and face-to-face communication.

Video mediated technologies were introduced under the assumption that visual signals are important in human communication, and that by allowing people to have access to visual information, communication can be more successful and rewarding (Whittaker & O'Conaill, 1997). This assumption can be quantified by three predictions concerning how video-mediated technologies might improve communication. One prediction is that access to visual signals supplies important non-verbal information to the observer (the 'non-verbal communication hypothesis,' Whittaker, 1995; 1996). A second prediction is that video allows information concerning the environment to be seen, specifically whether others are available for interaction; this in turn facilitates connection for unplanned communications ('the connection hypothesis,' Whittaker & O'Conaill, 1997). The third prediction is that access to visual stimuli also provides dynamic information about objects and events, which may be of importance for certain collaborative tasks ('the video as data hypothesis') (Whittaker & O'Conaill, 1997). The purpose of this section is to evaluate whether visual signals in video-mediated communication add value in terms of process and outcome by looking at experimental studies comparing V.M.C with audio-only communication and face-to-face communication and then looking at the added value of video-mediated visual cues.

A number of different theories provide useful frameworks when considering whether adding visual information to an audio channel is of benefit. These include:

1) The non-verbal communication hypothesis (Whittaker, 1995; 1996) (Previously discussed in section 1.8.3).

2) The connection hypothesis.

3) Media richness theory (Previously discussed in section 1.8.5).

4) Equilibrium theory (Argyle and Dean, 1965) (Previously discussed in section 1.8.4).

Where appropriate, quantitative and qualitative data will be interpreted using these theories. Therefore, any differences between face-to-face, video-mediated and audio-only communication will be explained in terms of these communication theories.

2.3.2 Task outcome measures

One way to measure whether video-mediated communication is successful is to look at task outcome. Research has been conducted comparing task outcome between V.M.C with face-to-face and audio-only communication on a number of different tasks, including negotiation tasks, problem-solving tasks and simple information exchange tasks. A review of the literature in this area will not only give an insight into the effectiveness of V.M.C in comparison to face-to-face and audio-only communication, but will also decipher which applications V.M.C is most suited to. Furthermore, although there is no standardisation with respect to the quality of systems used in the research, overall patterns and trends will be identified.

2.3.2A Collaborative problem solving and simple information exchange tasks

The most basic forms of task that can be used to make comparisons between media involve collaborative problem solving or the simple exchange of information. The successful completion of such tasks invariably requires participants to communicate with one another and therefore involve a social dimension. These types of task however will be compared with negotiation tasks in section 2.3.2B, where performance outcome relies on conflict resolution. Chapanis and colleagues (Ochsman & Chapanis, 1974; Chapanis, 1975) have carried out a series of studies comparing performance over varied media (handwriting, typewriting, audio, video, and face-to-face). Using problem solving and simple information exchange tasks, the researchers noted that media that offered access to audio information resulted in better levels of task performance than those which did not. For example, conditions that offered an audio channel of communication resulted in the exchange of a larger amount of information and faster task completion rates. The researchers, however, also noted that visual access had no impact on problem solving and simple information exchange. These findings agree with research comparing face-to-face with audio-only communication (for example Davies, 1971b; Reid, 1970). Such research indicates that audio-only communication is as effective as face-to-face interaction on tasks that involve problem solving or the simple exchange of information.

Doherty-Sneddon et al (1997) investigated whether high-bandwidth V.M.C 'videotunnels' (Smith et al, 1991) could replicate the advantages of face-to-face interaction. The face-to-face condition was replaced by a high-resolution video link, which could be set to allow direct eye contact between speakers and which included a full duplex audio link. The study compared the impact of three conditions in total: V.M.C where participants could make eye contact with their partner, V.M.C where eye contact was not possible and an audio-only condition. In all conditions the participants sat in separate but adjacent rooms

and completed the "Map Task" (Brown et al, 1984). Participants who were familiar with one another were recruited, and worked in pairs in order to tackle the task. The researchers hypothesised that if video-mediated communication were to confer the benefits that have been observed for face-to-face interactions, then the high-quality video-tunnel condition (which allows eye contact to take place) should elicit benefits in communication over audio-only communication. Results from this study indicate no difference in task performance between the three conditions (other measures of 'success' from this study will be discussed later in the chapter). This finding informs us that even when video technologies are near perfect (for example, no delay and supports gaze), a participant's ability to solve a problem is still no better than in audio-only conditions. This finding is not surprising in the light of research, which generally indicates no difference in problem solving outcome between face-to-face and audio-only conditions.

Anderson et al (1997b) asked participants to take part in a travel game, with the intention of planning an itinerary around the United States. The travel game was specifically designed to involve a more social dimension than the "Map Task" (Brown et al, 1984), however task outcome did not rely on participant negotiation. The researchers offered a cash prize for the best itinerary, which was judged upon factors including the number of visits planned to the cities and states of the United States. Each participant acted as a 'tourist.' All participants were presented with a map of the United States, which included states and main airports. The participants were required to communicate with a 'travel agent' in order to ascertain details of available flights and connections. In comparing the task outcome, the experimenters noted that there was no difference in the amount of destinations visited for subjects in the three conditions. Once more, this indicates that

access to visual signals had no effect on task outcome. Other measures of task success from this study will be discussed later in this section.

Research indicates that visual information has little or no impact on outcome for problem solving tasks or simple information exchange tasks. On the one hand, video-mediated communication offers little benefit in terms of task outcome when compared with audio-only communication, even with a high-resolution video link that supports gaze (Doherty-Sneddon et al, 1997). Research also indicates that in general participants in audio-only conditions carry out the end goal of problem solving tasks with the same level of success as participants in face-to-face conditions (for example, Davies, 1971b; Reid, 1970). Olson, Olson and Meader (1997) also reveal that remote high quality video can be as effective as face-to-face interaction even when the task is as complex as drafting the initial requirements for an automatic post office (i.e. "a collection of postal services offered through a stand-alone device similar to an ATM") (pg.161). The overall picture from such research is that access to visual images (whether they be co-present or mediated) have little impact on problem-solving or information exchange. In other words the end goal of communication seems to be carried out equally well in conditions which do not have access to visual information as conditions which allow access to visual information. As has been discussed in the previous chapter, however, access to visual signals in face-to-face communication effects how efficiently the communication is carried out. In other words, even though the end product may be the same, the manner in which the task is completed may be more efficient (for example, in fewer words, and in a quicker amount of time). The question therefore remains: does V.M.C offer similar benefits in terms of process? This will be discussed in section 2.3.3. Research from the previous chapter also indicates that participants are generally more satisfied with communicating face-to-face

than audio-only. A discussion will follow in section 2.3.5 in order to note whether video also confers such a benefit.

2.3.2B Negotiation tasks

Evidence suggests that visual information is of little benefit for task outcome in problem solving tasks and the simple exchange of information. Research from the previous chapter however indicates that access to visual information in face-to-face communication is of some benefit for task outcome in negotiation style tasks. This section aims to review the effects of visual signals on tasks that involve a greater degree of conflict in order to decipher whether V.M.C confers the same benefits as face-to-face communication.

Short (1974) looked at the effects of medium of communication on experimental negotiation. In this particular study pairs of subjects performed a simple negotiation task over one of three media of communication (face-to-face, closed circuit television, or a loud speaking audio link). One person was required to argue a case, which was consonant with his/her personal views, whereas the other person was required to argue a case that bore no necessary relationship to his/her personal views. Short (1974) was interested in finding out whether other means of communication could be used as successful alternatives to face-to-face interaction. Short was particularly interested in negotiation style tasks as research by Argyle (1969) had indicated that non-verbal cues are used mainly to convey information primarily about a participant's emotional state, as opposed to matters which are seemingly the subject matter of the meeting itself. Therefore removing certain types of information from the communication channel may result in an increased reliance on other sources of information. As the visual channel appears to be particularly concerned with socio-emotional information, its removal would be expected to lead to an increased dependence

on the more task-orientated verbal channel. According to Short (1974) the removal of the visual channel has consequences for the interaction outcome. In particular, a negotiation scenario might be expected to be particularly sensitive to variation in the medium of communication. As it is clear that a person's actions are highly dependent on perceptions of previous moves made by others, a lack of visual cues may diminish the information available, and therefore distort interpretations of previous actions. This may have potential implications for the outcome of the negotiation task (Short, 1974).

The task involved participants coming to an agreement on three cuts for a business out of a possible nine. Participant A was asked to rank in order which cuts should be made in terms of dispensability, and B's arguments focused on the exact opposite of A's as he/she was given a sheet with A's ranking's reversed. It was expected that A believed in his standpoint but B did not necessarily. Results indicate that for both face-to-face and video participant A was more successful than participant B, which seems to show that emotional factors are taken more into account when visual signals are available. Participant B was also more successful than participant A in the audio-only condition. This shows that the individual whose personal views were compatible with the case he was required to argue would be relatively more successful when visual signals are available, whereas the reverse would hold for the other (Short, 1974).

Daly-Jones et al (1998) have also conducted research into experimental negotiation. In one study, participants were split into pairs and linked from separate rooms by either audio-only or video-mediated technologies. Each member of the pair was given a different brief (one member of the pair had to argue for students with dependants, and the other had to argue for students with disabilities). Furthermore, participants had to decide which three

of the ten applications for a student hardship fund should be accepted. This task was chosen as it was considered to be particularly meaningful to the participants taking part (they were students themselves). The task was also considered to be typical of the kind of tasks for which people tend to prefer to meet face-to-face. Before completing the task the experimenters made it clear that it was not absolutely essential that agreement should be reached. Results of the research indicate that all pairs came to agreement and the mean time taken to do this did not differ between the conditions (Daly-Jones et al, 1998). Other measures from this research will be discussed later in the chapter. This study therefore indicates that visual information in remote communications has no effect on task outcome in negotiation tasks.

Daly-Jones et al (1998) also conducted research on experimental negotiation with groups (in this case a four party debate). Each group of discussants were organised into two pairs, with a pair of participants at each end of the audio or video link. Participants were recorded and data was derived from transcripts and post-experimental questionnaires. In addtion to this, the experimenters also held brief discussions after the experiment had taken place (reports from which will be discussed later). Akin to the one-on-one comparisons, the task involved jointly assessing 10 fictional application forms for a special support fund for students (Daly-Jones et al, 1998). The results from this research indicate no differences for mean time to agreement and task completion rates between audio-only and video-mediated communication. Once more this indicates that visual information in remote communications has no impact on task outcome for negotiation tasks.

Results from research on experimental negotiation are mixed. Short (1974) indicates that access to visual signals alters the outcome of negotiation, and this is true for both video-

mediated and face-to-face communication. Daly-Jones et al (1998) on the other hand indicate that access to visual signals in remote communications does not affect whether participants arrive at an agreement and the length of time it takes to form an agreement, and this is true for both one-to-one negotiation and group negotiation (as will be discussed later in the chapter however, the style of communication was affected). One possible explanation for the lack of difference in outcome in the Daly-Jones study may be linked to participants' feelings towards their arguments. In Short's (1974) study it was expected that one participant would argue a case that they agreed with and the other participant would argue a case that they didn't agree with. The person who believed in their case was more successful when visual cues were permitted. This would indicate that visual information allowed them to get their point across more effectively. In audio-only conditions the participants who didn't agree with their case were more successful as they were not influenced by visual information from the other participant. In Daly-Jones et al's research however it is likely that participants would have been sympathetic towards both the students with disabilities and the students with dependants, and this may have affected the outcome. In other words, because none of the participants may have actually agreed with one side more than the other, negotiation would have been made easier.

A number of researchers (for example, Short, Williams and Christie, 1976; Williams, 1977; Whittaker, 1995; Gale, 1989) have also reported benefits in terms of outcome in negotiation style tasks when access to visual images is permitted. This would indicate that in negotiation tasks where participants believe in their arguments, access to visual signals might be important. Overall, it would seem that access to visual signals in remote communication is beneficial for tasks that involve negotiation, and therefore in this case V.M.C duplicates one of the benefits of face-to-face communication. One way to explain

this difference is through media richness theory. Richer media (in this case media that allow access to visual signals) are better equipped to resolve differences and aid in the exchange of subjective interpretations (Dennis and Kinney, 1998). Access to facial expressions for example allows individuals to better judge the emotional state of conversational participants.

2.3.3 Process measures

Research indicates that video-mediated communication does not differ significantly from audio-communication with respect to one measure in particular: task outcome on problem solving tasks and simple exchange of information. Research does however indicate that visual signals are particularly important for tasks that involve conflict resolution, and therefore V.M.C and face-to-face communication are more successful than audio-only communication when participants are required to negotiate (but perhaps more noticeably when participants believe in their case). Considering that V.M.C has shown some benefits in terms of task outcome, it is also of interest to note whether V.M.C offers the same benefits with respect to process measures. The following process measures will be evaluated: making contact and availability information, turn taking and interaction management, and monitoring understanding and attention.

2.3.3A Making contact and availability information

One hypothesis that can be used to test the effectiveness of video-mediated communication in comparison to audio-only and face-to-face interaction is the 'connection hypothesis.' It can be hypothesised that, in comparison to audio-only communication, video provides availability information about the movements and interruptibilty of co-workers. This visible environment information can in turn facilitate connection for unplanned

interactions. The question is however, can video allow individuals to make contact as successfully as face-to-face communication? In order to test the connection hypothesis researchers have looked at two separate classes of video application in particular. One such application is the 'glance,' which allows an individual to briefly 'look into' the office of a co-worker in order that an assessment may be made as to whether they are available for communication. The other application is the 'open link,' in which video channels are constantly provided between two separate locations.

Fish et al (1993) tested the use of different types of glance in video-mediated communications. The researchers were interested in how successful each type of glance was for promoting chance interactions. The most frequently used type of glance (occurring 81% of the time) was a brief glance in the direction of a selected target. The researchers also indicated that 54% of these glances lead to a prolonged conversation. All other types of glance occurred less frequently and were less likely to result in extended interactions. One type of glance was applied in order to simulate the type of chance encounters that occur in face-to-face situations (for example, bumping into someone in a corridor). Such chance encounters are normally unintended, but often lead to extended communication. These system-initiated connections resulted in high failure rates. In fact, 97% of the connections were completed without conversation. Fish et al concluded that callers in a V.M.C environment prefer to have direct control over whom they connect with and when the connection is made. Further to this, people desired to use the glance as a preparation for communication. This differs from face-to-face communication where a glance is often used to see who is around. In this respect, glances that allowed people to look into another location without the choice of communicating were infrequent, accounting for only 12% of user-selected glances.

Tang et al (1994) also explored the relationship between glances and chance interaction. Participants were given the option of looking into another location, with the intention of discovering if another individual was available for communication. Tang et al found that resulting connection rates were no better than the telephone. This leads one to question why connections were relatively infrequent. Although many failures occurred when an individual was absent from a location (38%), or when a recipient indicated that they were unavailable for communication (4%), the reasons behind the remaining failures are unclear. Perhaps these failures can be attributed to the recipient being busy with another task or talking to another person. Tang et al did not report this data, however.

The connection hypothesis can also be explored using continuously open links between locations (Fish et al, 1993). This set-up is intended to simulate sharing the same environment as another, and it is therefore expected that persisting availability information should make connections between locations easy. Fish et al (1993), however, reported that only 5% of connections lasted more than 30 minutes and Tang et al (1994) reported that only 5 interactions (out of a possible 233) lasted more than 30 minutes. Evidence from both sets of research would therefore suggest that brief interactions are the main use for such a system. Continuous open links can be constructed in office environment or between public areas, for example a coffee area (Abel, 1990, Fish et al, 1993). Open links in public areas are intended to simulate the types of chance meetings that occur in such places in face-to-face contexts. Data from such research indicates that individuals connect more frequently for social greetings than other types of connection (Abel, 1990). Researchers therefore suggest that open links are ineffective in supporting work (Fish et al, 1993). Fish et al (1990) also compared the frequency of verbal communications as a result of sighting

someone over a video link with interactions following face-to-face sightings. The researchers found that sightings over a videolink were less likely to result in prolonged conversation than face-to-face sightings.

In conclusion, the results on glance and open links indicate a lack of evidence to support the video as connection hypothesis. Failure rates for glancing were as high as the telephone. Open links were chosen infrequently by users, and were less likely to promote communication than face-to-face sightings. Furthermore, open links are inadequate to support work (Fish et al, 1993). These failures, however, may be a result of confounding variables. In wide-area connection applications, for example, video quality is poor. Participants may therefore be less motivated to use video, as it will be clear to them that any ensuing conversation will be over low quality video. Also, in office environments, availability information may be accessible by other means, for example moving around the workplace. Such variables may have therefore lead to the reduced use of video for connection (Whittaker and O'Conaill, 1997). Nevertheless, the research in this area indicates that video is less successful than face-to-face communication for promoting unplanned interactions and does not differ significantly from the telephone.

2.3.3B Interaction management

A number of researchers have made comparisons between different media on how efficiently participants manage interactions. In order to achieve this, an analysis of the surface structure of conversation may be undertaken. This may include an analysis of how well participants negotiate turns, the number of interruptions made by participants, or the number of words or turns used by subjects. An analysis of such measures will give a clear indication as to how efficiently participants communicate.

Once an interaction has been set-up it needs to be managed. In face-to-face communication this is achieved with the help of a number of linguistic and non-verbal devices. Research has also been conducted which tests how effectively participants manage interactions in video-mediated conditions. In other words, are the same linguistic and non-verbal devices used in video-mediated interactions? Daly-Jones et al (1998), for example found no evidence of more fluent speech (turn-taking) in one-to-one interactions in video-mediated communication compared with the audio-only communication. The number of turns, the average length of turns, and frequency of overlapping speech also did not differ significantly between conditions. According to the researchers this would suggest that participants can co-ordinate talk as effectively in audio-only interactions as in video-mediated interactions.

Doherty-Sneddon et al (1997) also looked at interaction management in video-mediated communication. In their research they found that dialogues in the video-tunnel condition (eye contact) were significantly longer (11% more turns and 10% more words) than the normal video-mediated and audio-only conditions. The audio-only and normal video-mediated conditions however did not differ. Contrary to expectations, video-mediated communication with eye contact was significantly less efficient. Interruptions were more frequent in both video conditions, in comparison with audio-only (5.6% of turns compared to 3.6%). Doherty-Sneddon et al (1997) state that this finding is surprising considering that Boyle et al (1994) found that interruptions accompanied on average 8.7% of the dialogue in face-to-face conditions and 12% of the dialogue in audio-only conditions. Remote interactions even between separate rooms seem to produce a more formal and less interactive communicative style in general. In Boyle et al's (1994) study, when speakers

were physically present dialogues were more interactive, with a higher incidence of interruptions in general, but visual signals may allow speakers to time their contributions somewhat more precisely and reduce the incidence of overlapping speech. So, visual signals (in a face-to-face context) may encourage a more interactive style of communication and therefore more relaxed communication may result in increased overlapping speech.

In a comparison of length of dialogues Anderson et al (1997b) also revealed no difference between V.M.C and audio-only conditions. This result is surprising as the same study also compared face-to-face with audio-only, and the face-to-face condition produced 22% fewer words with the same task outcome, revealing a more efficient use of language. There were, however, significant differences in the amount of interruptions, with the video-mediated condition producing considerably fewer examples of overlapping speech. This lack of differences in conversational style has also been noted in an evaluation by Fish et al (1993), who noted that interaction over video systems are more like phone conversations than face-to-face meetings, with conversations tending to be brief and task focused.

According to Daly-Jones et al (1998) linguistic devices (for example, sentential, grammatical and clausal information) are adequate to allow participants to manage interactions successfully. This conclusion is supported by a number of researchers. Duncan and Fiske (1977), for example note that transitions from one speaker to the next are co-ordinated by the completion of a grammatical clause, the rise or fall in voice pitch at the end of an utterance, any drawl or expansion of the final syllable and the use of stereotyped expressions such as 'you know.' In this light it is not surprising that visual

cues are not necessary for co-ordinating turns at talk in dyads. Several researchers have suggested however that visual information is of benefit for interactions which involve groups of interactants (Backhouse and Drew, 1992, Bly et al 1993). In such instances, there is a greater degree of uncertainty with regards to who comments are aimed at, and therefore restricting visual access may prove problematic (Daly-Jones et al, 1998).

To test this hypothesis, Daly-Jones et al (1998) also investigated interaction management in group interactions over V.M.C. The data indicates that significantly more conversational turns were employed by each quartet during use of the video link. The mean length of spoken turns was also significantly longer under the audio condition and overlapping speech occurred to a greater extent in the video condition. According to the researchers these effects can be interpreted as evidence of greater fluency of conversation in the video condition. This finding indicates that video-mediated communication delivers some of the face-to-face benefits, and therefore in video-mediated communication with groups the non-verbal hypothesis can be accepted. In contrast to the previous experiment Daly-Jones et al also found more instances of vocal backchannels in the video-mediated interactions. This finding is surprising considering that in the absence of a visual channel to convey gestures and head nods, the audio-only participants should have been more reliant on vocal backchannels. The researchers offered one explanation for this finding. Because participants in the video condition contributed more spoken turns it may have simply been the case that there were more occasions when some form of acknowledgement signal was required (Daly-Jones et al, 1998).

Research by O'Conaill et al (1993) also indicates that system quality has an effect on interaction management. Research suggests that face-to-face communication is

characterised by three properties of communication channels: messages are received almost instantaneously by listeners; communication is two way, for example, feedback can be produced at the same time as the speakers' utterances; and both verbal and visual channels are utilised (Whittaker, 1992). The aim of O'Conaill et al's research was therefore to address how different channel properties of videoconferencing technology affect the nature of spoken conversation in real meetings by comparing interaction in two wide area systems with face-to-face conversation. The analysis also focused on the impact of channel properties on visual behaviour. Comparisons were made between: a) a videoconferencing system with half-duplex audio, transmission lags, and poor picture quality (ISDN); b) a high quality videoconferencing system with duplex audio, no transmission lags, and full bandwidth video (Live-net); and c) face-to-face communication (O'Conaill et al, 1993). The researchers compared five ISDN videoconferences, four LIVE-NET meetings, and five face-to-face meetings. All meetings were scheduled for work-related reasons and were not arranged for the study; hence a more naturalistic observation was achieved. Measures were taken of the number of utterances, the number of words per utterance, backchannels, interruptions, overlapping speech, and handovers. (O'Conaill et al, 1993).

Backchannels and interruptions have been discussed in the previous chapter, however it is necessary to define examples of overlaps and handovers. Overlaps differ from interruptions in that they are instances of simultaneous speech that follow signals given by the speaker, indicating that they are relinquishing the conversational floor (Levinson, 1983). Different types of overlap can occur, for example 'projection/completion' is an overlap where the next speaker anticipates that the current speaker is about to finish or tries to help the "forward movement" of an ongoing utterance (Clark & Wilkes-Gibbs, 1986).

'Floorholding' is another type of overlap and occurs where the next speaker tries to take the floor while the current speaker attempts to hold the floor while producing utterances that do not contain any information (Jefferson, 1984). 'Simultaneous starts' are a final type of overlap and occur when two or more participants concurrently begin a new turn, for example when the current speaker has relinquished the floor and each new speaker may believe that there are no contenders for the floor (Sacks et al, 1974). (O'Conaill et al, 1993). Handovers occur when speakers signal that they intend to relinquish the floor using explicit verbal cues, for example by asking explicit questions by the method of 'tagging' (the use of stereotyped questions such a 'isn't it?' or 'aren't they?' or statements such as 'you know' or by the addition of redundant information on the end of a turn) (O'Conaill et al, 1993). An analysis of such behaviours will therefore give a clear indication as to how efficiently participants communicate.

In all meetings in the O'Conaill et al (1993) study two judges were asked to identify every instance of backchannels, interruptions, overlaps, and formal handovers. Results indicated that, when compared with face-to-face meetings, spoken conversation patterns were disrupted over ISDN with its half-duplex line, transmission lags, and poor image quality. These disruptions took the following forms:

1) Listeners produced fewer backchannels and were interrupted less often.

2) Listeners were also less likely to anticipate turn endings.

3) Speakers also altered their behaviour, being more likely to hand over turns formally using a question or naming the next speaker. They were also less likely to hold the floor with redundant phrases.

4) The result of the listeners reducing interruptions and speaker feedback, combined with the general difficulty of switching speakers, was a formal lecture style of interaction, with long turns, handed over by a very deliberate process.

In the LIVE-NET condition, even when there was a full duplex line, immediate transmission, and a broadcast-quality image, the properties of the spoken communication still differed in many respects from face-to-face interaction.

These differences were as follows:

1) Although listeners interrupted as frequently in face-to-face, they were less likely to give backchannels.

2) Speakers used questions to formally hand over the floor more frequently, and they were also less likely to hold the conversational floor with redundant information.

Thus, although better quality video systems (LIVENET) produced more similar communication to face-to-face interaction than poor quality systems, high quality V.M.C was still characterised by highly formal conversational behaviours (O'Conaill et al, 1993). For several conversation characteristics, the high quality system (LIVE-NET) was more like the lower quality system (ISDN) than face-to-face. This result suggests that other channel properties are also critical here and according to O'Conaill et al (1993) "the account should be extended to include these properties and to determine which conversation characteristics they impact" (O'Conaill et al, 1993, pg 421). This research therefore indicates that participants manage interactions less efficiently in video-mediated conditions than in face-to-face interactions. There is some evidence however to suggest that higher quality systems result in a more efficient use of language than lower quality systems.

More efficient communication in face-to-face interactions can be explained in terms of access to non-verbal cues, specifically head turning and eye gaze. Research by Duncan (1972) has indicated that head turning and eye gaze play an important role in speaker switching. Duncan (1972) has also shown that head turning and eye gaze are reliant on directionality. Therefore, as theorised by O'Conaill et al (1993) the absence of directionality in ISDN and LIVE-NET may lead to changes in speaker behaviour, with speakers having to use the verbal channel to signal turn transitions explicitly and carefully manage speaker switches. O'Conaill et al (1993) further this sentiment by adding that this result may go some way to explaining the increased incidence of questions in ISDN and LIVE-NET and the reductions in floor holding. The difference between LIVE-NET and face-to-face therefore may arise because speakers rely on directional gaze in face-to-face to elicit backchannels, and this cue is removed in both LIVE-NET and ISDN. It was also observed that there were conversational characteristics that seemed to be unaffected by channel properties, for example, turn distribution. According to the researchers it may be the case that such examples are reflective conversational behaviours produced independently of the communication situation (O'Conaill et al, 1993).

O'Conaill et al (1993) also consider the implications about the kinds of tasks for which the current ISDN quality is appropriate. According to the researchers, the lecture-like style and the evident incapability of the ISDN system to support quick-fire exchanges indicates that it is unsuitable for certain types of tasks, for example conflict resolution, planning, or negotiation. The reason for this is that ISDN cannot support rapid clarification and feedback, which is a requirement for such tasks. The researchers also suggest that in future remote collaborators should attempt to choose appropriate communication technologies for

the task at hand and ensure that certain types of tasks, for example negotiation, should always be resolved if possible in a face-to-face setting. Although it is clear that improving the quality of a video system (low lags, full duplex channels) leads to improvements in communication, as evidenced by the superiority of LIVE-NET over ISDN, the LIVE-NET results also suggest that improving these properties alone will not exactly reproduce face-to-face interaction (O'Conaill et al, 1993).

The evidence for the ability of V.M.C to support interaction management successfully is mixed, however the resounding view is that video-mediated communication is less successful than face-to-face communication. Evidence also suggests that V.M.C is more like audio-only communication, although higher quality systems are better than lower quality systems and there is some evidence to suggest that interactions are managed more successfully with groups over V.M.C. There are a number of possible reasons for these findings. Although on the one hand many researchers agree that visual signals are not crucial for regulating turns, it is possible that non-verbal behaviours such as head turning and eye gaze are important in regulating speaker transitions, and further to this V.M.C systems do not support such behaviours. While not crucial non-verbal behaviours have an impact on interaction management. Therefore in this case the non-verbal hypothesis can be rejected: access to visual information in video-mediated communications does not allow participants to manage interactions as successfully as access to visual information in face-to-face interactions.

2.3.3C Monitoring understanding and attention

It can be hypothesised that with the use of video-mediated communication establishing mutual understanding should be easier than with audio-only communication because of

access to non-verbal information. Campbell (1998) states that such non-verbal information enables groups to create shared meaning and a shared understanding, which in turn can lead to acceptance and action. Rogers (1986) emphasises that, while mutual understanding is the main aim of communication, it can never be obtained in an absolute sense. Several exchanges of information over time may increase the level of mutual understanding about a topic, but complete understanding is unachievable. Communication is usually adequate when the required amount of mutual understanding is achieved to perform the task at hand. In conversation this means that conversational partners attempt to establish mutual understanding to the extent that the receiver has understood what the transmitter meant to a satisfactory criterion for present purposes (Clark and Brennan, 1991). This model of communication views the creation and sharing of information as a "convergence towards (or divergence from) a common point of mutual understanding" (Rogers, 1986, pg 330).

One way to test this is to investigate how well participants monitor understanding and attention in different media. In the Daly-Jones et al (1998) study, as predicted by the researchers, participants in the audio condition asked more questions than participants in the video condition. The use of direct questioning may have been served as a way of securing information concerning the other participant's reaction. Essentially, a lack of visual information made it difficult for participants to monitor understanding and attention. This is borne out in the manner in which participants responded to a number of questions on a post-experimental questionnaire. Respondents were more aware of the presence of their partner when visual access was possible. Participants also found it easier to monitor their partner's levels of attention in the video condition.

Evidence from group interactions also suggests that participants find it easier to monitor understanding and attention in video-mediated conditions compared to audio-only interactions. In an analysis of the questionnaire data Daly-Jones et al (1998) found that the mean ratings for awareness of the remote partners in the video condition were very similar to the mean rating for the awareness of the co-present partner in both audio and video conditions. However participants found it much easier to monitoring their partner's attentional status in the video condition. For example, there was a significant difference in favour of video for the following question: 'during our conversation I could readily tell when (partner) was occupied with the forms and not listening to me.' There was also a significant effect for video for the following question: 'I could readily tell when…was concentrating on what I was saying' (Daly-Jones et al, 1998). This result informs us that participant perceptions concerning the ability to monitor understanding and attention are better in video-mediated communications than in audio-only interactions.

Further evidence comes from Doherty-Sneddon et al (1997), who found that there were a significantly higher amount of Align games in audio-only condition than both video-mediated communication conditions (V.M.C with eye contact and V.M.C without eye contact). In other words, participants check that their partner understood more often, perhaps accredited to a lack of no-verbal cues. So, in this respect V.M.C delivered one feedback function associated with face-to-face communication. There is evidence here then that video-mediated communication is more successful at allowing participants to monitor understanding and attention. This can be credited to access to crucial non-verbal cues associated with understanding and attention, for example head nods and facial expressions.

Evidence from the literature therefore suggests that video-mediated communication is more successful than audio-only communication in providing participants the ability to monitor understanding and attention. This finding can be attributed to access to visual signals. Research has indicated that visual behaviours, for example head-nods, gazing behaviour and facial expressions, provide information concerning an individual's level of attention and understanding. It would therefore follow that the advantage of video over audio can be explained in terms of access to such non-verbal behaviours. Therefore, V.M.C offers one benefit associated with face-to-face communication, and in this case the non-verbal hypothesis can be accepted.

2.3.4 Non-verbal responses

Another measure that can be used to test the effectiveness of V.M.C concerns non-verbal responses to using different media. Non-verbal behaviour serves little communicative function in audio-only interactions, as conversational participants do not have access to visual information. The addition of a visual channel of communication in video-mediated interactions however would suggest that non-verbal behaviours could be transmitted successfully. The validity of this view will be discussed in more detail as follows.

Through the use of a post-experimental questionnaire Daly-Jones et al (1998) noted that participants were concerned with the loss of natural eye contact in a video-mediated setting. The following quotes illustrate some of the participants' concerns.

1) 'If I looked at the screen I found that I wasn't actually keeping eye contact.'

2) 'Even when they were looking at you they were looking at you on the screen, whereas you're getting the view from the camera so it was like you were looking down on them. It wasn't proper eye contact.'

Although users undoubtedly have access to visual images, it would seem that natural eye contact in a videoconference is impossible. The eye contact problem is due to the camera being mounted above the monitor. It is possible to rectify this by constraining the position of the eyes of the viewed partner to a particular region of the screen and then arranging for the position of the camera to correspond with this region using mirrors as in the Xerox EuroPARC Video Tunnel (Gaver et al, 1992). Research by Doherty-Sneddon et al (1997), however indicates that the use of such videotunnels results in an overuse of the visual channel. In fact, they noted that there were significantly more gazes in eye contact condition compared to no-eye contact (239 compared to 144), more than double that recorded in face-to-face dialogues. This over-gazing may in turn interfere with cognitive processing and speech planning (Beattie, 1981). Therefore, the benefits of having access to visual signals may be counterbalanced by users becoming distracted by their partner's face and so using significantly more speech to achieve a comparable level of task success. Doherty-Sneddon et al (1997) go on to argue that this over-gazing effect may be a direct result of the novelty of using such equipment, and are interested to note whether this would change over a period of time. Over-gazing may also explained in terms of equilibrium theory. According to Argyle and Dean (1965) there are many cues to intimacy (for example proximity, touch and eye contact), and in cases where such cues are restricted individuals may compensate with other available cues. In this case participants may have compensated by over-gazing. According to Daly-Jones et al (1998) in the long term people adapt quite quickly to a lack of eye contact and essentially learn to get used to it. While this concern is important, the researchers conclude that it must have had little or no effect on communication. In essence, if this problem were important, then the effects of interpersonal awareness and fluency of conversation would not have occurred (Daly-Jones

et al, 1998). Many researchers have however indicated that, in comparison to face-to-face interactions, communication is less efficient and less fluent in video-mediated interactions. The fact that participants have noted this problem to be of concern would also indicate that it still deserves some attention.

To truly preserve eye contact in a video system, the participant must look directly into the camera, but because the person must also be looking at the video image of the other person, or at the computer screen, eye contact must be compromised. The most typical compromise, both in computers and videotelephone and conferencing devices, is to place the users' camera above the screen. This causes people to appear to be casting their eyes downward, but this is better than other positions, as people appear suppliant when they gaze upward and deceitful when they gaze sideways (Kenyon et al, 1985). There have been several novel solutions to preserving eye contact including videotunnels and the clearboard project (Ishii & Kobayashi, 1992), which superimposes video of the other participant on a common whiteboard while a camera preserved eye contact using half-silvered mirrors (Angiolillo, 1997). These particular systems, however, are not readily available on the market, and therefore this problem with eye contact is still an issue.

Technical constraints in video limit the ability of users to make eye contact and use gaze awareness. Such constraints contribute to the inability of V.M.C to fully portray non-verbal signals, and according to Angiolillo (1997) are one of the major causes of the weakness of video to provide value to remote conferencing. As well as video not being able to provide users with the ability to make eye contact, video-mediated technologies also result in participants over-gazing at the screen. Because participants stare consistently at the image on the screen, none of the normal modulation of gaze behaviour that are

characteristic of face-to-face interactions are realised (Duncan, 1972). Similarly, Abel (1990) reports such 'T.V watching' effects. The end result is that speakers are presented with an image of a person staring almost directly at them. People tend to report that such behaviour is off-putting and even confrontational. Such behaviour also means that speakers will receive little in the way of local attentional feedback as local listeners are staring continuously at the remote site (O'Conaill et al, 1993).

As well as natural eye contact being compromised, Monk and Gale (2002) indicate that full gaze awareness is also difficult to achieve in a normal videoconferencing set-up and is dependent upon the scope of the image provided. Traditionally, V.M.C is set-up in such a manner that only an image of another person's head and shoulders is available. The problem with this is that individuals in a videoconference will be unaware as to where and at what their partner is looking at. Monk and Gale indicate that providing wider coverage to expand to the environment around participants may therefore be beneficial. In doing so, images of the participants' faces will become less clear, however, Daly-Jones et al (1998) indicate that benefits of facial expressions have been over stated in most task contexts.

Various novel attempts have been made been researchers to provide full gaze awareness in V.M.C. Velichkovsky (1995) and Vertegaal (1999) for example used eye-tracking devices to help detect where individuals were looking. Monk and Gale (2002) report the benefits of full gaze awareness in video-mediated communication. Using apparatus that supported gaze awareness (GA Display), a number of effects were found. In comparison to two conditions (V.M.C with eye contact and audio-only), the gaze awareness set-up resulted in a lower number of turns and words in order to complete the task. The researchers go on to suggest that this finding can be explained in terms of understanding. Essentially, full gaze

awareness provided an alternative non-linguistic method for checking one's own and another individual's understanding. In conditions where gaze awareness is not possible, individuals must signal understanding verbally, which is less efficient and therefore takes a longer period of time.

The same level of non-verbal interaction found in face-to-face interaction does not occur in the use of video-mediated technologies and has a number of effects. As well as the problem associated with gazing behaviour, it is evident that many systems use only head shots (talking heads) and therefore eliminating any benefit that non-facial non-verbal behaviours such as posture, and gesture have to offer (Videotalk, 1998). Indeed, O'Malley et al (1996) provide some evidence that video-mediated technologies are more effective when access to the head and shoulders is provided as opposed to a head shot only. As far as facial expressions are concerned the major foreseeable problems concern delay and the size of the image. In instances where delay occurs, turn-taking transitions become increasingly difficult and interruptions occur more frequently (Angiolillo, 1997, O'Malley et al, 1996). Perhaps this effect can be attributed to the disruption of visual signals, as research has indicated that eye gaze is particularly important in the turn-taking process (for example, Kendon, 1967). In instances where the size of the image is small, this makes it increasingly difficult to pick up crucial non-verbal signals, perhaps explaining user preferences for life-size images (Inoue et al, 1984).

When non-verbal cues are removed from communication the result is a lack of social presence (Rice, 1993). Participants are depersonalised and therefore are seen to be less like people and more like objects (Williams, 1977). According to Dennis and Kinney (1998), there is a general acceptance that media with fewer cues are less friendly, more

task focused, more impersonal and even depersonalising. In turn, this depersonalisation has effects on social and communicative processes. Sproull and Kiesler (1986), for example, note that depersonalisation can result in participants becoming more self-centred, and Siegal et al (1985) even noted that depersonalisation could lead to antisocial behaviour. Just as importantly, Williams (1977) noted that lack of non-verbal cues could lead to decisions based more on facts than personalities. Having access to fewer channels of non-verbal communication results in changes in social and communicative processes. V.M.C therefore does not allow participants to use many non-verbal cues as effectively as in face-to-face communication.

2.3.5 User perceptions

Another method that can be employed to test the effectiveness of video-mediated communication concerns user perceptions. Here, users' subjective feelings towards media are assessed. Williams (1977) for example found that preference ratings for individuals meeting over an audio link were lower than those meeting over a video link. These differences can be explained in terms of participant access to visual communication cues in V.M.C. Daly-Jones et al (1998) also asked participants to assess how suitable video-mediated communication was in comparison to audio-only communication. One question in particular asked whether participants agreed with the following statement: "during our conversation I was able to focus on the task at hand." Results from this question indicate no difference between video-mediated and audio-only communication, although the researchers argue that this could be due to a ceiling effect. However, other posed questions ('the mode of communication aided collaborative work,' and 'rate below your overall evaluation for the communication facilities provided') both showed large and significant effects in favour of the video condition. The experimenters also found a significant

difference for the question: 'please rate each mode of communication according to how much effort you felt it was to converse effectively'. This indicates that participants feel that they need to make less effort in order to communicate effectively in a videoconference than in audio-only communication. Participants also rated the video-mediated condition as being higher in terms of co-presence. Furthermore the video link was considered to be more suitable for the task than the audio link. Overall, these results indicate that high quality video links are viewed more favourably than high quality audio links (Daly-Jones et al, 1998).

Anderson et al (1997b) also conducted a post-task questionnaire, which probed the participant's satisfaction with audio-only, video-mediated and face-to-face communication. When asked how easy it was to make changes to the itinerary, 30% thought it was very easy in the VMC condition compared to 20% in the audio condition, but 70% found it very easy in the face to face condition. Therefore, there is some evidence here that richer media result in higher evaluations.

Even though there are many communicative drawbacks associated with the use of V.M.C, one resounding finding is that people nevertheless tend to prefer video-mediated communication to audio conferencing (O'Conaill et al, 1993). Participants taking part in O'Conaill et al's 1993 study indicated many reasons for this preference. One advantage of using videoconferencing over audioconferencing lies purely in the visual channel: knowing who was at the remote site and knowing who was speaking, although it should be noted that this is often dependent on the quality of the image, with an increased difficulty in discriminating speakers being noticeable in poorer quality systems. Another stated advantage was the feeling of 'not talking into a void.' O'Conaill et al's participants,

however also commented that they found videoconferencing appropriate for only certain types of meeting such as information exchange and project updates.

Evidence from the literature also suggests that the type of task interacts with medium of communication with regards to participant evaluations. In a study in 1998, Campbell investigated the effects of technology quality on willingness to participate in future videoconferencing meetings. Campbell was also interested, however, in how the type of application effects user dispositions. Campbell theorised that if video-mediated meetings are constrained by the technology, then users will indicate less commitment or willingness to participate in future videoconference meetings. Therefore, the higher the perceived quality of the system interface, the more positive the user disposition towards using videoconferencing systems. Campbell also noted that there is a strong relationship between the type of communication activity performed and the user's perception about the quality and outcome of that meeting. Therefore Campbell also expected that the 'richness' of videoconferencing would be better suited to some specific communication tasks more than others.

One's decision to try or indeed avoid the use videoconferencing is primarily effected by the individual's expectations of improved productivity. Therefore, according to Campbell (1998) the consequent adoption of videoconferencing technology is largely dependent upon the communication and information requirements of the user group and the ability of the technology to facilitate these requirements.

McGrath (1984) identified four basic types of group tasks.

1) Generate ideas and plans

2) Choose correct of preferred answers

3) Execute performance tasks

4) Negotiate conflicts of interest and viewpoints

For his experiment Campbell used these four types of task. Participants were made aware that they were allowed to choose the type of task that they wanted to use the equipment for.

Campbell's findings indicate that videoconferencing does not facilitate the desired level of interaction during meetings involving negotiation, conflict, or the generation of ideas, however video-mediated communication seems to be more suited to choose/execute style tasks. There is also evidence from the research that the constraint on interaction tends to increase the task focus of the group, which can, in some instances, lead to higher quality outcomes. For example, groups that did take part in negotiation style tasks had positive feelings about the results despite the fact that, compared to face-to-face interaction, there was a lessened opportunity for interaction.

Overall, evidence from the literature would suggest that participants prefer the use of video-mediated communication to audio-only communication. There is also some evidence to suggest that face-to-face meetings are preferred over video-mediated interactions. This would indicate a direct relationship between media richness and user preference. The preference for video-mediated communication over audio-only communication can be explained in terms of participant access to visual information. The fact that users prefer face-to-face contact over V.M.C, however would indicate that preference ratings are affected by other factors other than access to visual information.

The differences between media in user preferences can also be explained in terms of social co-presence. V.M.C should result in improved perceptions of co-presence in comparison to audio-only communication. This can be explained in terms of access to visual information. Face-to-face communication however takes places in a co-present environment, and therefore perceptions of co-presence should be higher in comparison to video-mediated communications. Evidence for this comes from Gale (1989), who noted that as bandwidth increased, so did feelings of social presence. In particular he noted that higher perceived levels of co-presence were achieved with a shared whiteboard with audio and video capabilities, in comparison to lower levels of perceived co-presence with shared whiteboard only. Habash (1999) found that the use of group decision support tools (GDST) had a significant impact of user perceptions of social co-presence. When GDST tools were not used, video-mediated communication was perceived as being more effective and supporting higher levels of social co-presence than audioconferencing. However, when GDST tools were used, this difference was not found, suggesting that GDST tools can help to bridge the gap between audio and video conferencing systems. Furthermore, social co-presence has been shown to affect media choice. Straub and Karahanna (1998) suggest that social co-presence is a good predictor of media choice. King and Xia (1999) however found evidence to the contrary. They suggest that choice of media for a task is not affected by levels of social co-presence, but rather by the individual's experience with the media.

According to Campbell (1998), there is a general agreement within in the literature that, as bandwidth narrows (for example, from face-to-face communication to audio-only communication), the communication becomes less appropriate for interactions that are socially complex and interpersonal communication. Schweizer et al (2001) compared varying degrees of social co-presence and its effects on communication. The researchers

conducted a field study on distance learning, comparing 4 conditions: a tutor mediated by written information; a tutor mediated by written information and various personal views; a tutor mediated by written and spoken information; and a tutor mediated by text, views and spoken language. Results indicated that in conditions with less co-presence, participants relied more on emotional evaluations as well as adopting a more task-orientated and informal approach and resulting in tense reactions. Paechter et al (2000) noted that participants who experience a tutor with a high degree of social co-presence are significantly more likely to achieve better learning outcomes. They also gave better evaluations of the tutor and the learning process.

Further evidence for the effects of social co-presence on communication comes from Anderson et al (1997b). As well as the other measures of success (discussed earlier) the researchers were also interested in exploring the impact of visual signals on the participants' decision making. Comparisons were made between the number of optional changes made in each condition. Optional changes are those in which "the client decided on a new destination from choice, for example, by revising an earlier part of the itinerary, asking for alternative possible destinations, or using the surcharge option in the rules of the game to backtrack to a previous destination" (pg. 148). Such changes required more work on the travel agents part and therefore reveal a higher level of shared social presence. The results indicate that there was no significant difference between audio-only and video mediated conditions but face-to-face was noted to be better than audio-only with an average of 6.6 optional changes per dialogue compared to 2.7. Although in many ways, video mediated communication had advantages over audio-only communication, the results from this experiment inform us that it fails to deliver the same benefits as face-to-face communication. Logically, this indicates that perhaps the visual processes that

facilitate communication in face-to-face interaction may not transfer to video mediated communication.

Social co-presence (or indeed the lack of) has been noted to have a number of communicative and social effects. As well as altering the social processes and the communicative process, lack of social co-presence also seems to effect task performance. According to Dennis and Kinney (1998) social facilitation research implies that working on a shared task in the presence of others should increase the speed of a person's task performance on a simple task and impair performance on more complex tasks (Levine et al 1993; Sana, 1992). Therefore Dennis and Kinney theorise that a lack of social presence could decrease social facilitation effects and increase the time required to complete a task. If this is indeed the case then a lack of social presence in a videoconference may also have negative effects on task performance. According to Chapanis (1988) as the multiplicity of cues decreases then decision time increases. Evidence for the effects of decreased cues on decision quality and task performance is mixed, however, as no clear overall pattern emerges (Dennis and Kinney, 1998). Therefore access to more cues seems to have an effect on time to make a decision and efficiency, but not necessarily task performance.

2.4: Real world applications: Evidence for the differences between V.M.C when compared with audio-only and face-to-face communication

Many researchers have looked into the effectiveness of V.M.C for real world applications, for example distance learning. Although in many cases such research does not offer face-to-face comparisons, they give useful insight into whether V.M.C can be used effectively for applications that are normally expected to be completed face-to-face. Research on real world applications are often measured in terms of task outcome (for example, do students

complete a course successfully via video-mediated technologies?), but also rely heavily on user perceptions. Although users perceptions have been discussed more generally earlier in the chapter, they will also be discussed in relation to task outcome measures in this section. The reason for this is that user perceptions often go some way to help explain differences in performance between media.

Currently, corporate and academic sectors appear to be making the most use of videoconferencing technology. In a business environment, video-mediated communication has been put to many different uses. "Business television," for example is a corporate communications tool, which involves the transmission of information via videoconference and is commonly used for meetings, product announcements and training (Videotalk, 1998). Video-mediated communication may also be used to conduct interviews. Bankers Trust, for example, regularly conduct videoconferences with business-school students when recruiting for the New York-based bank's Asian Branch offices. According to Julie Dines, assistant Vice President for Bankers Trust "videoconferencing is a lot more cost-effective than flying all the students across the Pacific" (Strom 1997, pg 1). Julie Dines initially screens each applicant in person and then conducts an interview via videoconference with her branch managers (Strom, 1997).

Many companies also apply videoconferencing in order to accelerate the feedback process in their manufacturing cycles. This particular use comes about, as often the managers are located in different cities from the plants that produce their products. According to Gary St Onge, vice president of Everett Charles Technologies "videoconferencing is extremely valuable in situations that require a high level of detail and it can eliminate many misinterpretations" (Strom, 1997, pg 2). Dennis Miniero, executive vice president of

Radiant House Europe Ltd, states, "we use video for viewing documents and editing them together, as well as showing finished samples of items that are under development" (Strom, 1997, pg 2). Strom (1997) also notes that there are numerous cases where video enables the development of new applications. According to Jim Knox, a multimedia-services manager for the communication-services department of Stanford University, "many of our conferences involve activities that would not otherwise have occurred: music students at Princeton, San Jose State, and Stanford sharing professors and projects for a computer-music course. Or departments bringing in guest lecturers for whom they couldn't afford to pay travel expenses" (Strom, 1997, pg 2). Strom also cites another example of video-mediated communication enabling a novel application: the Gertrude Stein Repertory Theatre in New York City is using experimental videoconferencing system from Lucent Technologies to collaborate on joint performances involving American and Russian casts (Strom, 1997). Evidence from this interview data would indicate that the use of video technologies in a business environment is more cost effective than flying people around the world, supports numerous functions and enables applications that without this technology would normally not have occurred. Video-mediated communication is also used prominently in academic environments. The following research considers the effectiveness of video technologies for educational purposes.

The most basic form of video-mediated communication as an educational application is "instructional television." Instructional television involves a pre-recorded lecture/seminar being viewed after the lecture has taken place by those students who could not attend the live lecture. Instructional television may not be regarded as videoconferencing in the traditional sense, as it does not involve a live broadcast and therefore is not a dynamic communicative process, with no feedback from the lecturer/speaker being possible.

Instructional television, however, is a method of video-mediated communication and has much in common with videoconferencing. Much of the dynamics are essentially the same as distance learning, for example the lecture is viewed on a monitor and the lecturer is not physically present in the same room as his/her students. Any problems with such a method of transmitting information may therefore aid in understanding whether video-mediated communication is effective as an educational tool.

The literature in this area, for example Smith and Nagel (1972) and Trent and Cohen (1973) implies that instructional television is not significantly different from traditional instruction with regards to one goal in particular: the efficiency of learning. Dublin and Hedley (1969) reviewed the effectiveness of instructional television at University level. The results of this study indicate that on 191 different comparisons: 102 participants favoured instructional television, while 89 favoured traditional instruction. However, the majority of differences (90%) were not significant. The authors' interpretation of these results was that there seems to be no significant differences between instructional television and traditional instruction at the University level. According to Sullivan et al (1976) these studies and reviews, although on one hand are reassuring concerning the general effectiveness of instructional television, are not particularly conclusive or helpful, especially in providing guidance for the most effective techniques of production. Essentially these studies had been carried out in an applied setting during a whole semester. Sullivan et al (1976) consequently theorised that it was not possible to isolate and manipulate the relevant independent variables to provide for the necessary practical and experimental controls and to measure the dependent variable (i.e., the amount of learning that had taken place) with sufficient precision.

Sullivan et al (1976) therefore investigated whether instructional television could produce as high a level of achievement as live lectures. The researchers were also interested in ascertaining under what circumstances instructional television would be most effective (Sullivan et al, 1976). The experimenters prepared three different methods of instruction. All lectures came from an introductory psychology course and therefore the lecture material was standardised. The first method consisted of a live lecture, which was recorded at the time of presentation. This recording was used as the second method of presentation. Finally, the third method of instruction involved a studio-produced videotape in the absence of an audience. In order to measure the effectiveness of each individual method of instruction, the students were expected to write a five minute pre-test and a five minute post test and therefore the basic measurement of achievement was gain scores as measured by the difference between pre- and post-tests (Sullivan et al, 1976).

The results indicated clearly that live instruction produced a significantly higher level of achievement than did a videotape of the same live instruction. Also, the videotape of live instruction produced a significantly higher level of achievement than did the studio-produced videotape. (Sullivan et al, 1976). The attitudes of the students towards each method of instruction were also assessed through the use of a post-experimental questionnaire that was administered to each student individually. The correlation between actual effectiveness and perceived effectiveness was invariably higher than +.70 and frequently higher than +.80. The correlation between effectiveness and liking for a particular mode was always positive but somewhat lower (+.40 to +.65). These results indicate that it is possible to demonstrate that live lectures produce a significantly higher level of achievement than do instructional videotapes (Sullivan et al, 1976).

Sullivan also notes that the basic difference between live and videotaped instruction concerns the greater variability in post-test scores following videotaped instruction. Sullivan theorises that this increased variability can be associated with an individual's personality traits. For example, students considered anxious and introverted did equally as well as others after hearing a live lecture, but showed significantly lower scores following instruction by videotape. Sullivan goes on to state that differences may be accounted for, in part at least, by the likelihood that in a live lecture the instructor is able to deal more effectively with the differing needs, personalities, and learning styles of students. Sullivan notes that the greater variability in scores could also be explained through the idea that anxious and introverted subjects may become apprehensive when taught by a new method of instruction and therefore would inevitably learn less effectively. Sullivan proposes that more than likely it is a combination of such factors (Sullivan et al, 1976). This finding isn't surprising if we consider theories of communication and learning, in particular grounding theory. Clark and Marshall (1981), for example, argue common knowledge is crucial for achieving an important aspect of communication, namely reference. Reference enables conversational participants to jointly identify the objects and events that they want to talk about (Clark and Marshall, 1981). Having access to a shared environment would therefore result in reference to be achieved more easily than communication that takes place at a distance.

Results from experimental research in this area are mixed. On one hand, research would have us believe that Instructional television is as effective as live lectures, however, research by Sullivan (1976) suggests that live lectures are a more effective means of presentation than Instructional television. This finding is partly credited to the idea that feedback from the lecturer is impossible during a pre-recorded videotape, as the lecturer is

not present to answer questions or ease student worries. If this is the main reason for the lack of effectiveness of this method of communication, then it is possible that live broadcasts over video-mediated technologies may be a more effective method of instruction.

In the academic sector, "distance learning" is a relatively new application, which incorporates audio and or video technologies for educational purposes, so that widely dispersed students can attend training seminars and courses without travelling to where the course is being presented (Videotalk, 1998). Distance learning differs from instructional television in the sense that the video broadcast is live and all students have live access, be it co-present or mediated, to the instructor. Considering that audio and video systems are now applied more regularly to distance learning settings, it would be sensible to ascertain the relative effectiveness of these systems for such a task. Many studies have been conducted to test the effectiveness of distance learning, for example Robinson and Jones (1985) studied the effects of using audioconferencing as a distance-learning tool. Students studied for a Postgraduate Diploma in Continuing Education at a distance site via audio links. Robinson and Jones (1985) found that the students successfully completed the course, however, through means of post-experimental questionnaire it was discovered that the students did not recommend audio-only communication as an effective form of distance education. The students explained that they expected some sort of visual communication, as they found it difficult to maintain attention and the lack of visual cues restricted interaction between the origin and remote sites. Students also noted that local accents appeared to be exaggerated over the telephone.

A major problem associated with distance learning via audio-only technologies seems to be that a lack of visual cues restricts interaction. For example Webster and Hackley (1996) have found that students feel at a disadvantage at a remote site in that they feel less included in the course, report lower levels of interaction and find it difficult to contact the course tutor or instructor. It would therefore seem appropriate to test the comparative effectiveness of video-mediated technologies for such an application, in order to test whether access to visual signals improves communication in a distance-learning environment. Armstrong-Stassen et al (1998), for example, studied the suitability of using video-mediated technologies in a distance-learning environment through the means of questionnaire directed at students at the origination site (same room as tutor using V.M.C equipment) and the remote site, at the beginning of the semester and the end of the semester. Since many of the tutors had no experience in using such equipment, all tutors were given a manual entitled "Hints for Videoconferencing Instructors" (Armstrong-Stassen et al, 1998 pg.154). This manual included a section on "personal presentation (looking at the camera, projecting one's voice, not moving too quickly) and a section on encouraging interaction" (Armstrong-Stassen et al, 1998, pg.154). The justification behind using such a manual is that it has been found that students in remote sites often feel isolated and therefore encouraging interaction and looking up at the camera (giving the impression of eye contact) will help to relieve these feelings of detachment.

Results from the Armstrong-Stassen study (1998) suggest that there was little difference in attitude towards videoconferencing whether students were at the origination site or at the remote site. Although we would expect the remote students to have more negative views and feel more isolated, the fact that there is no difference isn't surprising. Students at the origination site will also feel isolated as the tutor will in essence be talking to the camera

and monitor and not necessarily directly to the students in the room. What the experimenters did find was a less positive attitude towards videoconferencing and an increase in anxiety. Armstrong-Stassen et al suggest that feelings of anxiety may be more to do with lack of experience and therefore it is as important to prepare students for videoconferencing as it is to prepare the tutors. Although the experimenters did find that those who were anxious and had a negative attitude at the beginning of the semester were also negative and anxious at the end of the semester, which implies that experience with the equipment did not alleviate this problem.

On the whole student problems fell into one of four categories: technical difficulties, accessibility of instructor, distracting/impersonal environment, and participation constraints. Many comments were made for each of these concerns. Examples include:

1: Technical difficulties: "background noise," "difficult to hear," "unclear video image."

2: Accessibility of instructor: "loss of direct contact with instructor," "can't 'connect' with the professor," "difficult to relate to image on the screen."

3: Distracting/impersonal environment: "creates a sterile environment," "classroom setting impersonal," "difficulty maintaining attention span," "shifts focus from the subject of the course to the technology."

4: Participation constraints: "stilted discussion; loss of classroom exchange," "less willing to ask questions in fear of camera being turned on him/her," "people at remote site feel like observers rather than participants," "intimidated by camera; self conscious about seeing self on camera."

Leidener and Jarvenpaa (1993) found that the course subject matter relates to different types of instructional methods, in other words certain technologies are better suited to

certain types of information. One example quoted includes that less anxiety was felt in the communications studies course in the Armstrong-Stassen et al study (1998). This would suggest that certain courses are well suited to videoconferencing but others are not, implying that some courses may require greater modifications in order to make them more suitable. Perhaps also, it is the case that communications tutors are better suited to using such technology because they have more appropriate knowledge, or even that they understand better how to communicate successfully under such conditions. An overall picture from Armstrong-Stassen study (1998) would suggest, "there is a tendency to force users to adjust to the technology rather than designing the technology to the needs of the users" (Armstrong-Stassen et al, 1998 pg.163). Robinson (1995), for example has noted that conferencing technology has yet to be designed that suits the requirements of the teachers using them. Fussell and Benimoff (1995) further this sentiment by stating that design decisions are based upon technical limitations and cost factors as opposed to what would suit a certain application. What must be considered then is that until prices come down or technological advancements are made users must work with what is available to them. In other words users could be trained or advised so that they can make the best use out of imperfect, yet highly useful media.

Abbott et al (1993) also studied the effects of using video-mediated technologies for distance learning. The experimenters collected data from pre and post-course questionnaires and from in depth interviews with both the students and tutors. Induction began at the time of registration and it seemed appropriate for the tutor to do this by videoconferencing. According to the experimenters using the videoconferencing equipment for the first meeting proved to be a mistake as half the students later reported that they were 'thrown in cold.' One student in particular stated that "the induction was

very explicit but it didn't ease my fears. If anything, they increased" (Abbott et al, pg 88, 1993). Some further induction took place face-to-face, which proved to be more appropriate than videoconferencing for the start of a new learning experience. These face-to-face meetings were later to be considered as essential. For example, one spontaneous face-to-face meeting took place in week 5, and this enabled student uncertainties at that stage to be removed. According to Abbott et al (1993) for more effective teaching and learning, the use of videoconferencing in a distance form of delivery required students and tutors to have face-to-face contact with one another for a quarter of the total classes in the year. Although this would entail students travelling on occasion to the campus, the experimenters felt that this would be worthwhile. According to the researchers, the fact that face-to-face contact was required highlights the idea that in terms of effective learning, the use of videoconferencing is a means to an end and does command a delivery of its own. In the opinion of Abbott et al (1993) it would seem that videoconferencing needs to be placed alongside physical contact in order for such distance meetings to be effective.

Results of the questionnaire data and interviews with students and tutors were interesting. At first students were apprehensive about using the technology, partly credited to the novelty of using such equipment and also lack of technical expertise. At the end of the semester, however, increased confidence was duly noted as all students were asked to lead the seminar and all students were willing to give it a go. No student reported any feelings of inhibition, but almost all felt some sort of frustration directed at their own ineptitude with the technology. Group size was also a factor that the students believed to be of importance. In the experiment eight students took part at the remote site. Seven of the eight students believed that their group was of optimum size for the kind of co-operation they had achieved on collective tasks, however, the students also were of the view that a

larger group would have had an inhibiting effect on their learning (Abbott et al, 1993). This sentiment is also borne out by Baron and Byrne (1984) who have stated that more than five in a group can reduce the willingness "to go along." Although videoconferencing can often be regarded as an intimidating method of communication, it was noted that this indeed helped to promote a rapid bonding within the group. According to the researchers, tackling a new system resulted in a sense of urgency right from the start (Abbott et al, 1993).

The fact that the tutor was not physically present at the sessions gave the students 'the impression of being alone,' however this decreased social co-presence resulted in the students being able to discuss the issues more openly amongst themselves: in a sense there were feelings of privacy as well judgements of being monitored. One student stated that: "videoconferencing was at its most successful when the group was carrying out an exercise without the need for tutor intervention" (Abbott, 1993, pg 90). Another student, however, noted that videoconferencing might result in the less active participant hanging back and not engaging in group discussion. It seems that videoconferencing may be advantageous over face-to-face meetings in that it encourages the students to be more independent and can encourage participation. As has been discussed, however, this is dependent on the individual, and with less physical presence many students may be less willing to join in conversation.

Over and over we see evidence for a loss of co-presence in V.M.C. If this is the case then potentially there are a number of psychological and communicative effects of using this technology. This particular theory is borne out by Doherty-Sneddon and McAuley (2000). They found that children were better able to disagree with adults in a video-mediated

interview than in a face-to-face interview. This finding is quantified by the fact that the children in the video condition were less likely to answer incorrectly in response to closed questions and were also more resistant to misleading questions. This result agrees with Goodman et al (1998) who indicate that children were more resistant to misleading questions in a mock video-mediated court trial than a mock face-to-face court trial. Doherty-Sneddon interpreted their findings to indicate that children feel more comfortable disagreeing with an adult when there is distance between them. In this case a feeling of increased distance between child and adult was noticed in the video-mediated condition. Davies and Noon (1991) also found that judges were less empathic to live link witnesses than face-to-face child witnesses, indicating that they are being judged more on content than emotion. These studies indicate that there is a decrease in social and emotional pressure as distance is increased.

Overall, like the audioconferencing experiment conducted by Robinson and Jones (1985), the students in the Abbott et al (1993) study completed the module as effectively as those students who attended tutorials face-to-face. Although there were many technical difficulties, the experimenters believed that the occasional face-to-face meetings and group cohesion were the main factors in achieving the objectives of the module. According to the external examiner the use of videoconferencing did not "impair the quality of modules and in some significant ways enriched the opportunities for learning – a view shared by the students." (pg 90). Abbott et al compare these results to a previous experiment in which two groups were each trying to cohere simultaneously. According to the experimenters this resulted in a sense of separation rather than unity. In a setting in which two groups were operating at once, it was easier to remove oneself from the interactive process than it was in the single, distinct group context. Like Armstrong-Stassen et al (1998) Abbott et al

(1993) also regarded the quality of technology to be an important factor, stating that universities must be willing to invest in the best possible equipment.

Temple (1999) also reports that videoconferencing has been used successfully in order that business students can develop new products at a distance. This research indicates that other forms of distance communication (for example e-mail or telephone) were unsatisfactory, as the students were uncertain as to whether their message had been received and also as to whether the message was being acted upon.

On the whole, results from the real world application experimentation indicate that students can successfully complete a course even when the tutor/lecturer is not physically present. This is true for both audioconferencing and videoconferencing. Unlike Instructional television, distance learning is advantageous in the fact that feedback can be given by the lecturer at the time of delivery. It is also agreed that video-mediated communication in this context is more successful when supplemented by face-to-face meetings. Overall it would appear that video-mediated communication in a distance learning environment is less satisfactory than face-to-face learning, however, it can offer it's own benefits, for example improved bonding between students, and improved levels of team-work. Although many problems can exist in the use of videoconferencing as a distance learning tool, according to Dallat et al (1992) when video-mediated communication is judged by the standards of face-to-face teaching and learning, it can approximate but not equal it. When judged by the standards of distance learning, however, V.M.C "has a potential, and a challenge of its own" (pg. 88).

2.5: Advice on good practice

Based on a review of case studies on distance learning, Alexander et al (1999) argue that there are certain practices which can be employed in order for more successful communication over a videoconference. Aside from the technological considerations, users can achieve more satisfactory communication by paying heed to the following advice. The presenter should plan and rehearse his/her presentation and should take considerable care to involve both sides of the link in discussion (for example, through asking questions, or pausing to allow others to interject). Before the videoconference, users are also encouraged to initiate a warm-up session, with an emphasis on informal interaction. This could include: personal introductions, games to break the ice, or panning the camera around the room. Through adhering to such advice users may experience higher levels of co-presence. Indeed, personal introductions may go some way to relieve the feelings of isolation that users experience during video-mediated interactions.

During the videoconference users are advised to take into account the differences between visual and audio quality in a videoconference in comparison to face-to-face interaction. Users should speak clearly, pause for reflection, allow time for delay (if appropriate) and indicate clearly when they have finished their turn. This advice would seem appropriate, as research implies that it is more difficult to manage interactions in video-mediated communication than in face-to-face communication. Indeed, through indicating clearly when a turn is finished, users can compensate for the lack of non-verbal cues (for example eye gaze and head turning) that are used in the turn-taking process. Researchers such as Sellen (1992), propose that the implementation of directional audio and video might also bridge the gap between face-to-face and video meetings. A second strategy would be to modify existing videoconferencing systems by acting on the comments made by O'Conaill

et al's (1993) participants. These participants suggested that providing remote audio and video controls would be beneficial, so that remote participants can choose what they want to see and hear as opposed to having these choices made for them.

In the visual domain, users should take into account that subtle expressions will more than likely be lost. They should also avoid the use of excessive hand movements, however hand signals should be used to draw attention to who is speaking. After the videoconference, it is also advised that time should be given for an informal 'warm down' session (Alexander et al, 1999). Research by Hackman et al (1990) has indicated that such advice can result in more favourable communication. Instructors who engaged in behaviours such as encouraging involvement, offering feedback, maintaining relaxed body posture, and speaking clearly were viewed more favourably by students in terms of satisfaction and perceived learning outcomes.

O'Conaill et al (1993) also suggest that remote collaborators should attempt to choose appropriate communication technologies for the task at hand and ensure that certain types of tasks, for example negotiation, should always be resolved if possible in a face-to-face setting. Although it is clear that improving the quality of a video system (low lags, full duplex channels) leads to improvements in communication, as evidenced by the superiority of LIVE-NET over ISDN, the LIVE-NET results also suggest that improving these properties alone will not exactly reproduce face-to-face interaction.

2.6: Chapter conclusion

Despite the multimodal physical nature of face-to-face communication, the most pervasive and successful technology for communicating at distance is the telephone, which relies

solely on the voice modality (Whittaker and O'Conaill, 1997). According to Doherty-Sneddon et al (1997) video-mediated communication potentially offers us the benefits of face-to-face interaction. Doherty-Sneddon et al (1997) note that visual signals are used in many ways in face-to-face communication and V.M.C systems should be designed to emulate these functions. For example speakers use visual cues in face-to-face communication to judge that communication is proceeding smoothly and hence need to elicit verbal feedback less often. Second, listeners are more confident and check their understanding of messages less often when they have access to visual signals. Third, visual cues may be used to establish a sense of social co-presence that makes us feel at ease with our conversational partner and so feel able to ask for additional information or help in a shared task (Doherty-Sneddon et al, 1997).

Evidence from the literature would suggest that even high-quality audio and video does not replicate face-to-face processes (for example Doherty-Sneddon et al, 1997), although there is also some evidence that high-quality video systems produce effects more similar to face-to-face interactions than low quality systems (for example, O'Conaill et al, 1993), and therefore supporting the media richness hypothesis. O'Conaill et al (1993) argue that one possible explanation for the difference between face-to-face and V.M.C is that video systems do not accurately simulate the presentational characteristics of face-to-face communication. A review of the literature indicates that V.M.C offers some benefits over audio-only communication (most notably in levels of satisfaction and task efficiency), however there is little evidence for task outcome differences (with the exception of tasks which require an element of negotiation). In comparison to face-to-face interactions, video-mediated communication has been shown to be less satisfactory, less efficient and results in lower levels of social co-presence. The non-verbal hypothesis can also be used

to explain these differences. Research clearly indicates that the same level of non-verbal communication cannot be transmitted via video as in face-to-face communication. Considering the importance of non-verbal behaviours in human communication (discussed in previous chapter), this lack of non-verbal signals may feed into the differences found between the two mediums. Evidence from the literature also indicates that video cannot be used as successfully as face-to-face interaction to supply availability information (for example Whittaker, 1996).

Overall, video-mediated communication does not produce the same benefits as face-to-face communication, and indeed a number of problems are associated with video-mediated communication: both human and technological. In order to ascertain which problems are most pertinent to current users a questionnaire was circulated to users of video technology and an interview was conducted with two 'experts'. The results of the questionnaire and interview are discussed in the following chapter. Furthermore, the interview and questionnaire data will be used to back up current theories for the lack of success of video-mediated communication and to validate the experimental focus of this thesis. As with researchers such as Olson, Olson, Carter and Storrosten (1992), and Strauss and McGrath (1994), this thesis will take a multi-dimensional approach to studying the effects of video-mediated communication. Both quantitative (for example, measures of task success, including time to completion, performance measures, and analysis of the surface structure of conversation) and qualitative measures (for example, user preferences and satisfaction levels gauged through questionnaire and interviews) will be undertaken.

Chapter 3: User Perceptions of Video-Mediated Communication and Expert Analysis of Videoconferencing

3.1: Introduction

3.1.1 Aims and objectives

This thesis will investigate the psychological and communicative effects of using video-mediated technologies, with the aim of determining whether certain practices improve the use of such technologies. This chapter will report preliminary qualitative data, which will help set the scene for the detailed quantitative analyses that will follow in the next two chapters. First, user preferences and perceptions are reported. This will help to identify whether users of current technology view V.M.C as a viable alternative to face-to-face communication. User perceptions will also help to identify the problems that exist in the use of such technology and methods that can be used to improve the use of the technology. Second, the perceptions and preferences of two experts in video-mediated communication will be reported. This will give the perspective of suppliers on the cutting edge of the technology and offer alternative viewpoints to those expressed by the users in the questionnaire data. A qualitative approach will help tap user perceptions and preferences in order to decide upon empirical studies that will follow on from this research later in this thesis. In other words, although key areas of study have been notified through a review of the literature, users' opinions pertaining to the technology will also help to identify avenues for exploration. Therefore, the aim of this chapter is to validate existing findings and theory and to discover pertinent communicative difficulties caused by the use of video-mediated technologies in order that practices may be undertaken to improve the use of the technology.

3.1.2 Previous research on V.M.C

In a review of the literature, Egido (1990) concluded that certain video-mediated communication systems fail to achieve acceptability as substitutes for face-to-face co-

present interaction. For example, Egido points out that visual images of other participants in a meeting add little information and that the use of video teleconferencing for scheduled meetings ignores informal interaction through which much communication in organisations takes place. Egido does, however, identify certain situations in which videoconferencing has been found to be useful, notably lecture-style presentations in which structured one-way communication with a large audience is desired. Researchers such as Armstrong-Stassen et al (1998) and Abbott et al (1993), have also indicated that videoconferencing is a useful tool in a distance learning environment, concluding that video-mediated communication can be as successful as face-to-face communication for teaching purposes. Success in these examples, however, was measured by the satisfactory completion of a course. It should also be noted that in the Abbott et al (1993) study, video-mediated communication was supplemented by face-to-face meetings, and in the Armstrong-Stassen et al (1998) study students complained that it was difficult to connect with the lecturer, the environment was distracting and impersonal, and there was a noticeable loss of classroom exchange. Therefore, clearly a multi-level approach to measuring success is required, that involves user preferences and task performance.

Olson, Olson and Meader (1997) have also noted that remote high quality video can be as effective as face-to-face interaction even when the task is as complex as drafting the initial requirements for an automatic post office. However, for problem solving tasks having access to visual signals is typically of little benefit, for example participants usually complete problem solving tasks as successfully in audio-only conditions as they do in video-mediated conditions. V.M.C has been noted, however, to be extremely

useful for tasks that are more socially orientated such as bargaining and negotiation (Anderson et al, 1997): in such tasks access to visual information is important.

Evidence from the literature also suggests that many problems exist in the use of video-mediated communication, dependent upon the type of system being used, the amount of bandwidth, the application and also the set-up. Evidence from O'Conaill et al (1993), for example, suggests that better quality video systems (LIVENET) can produce more similar communication to face-to-face interaction than poor quality systems (ISDN), however high quality V.M.C is still characterised by highly formal conversational behaviours. A review of the literature also suggests that different systems or set-ups are more preferable than others. Sellen (1995), for example, noted that users preferred using a Hydra set-up as opposed to a Picture in a Picture set-up or a voice activated switching mode set-up. Rather than proposing to advise users to purchase technology to suit the application, however, it would seem more appropriate to train or advise users so that they may be able to perform more efficiently when using any videoconferencing set-up. Consequently, the main aim of the current research is to construct an overall framework for improving video-mediated communication irrespective of set-up or system quality. Although technical issues are important, the focus of the research is to pursue a deeper understanding of the social and communicative factors involved in video-mediated communication.

The first stage of this thesis was to identify the particular problems associated with videoconferencing as it exists currently, with an overall aim to provide suggestions for the improved use of such technology. In order to identify the difficulties and in turn decipher methods to improve communication, user perceptions were evaluated and an

expert analysis of videoconferencing was undertaken. Users perceptions of video-mediated communication were gauged through the use of a questionnaire looking at user perceptions of the technology and an expert analysis was undertaken through the means of an interview with two senior members of a videoconferencing facilities supplier.

3.2: Study 1a: Questionnaire

Existing literature shows that the problems associated with video-mediated communication are varied. In order to narrow the focus of the empirical studies, a questionnaire was designed in order that the most important or most reported problems by the users themselves could be given attention. This questionnaire was circulated to current and former users of remote communication technologies. In order to get a representative picture of the kinds of problems experienced by users of video-mediated technologies, respondents from a wide range of backgrounds were required. Participants were therefore recruited from a number of different sources. E-mails requesting participation in the study were sent out to a number of sources: these were a large computer software company in the United States, academic departments at both Stirling and Glasgow University and a number of videoconferencing suppliers. Questionnaires were sent to these locations only after a senior member of staff had granted permission. All returned questionnaires were used in the analysis. In addition to this a number of participants were opportunistically sampled from a collaborative research project between Stirling and Glasgow University.

This questionnaire employed the use of open-ended questions as opposed to multiple-choice questions. The value of using open-ended questions is in the richness given in

the answers. There is also less chance of ambiguity as respondents can say exactly what they feel as opposed to responding to a fixed set of items. It was considered important that participants were allowed to respond freely with little or no restrictions. Multiple-choice questions by their very nature restrict participants to choosing from a set number of possible answers. Also, much research in this area has employed the use of video-mediated systems in laboratory settings. Furthermore, in many cases participants have been asked to complete tasks that are not representative of the kinds of tasks that such equipment is used for in the real world. For this reason it is possible that many participant perceptions have not been reflective of the kinds of problems encountered by individuals who use video-mediated communication in real world settings. Open-ended questions would therefore take into account the possibility that problems exist in the use of videoconferencing systems that have not yet been discussed in the literature.

3.2.1 Design

The questionnaire was designed with the intention of eliciting free reporting. Questions were chosen specifically to target users' opinions relating to how well V.M.C replicates face-to-face interaction, the problems that exist in the use of the technology, and techniques that may be employed in order to counteract these problems. All participants completed the questionnaires anonymously.

3.2.2 Participants

In total 48 questionnaires were returned from 5 different groups of users. Twenty-three participants stated that they used video-mediated communication for business purposes and frequency of use ranged from very occasionally to weekly meetings. Applications

included status meetings, collaboration on projects, company announcements, group business meetings, and one-to-one business meetings. These participants were contacted by e-mailing the questionnaire to their place of employment. Subsequently, the completed questionnaires were returned via e-mail. Eleven participants were first time or infrequent users who had taken part in a videoconferencing experiment. This experiment involved one-to-one remote collaboration between participants at Stirling University and Glasgow University. These participants were given questionnaires after the experiment had taken place and were informed to fill in the questionnaires in their own time and return them to a designated pigeonhole. Three participants stated that they were suppliers of conferencing equipment or conferencing facilities. Two participants were noted as being researchers into video-mediated communication. Both the suppliers and researchers were contacted by e-mailing the questionnaire to their place of employment. Finally, the last 9 participants were academic videoconferencing users at the University of Stirling. These participants either used videoconferencing for teaching purposes or were students who regularly took part in remote lectures. These participants were contacted face-to-face and were asked to return questionnaires via internal post at the university.

3.2.3 Materials

A questionnaire containing six questions was circulated to all participants. The questionnaire contained the following questions.

1) Under what circumstances would video-mediated communication be most beneficial?

2) Do you believe that video-mediated communication has advantages over communication over the telephone? What are these advantages?

3) Do you believe that videoconferencing systems can replicate face-to-face communication? Why?

4) Did you note any problems when using video-mediated technologies?

5) If you could make any changes to conferencing systems for the purpose of more successful communication what would they be?

6) Please add any further comments.

3.3: Results

3.3.1 Question 1: Under what circumstances would video-mediated communication be most beneficial?

Question 1 was included to gather information concerning opinions relating to the ideal applications for remote communication technologies. This should give an insight into the types of applications that users believe videoconferencing to be most suitable for. From this, a better understanding can be gained of the limitations and advantages of video-mediated communication. The results from this question are summarised in the following table, with the number of responses for each suggestion.

Table 3.1: Frequency of suggestions made for circumstances under which video-mediated communication would be most beneficial

Suggestion	Freq	Suggestion	Freq
1) Interview	2	11) Group meetings	10
2) In house use	1	12) Ideas session	2
3) Collaborations	2	13) With familiar person	2
4) Initial meetings	2	14) Demonstrations	1
5) Everyone has same technology	1	15) Business	3
6) Travel costs/geographically dispersed	14	16) Visuals required	1
7) Distance learning	6	17) Small groups	2
8) Internal meetings	1	18) No alternative	3
9) 1 to 1	5	19) Meeting of sensitive nature	4
10) Project reviews	2	**Total**	**64**

A total of 19 different suggestions were made concerning the most beneficial applications for video-mediated communication. Many suggestions were only mentioned by one participant, for example internal meetings and in house use. Other suggestions, such as for group meetings, were mentioned more frequently.

Ten participants noted that video-mediated communication is useful for group meetings, indicating that video-mediated communication is well suited to the dynamics of group meetings. Further comments made by these participants, however, suggest otherwise. One participant stated that group meetings should only be conducted over videoconference when people are geographically dispersed, in other words if there is no alternative. Two participants indicated that group meetings are beneficial only if there

is a "defined protocol." Two more participants asserted that group meetings are only successful if there is an agenda, with one participant in particular stating that it is more efficient if "participants are already familiar through real contact." Another participant stated that one needs to "make sure all notes/paper is forced to the other links prior to the session" in order for group meetings to run smoothly. These comments highlight the idea that a certain level of structure is needed for group videoconferencing to be effective. Although it is clear that group meetings are a favoured application, 6 out of the 10 participants say so only because it saves on travel costs and time, or point out that a high level of structure is needed in order for communication to be effective. Perhaps then it could be argued that group videoconferencing is an application used out of necessity as opposed to in relation to its effectiveness. It appears as though group meetings are only favoured if there is no alternative (for instance, if the members are geographically dispersed). Furthermore, the necessity of having a "defined protocol" or "clear "agenda" is evident. This sentiment reflects evaluations of video-mediated interactions by Fish (1993) and Angiolillo (1997), who concluded that V.M.C users adopt a formal, task-orientated approach. In essence group communication over videoconference is not so much of an ideal application for the technology, but more of an ideal method for communicating when physical boundaries, such as distance, are imposed. This sentiment is borne out by Abbott et al (1993) who noted that face-to-face contact is essential at some point when communicating at a distance for prolonged periods of time.

Five participants indicated a preference for one-to-one communications with an emphasis being placed on the difficulty of videoconferencing for group interactions. One participant stated: "I think I would remind people that effective communication

tends to decline as a function of the number of active participants. Having a 5-site multi-point VC (videoconference) with 4 or 5 members in each group is just impossible. None-the-less administrators insist on promoting this as an effective means of communicating." Research suggests that as the number of participants increases then communication becomes more difficult, for example Baron and Byrne (1984) have stated that more than five in a group can reduce the willingness "to go along." However, group size is not always important and is dependent upon the particular application. For lectures over videoconference, for example, one person is talking to many and communication is directed mainly in one direction. In group meetings the dynamics are different as both sides usually wish to communicate. One reason for the problems associated with interactions in groups is that when large groups on either side are used the images of each participant are small. It therefore becomes increasingly difficult to judge someone's reactions, not only on the other side of the link but also in the same room as focus is on the monitor and not on those people sitting next to you.

Two participants stated that initial meetings would be a favourable application for videoconferencing. One of these participants, however, also noted that initial meetings work only "if direct eye contact is not absolutely required or feasible...a vision of the work environment tends to lessen the feeling of distance and unfamiliarity...however, direct contact is a must at some point."

Although this participant reported that video-mediated communication is useful for initial meetings it also clear that there are stipulations attached. We note in particular the importance of eye contact in a business meeting. It appears that the user believes eye contact cannot be achieved when engaging in a videoconference and notes that in

many meetings such eye contact is necessary. The user also noted that there is a problem with co-presence and distance, in so far as he suggested that this could be overcome with "a vision of the work environment." However, the user asserts that this "tends" to help, implying that while being one option to reduce feelings of distance, this method is not completely successful. Perhaps then it is necessary to formulate another means of creating a sense of co-presence between users and therefore reducing feelings of distance. What must also be considered is that videoconferencing is not always used as a business application. Are feelings of isolation noticeable with other applications and if so, how for instance, could one counteract feelings of distance in an academic videoconference? Research by Abbott et al (1993) suggests that feelings of distance are pronounced in distance learning environments, and indeed one way to overcome this is to supplement V.M.C interaction with face-to-face meetings. An increase in familiarity therefore seemed to bring participants closer together when they interacted remotely.

Two participants proposed that videoconferencing is most effective when users are familiar with one another and have had prior face-to-face contact. This sentiment is clarified further with the quote "not good for external meetings, particularly when people have not developed a personal relationship." This quote illustrates that feelings of distance are present in the use of video-mediated technologies. Furthermore, feelings of distance can be alleviated with personal contact prior to videoconferencing.

Related to this are the comments made by four participants, who stated that videoconferencing is useful when a meeting is of a sensitive nature. For example in a "diplomatic role where people would not be able to face each other without risk of aggression." Two other participants also expressed a preference for videoconferencing

in "courtroom" situations. These issues seem to imply that interaction alters in some way during the use of video-mediated technologies. People regard the use of videoconferencing as different to face-to-face interaction in the respect that the same degree of personal contact is not available and therefore it is easier to remain emotionally detached from the subject matter that is being discussed. This sentiment is borne out by Williams (1977), who asserted that a lack of non-verbal cues could lead to decisions based more on facts than personalities. There are obvious advantages to remaining emotionally detached in situations such as courtroom cases and diplomatic meetings in the sense that it will be easier to remain impartial and not be swayed by ones emotions. Further to this, research by Doherty-Sneddon and McAuley (2000) also highlights this point. Results indicate that when giving evidence concerning a neutral event, children were more resistant to misleading information in the video condition in comparison the face-to-face condition. The researchers go on to explain this finding by stating that these effects are due to a decrease in social co-presence in V.M.C, which ultimately resulted in decreased feelings of intimidation and increased feelings of confidence on the part of the child witnesses.

3.3.1A Discussion of question 1

Informants from this particular sample perceive video-mediated communication to be an appropriate application for groups. Six out of the 10 participants who reported videoconferencing as being particularly suited to group meetings, however, indicate that this is primarily because it saves on travel costs and time, or point out that a high level of structure is needed in order for communication to be effective. Participants who use videoconferencing for academic purposes particularly note the ineffectiveness of V.M.C for group communication with many statements reflecting that video communication

would be best put to use with small groups (2) or on a one to one basis (5). This particular sentiment is interesting, as this is the opposite of what they are using videoconferencing for. Therefore we can assume that the technology is clearly not serving the requirements. Remote communication technologies, especially in a business environment, are mainly used as alternatives to face-to-face meetings, because travelling is more costly and time consuming. In fact 14 informants report the benefit of videoconferencing when people are geographically dispersed. The idea that videoconferencing saves on time and money is expressed by Dallat et al (1992), who noted in their research that lecturers saved 63 hours of travel time through alternating the teaching venue. Also a total of 126 hours of travel time were saved in one year for three modules. Dallat et al also pointed out that from the point of view of the staff the avoidance of the fatigue associated with long evening journeys was also a bonus. Perhaps then, remote technologies are used to save money and relieve the burdens of travel, rather than in relation to their effectiveness as a meeting application tool. This opinion is supported by 3 informants, who thought that videoconferencing should only be used if there is no available alternative.

The data also indicates remote conferencing may result in a lack of social co-presence between locations. Further to this many methods may be employed to help alleviate lack of co-presence, including initial face-to-face meetings, a view of the room, and replicating eye contact. The benefit of initial face-to-face meetings to supplement video-mediated meetings has been demonstrated by Abbott et al (1993).

It would seem then that video-mediated communications are on a needs must basis, and are used as alternatives for face-to-face communication because they save on travel time

and fatigue. One problem associated with V.M.C, however concerns a loss of social co-presence, although a suggestion to alleviate this problem would be to increase familiarity through initial face-to-face contact, and increasing 'personal closeness' by increasing levels of eye gaze. Both of these ideas are taken forward in the empirical work that follows in chapters 4 and 5.

3.3.2 Question 2: Do you believe that video-mediated communication has advantages over communication by the telephone? What are these advantages?

This question was included to investigate the ways users perceive advantages that videoconferencing has to offer over audio-only communication. In essence, it was necessary to ascertain whether the addition of a visual channel improves remote communication or whether audio-only communication is sufficient. Considerable literature exists that suggests that the addition of a visual channel is advantageous since one can pick up important non-verbal signals such as gestures and facial expressions. This question was set in order to determine whether these users agree that access to visual signals result in more successful communication. The literature is mixed in relation as to whether this benefits task outcome or the processes of communication. This will be discussed later, but here user perceptions are of interest.

A total of 45 participants stated that videoconferencing has advantages over the telephone, whereas only 3 participants stated that videoconferencing has no advantages over audio-only communication. User perceptions therefore reflect the communication literature in that there is a feeling that non-verbal communication ought to be beneficial. Although it is clear that the vast majority of participants believe that videoconferencing

has advantages over the telephone, many participants also noted that there are benefits associated with audio-only communication. The resounding view, however, is that there is added value to having access to visual information. The following table sums up all the advantages that were proposed with the frequency of each proposal made.

Table 3.2: Frequency of suggestions made for advantages that video-mediated communication has to offer over telephone.

Suggestion	Freq.	Suggestion	Freq.
1) Encourages familiarity	3	10) Display images	4
2) People like to see who they are talking to	6	11) Maintain sense of community	1
3) Facial expressions	6	12) View reactions	6
4) Share documents	3	13) More friendly	2
5) Collaboration is easier	1	14) Only for personal	1
6) Body language/gestures	10	15) Visual feedback/non-verbal	8
7) Forced into participation	1	16) Connect names with faces	1
8) Aids understanding	2	17) More than one can talk at once	1
9) More personal/human	3	**Total**	**59**

Participants noted a total of 17 different advantages that video communication has to offer over audio-only communication. Various advantages, such as having access to body language, were pointed out by numerous participants, however other advantages, such as being able to connect names with faces, were only mentioned by one or a few participants. It would therefore seem that some advantages are more prevalent than

others. These advantages are primarily related in some way to having access to visual signals, for example access to facial expressions, gestures and 'reactions.' So, user perceptions reflect what we know about human communication: visual communication signals matter, and therefore adding support to the non-verbal communication hypothesis. Also mentioned are the 'video as data' issues, such as displaying images and sharing documents. Access to visual communication therefore allows:

1) Access to crucial non-verbal signals.

2) Higher levels of social and human involvement (increased sense of social co-presence)

3) Other visual images, such as documents, can be shared and seen.

Each of these issues are discussed more fully as follows.

3.3.2A Non-verbal communication

So having access to non-verbal information is seen by users to be an advantage associated with video-mediated communication. Such non-verbal cues allow video-mediated communication to be a better alternative to face-to-face interaction than audio-only communication. All advantages proposed by the participants that fit into this category are listed below with the number of participants who stated each advantage in brackets.

1) Facial expressions (6)

2) Body language / gestures (10)

3) View reactions (6)

4) Visual/non-verbal feedback (8)

5) Connect names with faces (1)

6) Aids understanding (2)

Thirty-three comments, out of a total of 59, were made relating to non-verbal communication. The fact that such a high number of participants stated that access to non-verbal information is an important advantage for videoconferencing implies that such signals are crucial in human interaction and having access to them improves communication. In respect to frequency of comments made it would seem that having access to facial expressions (6) and body language (10) are major advantages associated with video-mediated communication. These comments correspond with Short et al (1976) and Williams (1977) who have noted that users prefer a view of the face when bargaining. It would seem that many of the non-verbal repertoires associated with face-to-face communication could be replicated in a videoconference (however, this is dependent upon factors such as the size and quality of the image). Being able to view an individual's reactions would also seem to be important in communication, as 6 participants have noted this as an advantage. When using the telephone reactions can be made verbally. If a person does not speak, however, it would be impossible to know how they have reacted to a particular piece of information. This would explain the increased number of verbal backchannels (for example 'mm-hmm') found in audio-only conversations by Boyle et al (1994). Without access to visual indicators of attention (for example head nods), users must signal verbally that they have understood the message. An individual's facial expressions is just one of many ways of determining how someone has reacted and is important in determining if someone has understood or not (Whittaker and O'Conaill, 1997). This sentiment is expressed by the following two participants' further comments.

"It is possible to gauge more easily what another person is saying, even thinking, from looking at the body language etc....perhaps in certain situations this may make the outcome more likely to favour one of the persons involved (e.g. negotiation) which might not otherwise have happened through the telephone."

"Yes it is better for making judgements about context and knowing what the other person really thinks...sometimes it is better not to know."

So it would seem then that access to non-verbal information helps individuals to better monitor understanding and attention. These perceptions are borne out in the literature. Indeed researchers such as Daly-Jones et al (1998) and Doherty-Sneddon et al (1997) indicate that participants in video-mediated interactions are better able to monitor understanding and attention than participants in audio-only interactions.

To sum up a quote from one particular participant:

"A view of physical characteristics and body language gives each of the participants another dimension of communicating."

What must be considered, however, is that eye contact has not been mentioned here as an advantage. From what we already know of the technology, natural eye contact cannot be replicated without the use of specialised mirrors ('videotunnels,' Smith et al, 1991). Therefore, although users agree that V.M.C is useful because it allows access to non-verbal signals, it is clear that not all elements of non-verbal behaviour can be

replicated in a videoconferencing environment. As will be discussed in further detail later in the chapter, not having access to eye contact can cause difficulties in communication.

3.3.2B Social factors

Another advantage that video-mediated communication has to offer over audio-only communication is that access to visual signals results in an increase in social involvement. All advantages that fit into this category are listed below with the number of participants who stated each advantage in brackets.

1) Encourages familiarity (3)

2) People like to see whom they are talking to (6)

3) Collaboration is easier (1)

4) Forced into participation (1)

5) More personal/human (3)

6) Maintain sense of community (1)

7) More friendly (2)

8) More than one can talk at once (1)

Eighteen comments, out of a total of 59, were made concerning the idea that access to visual images has an effect on the social factors of the exchange. Six participants stated that they like to see whom they are talking to.

"Communication over the video broke down barriers as you could see the person, it was also easier to have a face to face discussion rather than imagining what the person at the other end will look like."

It would appear that having access to what a person looks like facilitates interaction and increases feelings of social co-presence.

Three participants stated that video-mediated communication was friendlier, whereas another 3 participants stated that it was more personal.

"Depending upon the setting, number of people, topic etc., it can capitalise upon the 'person' or human factors of the exchange."

Video-mediated is regarded as being more 'personal' and more 'friendly' because one has access to an image of the other person. There is therefore a tendency to see the telephone as rather detached and isolated.

There are also benefits to communication, as well as being social benefits. One participant stated:

"In a situation of multiple participants on the video call, people are almost forced into active participation, or why be there."

Another participant noted that collaboration becomes easier because one has access to visual information.

"It (videoconferencing) requires interaction whereas the phone does not necessarily."

So, access to visual information goes some way to decrease social distance, corresponding with Argyle and Dean's (1965) equilibrium theory, in that access to visual images increases feelings of intimacy and therefore decreases feelings of distance.

3.3.2C Other visual signals

Seven comments, out of a total of 59, were made relating to videoconferencing allowing other visual images to be displayed. Therefore one advantage of V.M.C is that is allows for the sharing of documents or visuals. While not pervasive, these findings correspond with Gaver et al (1993), Anderson et al (1999) and Nardi et al (1993), who noted that participants preferred Teledata to views of the face and body.

The following advantages were highlighted:
1) Share documents (3)
2) Display images (4)

One participant in particular stated: "documents can be displayed, shared and highlighted on a common vantage."

Although there is considerably more evidence that users view access to non-verbal information as an important benefit of V.M.C, it is also clear that users consider images of documents to be important. While there is insufficient evidence to accept or reject

the video as data hypothesis, Anderson et al's (1999) discussion of the application's influence on the need for a view of the person (Telepresence) or a view of relevant objects (Teledata) is a good one. Telemedical applications, or engineering or training may be more reliant on Teledata (Egido, 1990, Nardi et. al, 1993). On the other hand purely social interactions among remote family members will benefit more from a view of the participants' faces rather than views of other images (Anderson et al, 1999). The relative infrequency of mention of video as data in the current sample may therefore reflect the sorts of tasks the current set of users employ video-mediated technologies for. Business meetings, for example, may require more socially orientated skills, for example negotiation or bargaining, for which access to images of the person are important.

3.3.2D Comments concerning video not having advantages over audio-only communication

Five comments were also made by those who believe that videoconferencing does not have advantages over the telephone and also by those who also believe that videoconferencing has its own advantages.

"I personally think that the telephone is good and has advantages of being detached."

"In group discussion yes. Probably not for one to one discussions."

The previous statement is reinforced by Daly-Jones et al (1998), who noted greater fluency in conversation when comparing group videoconferencing to group

audioconferencing. This effect was not realised in a comparison of one-on-one communication.

"No. The telephone is very easy to use, the voice quality is adequate for most purposes, the instrument is on the desktop, and we have learned culturally how to cope with the absence of a visual channel."

This statement is interesting as it relates to our adaptability to different communication media. According to this user it is possible to adapt to communication without visual signals. Adapting to technology does not necessarily imply that one is getting the best use out of it, and indeed the user states that audio signals are adequate for 'most purposes' perhaps indicating that there are some situations in which visual images are required. Nevertheless, the fact that this user believes that it possible to adapt to imperfect communication media is reassuring.

"Appointing spokesperson on telephone is as effective."

"I am not yet convinced. It may well be okay for brief discussion, but detailed argument is not very successful."

3.3.2E Discussion of question 2

From the suggestions made informants perceive that video has much to offer over audio-only communication. Although it is clear that audio-only communication also has advantages in certain contexts, for example users can become more detached from emotional content, video-mediated communication is associated with more advantages

due to the availability of visual images, especially visual communication cues like gesture and facial expressions. This feeds into another advantage; the social facilitation of V.M.C. Videoconferencing seems to encourage familiarity between users more effectively than through the use of the telephone. Video-mediated communication is noted as being friendlier and more personal. One informant reports that video-mediated communication capitalises on human or person factors, suggesting that there is a higher level of social co-presence than with the use of the telephone.

Evidence from the literature suggests that when comparing face-to-face dialogues with audio-only and video-mediated dialogues, the face-to-face conditions usually result in more speaker exchanges (turns), shorter turn lengths, and more interruptions (for example, Cook and Lalljee, 1972, Rutter and Stephenson, 1977, Cohen, 1982, and O'Conaill et al, 1993). A common interpretation of these findings is that face-to-face interactions are less formal and more spontaneous than audio-only and video-mediated conversations. So what could cause a lack of co-presence in video-mediated communication? The process of communication changes when it is mediated. Perhaps this change has socially related effects, for example less interaction may reduce the likelihood of seeing oneself as being part of a spontaneous exchange. Rutter (1987) has argued that videoconferencing attenuates the visual cues that are present in face-to-face interaction. In cases where transmission is less than ideal this can be a result of poor resolution, bandwidth or framerate.

Video-mediated communication is also regarded as superior to audio-only communication due to the availability of important non-verbal signals. A lack of cues may result in the people with whom one is communicating seeming less like real people

and more like objects (Williams, 1977). This depersonalisation can encourage people to become more self centred (Sproull and Kiesler, 1986) and exhibit antisocial behaviour (Siegal et al, 1985), but it can lead to decisions based more on facts than personalities (Williams, 1977). Informants report that they like to see whom they are talking to and can pick up non-verbal cues such as facial expressions and body language. From the data then it seems that the addition of a visual channel is advantageous, allowing for improved communication. Video-mediated communication is a step closer to face-to-face interaction than teleconferencing due to the inclusion of a visual element of communication. A review of the literature by Fish et al (1993) has indicated that video-mediated communication is more similar to audio-only communication than it is to face-to-face interaction, however, the informants here perceive advantages. Even with higher quality media, however, non-verbal communication is still attenuated as natural eye contact cannot take place, and dependent upon the size of the image body language may be difficult to read. All of these effects will result in an increase in social intimacy distance (Argyle and Dean, 1965). So attenuation of visual cues and the decrease in intimacy both seem to lower social co-presence. It should also be noted, however, that a number of studies have found no differences in the number of turns, turn length, and number of interruptions between face-to-face interactions and audio-only and video-mediated interactions (for example, Argyle et al, 1968; Rutter et al, 1981; Sellen, 1995). This would indicate that the process communication does not always change when communication is mediated.

There is also an issue in V.M.C relating to shared-space. When using videoconferencing equipment, for example, people only have access to audio and visual information and therefore cannot manipulate a common environment. This makes it impossible to pass

items to another, or even follow another's attention to a painting on the wall for example. This contrasts with virtual reality systems, where the users can share a common environment and can even manipulate joint virtual objects (for example Ishii and Kobayashi, 1992).

3.3.3 Question 3: Do you believe that videoconferencing systems can replicate face-to-face communication? Why?

This question was included to assess whether video-mediated communication is as effective as face-to-face communication considering the addition of a visual channel, or whether, as suggested, video communication will never replicate face-to-face communication because of the lack of personal contact.

Table 3.3: Total responses for whether videoconferencing can replicate face-to-face interaction

No	Yes	Maybe	Unanswered
30	10	5	3

It seems then that the majority of users believe that video-mediated communication cannot replicate face-to-face interaction. The reasoning behind these answers are listed below.

3.3.3A Comments by users who believe that video communication cannot replicate face-to-face communication.

"Many people do not feel at ease in front of a camera; cannot always see body language; cannot shake hands; cannot go for a beer."

"A handshake and eye to eye contact cannot be replaced."

"Face to face always required."

"Cannot go for lunch after a videoconference or build up personal relationship."

"Most effective with planned agenda and meeting leadership."

"Like saying machines could replace people."

"Lose element of personal contact and interaction."

"Only a distinct representation."

"Difference in psychological space...not the same degree of social interaction...can't pass someone a jug of water for example."

"Face to face much better"

"It can enhance what would normally be an audio call, but it can't replace live interaction."

"Not replace because humans still need the face to face meeting. Video cannot fully replace the socialising impact of face to face. Should be used to augment regular meetings where bringing people together is expensive."

"Face-to-face is critical for sizing up individuals in a common environment and balancing that against the spoken word and written document. It does not replace, but it can reduce the need for all communications to be in person."

"No it will never replace face to face. You need personal interaction to establish relationships once established these mediums may work."

"No you can be more informal in face to face situations, you can more easily talk only to selected individuals (if necessary in private) and you can put more pressure on people. The group dynamics are different."

3.3.3B Comments by users who agree that video communication can replicate face-to-face interaction.

Many users believed that video-mediated communication could replicate face-to-face communication successfully.

"Videoconferencing is face-to-face communication. Where the difference lies is with the texture and time delay. If you operate a quality facility, then I do not believe that the differences has any bearing on the positive applications and outcomes."

"Providing that a user is not trying to lip-read as delay makes this virtually impossible."

"Only the handshake is missing."

"For small meetings and you are still speaking face-to-face even though it's over a video screen, so therefore no loss of personal touch."

"Only if parties have a previous history and relationship. Not good for initial introductions and meetings."

3.3.3C Comments by users who agree that video-mediated communication may be able to replicate face-to-face communication.

Many users believed that video-mediated communication could replicate face-to-face interaction, but only if certain developments in the technology were made.

"When technology gets up to scratch."

"Only if a culture of acceptance develops."

"Not until the eye-to-eye thing is sorted."

"Not yet. Perhaps with more fine tuning (e.g. allowing eye contact to take place) and through longitudinal use of such technologies to familiarise users."

3.3.3D Discussion of question 3

One general comment illustrates that for certain applications people can achieve the same level of success with videoconferencing as with face-to-face interaction. "If we look at the outcomes for students who have used this system of communication for their Family Practice Clerkships we find no difference in their end of clerkship exams as compared to students taught in more traditional settings." One is compelled to question here the importance of face-to-face interaction in such a scenario. Students can possibly achieve high levels of success without attending lectures but rather through obtaining lecture notes off the web and reading the course material. However the important point here is that this informant thinks that there essentially is no difference between face-to-face and video-mediated communication.

A major theme that is discernible from this data is that nothing can replace personal contact. It seems that there will always be a need for face-to-face interaction, unless technology is invented that can truly replicate face-to-face interaction. This data, along with other research (Abbott et, 1993) would suggest that supplementing video-meetings with face-to-face meetings improves communication. The current study reinforces the idea that there is a hierarchy of communicative preference due to the richness of media. In other words, media with a high capacity for communication cues, rapid feedback, a personal focus are more likely to result in higher levels of co-presence (Dennis and Kinney, 1998). These findings also agree with Campbell's (1998) assertion that as bandwidth narrows, it becomes increasingly more difficult to complete complex social interactions. One informant reports that with video-mediated communication there is a loss of personal contact and interaction. Another informant reports that one cannot build a personal relationship over videoconference. One cannot "shake hands", "go for

a beer" or even "go for lunch." It seems that such personal contact is important in human communication and therefore a lack of personal contact causes difficulties. Just how problematic will depend on the type of tasks and goals that the equipment is being used for. Although it is clear that video communication does not as yet replicate face-to-face interaction we can attempt, however, to create a more personal and friendly environment for videoconferencing users. If it is possible to eliminate feelings of distance then it is also probable that many of the reported difficulties will also be reduced.

A noticeable hierarchy of communicative preference is apparent. The majority of informants in this sample report a preference face-to-face interaction to other methods of communication. Videoconferencing is regarded as an inferior means of communication in comparison to face-to-face interaction, however audio-only communication is the least favoured of all. These results reflect general trends in the literature, for example O'Conaill et al (1993) and Daly-Jones et al (1998) reported that users preferred using video-mediated technologies to audio-only technologies, and Anderson et al (1997b) indicated that users show a preference for face-to-face interaction over video-mediated communication. If we regard this communicative preference in terms of media richness then the reasoning behind such preferences are evident. Media richness dictates that communication is more satisfactory and effective when rapid feedback cues are available, there is a high capacity for communication cues and the medium is more personal (Dennis and Kinney, 1998). Figure 3.1 illustrates some of the fundamental differences between audio-only, video-mediated and face-to-face communication.

Figure 3.1: Three levels of communication

FACE-TO-FACE	VIDEO- MEDIATED	TELEPHONE
Audio signals	Audio signals	Audio signals
Visual images	Visual images	
Personal contact/ shared emotion		

As illustrated in figure 3.1, the main difference between face-to-face and video-mediated communication is in the personal focus of the medium. In other words, face-to-face communication allows for personal contact and a shared physical environment, whereas video-mediated communication does not. Some V.M.C systems however are characterised by poor audio quality (for example delay) and poor video quality. Therefore although V.M.C has access to audio and visual signals like face-to-face interaction, the quality may be reduced. Depending upon whether there is a delay, video-mediated communication should equal face-to-face in speed of feedback. A study by O'Malley et al (1996) has indicated that in cases where a delay is present between audio and video (in this case 500msec), poorer task performance is evident and a higher number of interruptions occur. Many video-mediated systems also employ compression techniques in order to increase the speed of transmission. Compression techniques, however, reduce the number of frames per second and the amount of information in each frame of video. The result of these techniques is low picture definition and jerky pictures. This indicates that in some cases V.M.C is also inferior to face-to-face communication with regards to the quality of visual and audio signals. Therefore, at a glance it would appear that V.M.C offers identical communication

facilities with respect to the audio and visual channels. Poorer audio and visual quality however results in poorer communication, and therefore (dependent upon the quality of the system), V.M.C does not realise the benefits of face-to-face communication.

One particular problem that has been noted in the data concerns co-presence. It has been suggested that it is easier to remain detached over the telephone than through the use of video-mediated communication. It has also been noted that levels of social co-presence are lower when using video communication as opposed to engaging in face-to-face interaction. Social co-presence seems to be affected by access to audio information, visual information and personal contact, although it is seems that personal contact is the most influential factor. It would follow that having access to all three of these elements of communication may result in higher levels of social co-presence as in face-to-face interaction. Levels of co-presence would be lower over video-mediated communication, as one would have access to visual signals and audio information only. Levels of co-presence would be lowest over audio-only communication, as one would only have access to the audio level of communication. Having access to fewer of the elements of communication can also be advantageous. If we consider the case of emotional detachment, mentioned also in the data, then the hierarchy would be the reverse of the familiarity example. Using the telephone would allow for higher levels of detachment, followed by video-mediated communication and finally face-to-face interaction.

Non-verbal communication and personal contact increase intimacy levels. An increase in intimacy has been noted to decrease the negative effects of lack of co-presence in video-mediated communication.

3.3.4 Question 4: Did you note any problems when using such technologies, for example breakdown in communication or not being able to get your point across?

Question 4 asked participants to discuss the problems that were encountered when using video communication technologies. The purpose of this question was to assess the most important problems with a view to compensate for these difficulties through training and advice on good practice. The results from this question are summarised in the following table, with the number of responses from each group for each problem that was reported.

Table 3.4: Frequency of suggestions for problems associated with videoconferencing usage.

Suggestion	Freq.	Suggestion	Freq.
1) Audio delay	5	13) Visible people dominate	1
2) Poor picture quality	5	14) Motion problems	1
3) Lose signal/networks go down	7	15) Turn taking difficult	1
4) No eye contact	3	16) Overwhelming	1
5) Unreliable	2	17) Too polite	1
6) Interruptions	3	18) Difficult to control cameras	1
7) None	14	19) Have to pass Mic about	1
8) Compatibility	1	20) Competition for Mic	1
9) Sound	11	21) Difficult to get point across	1
10) Feel isolated	1	22) Person not actual size	1
11) Establishing links	5	23) Aware of self presentation	2
12) Distracting	1	24) Silences run for longer	1
		Total	**72**

Problems associated with video-mediated communication are varied and seem to fall into two categories.

1) Problems caused by technological constraints.

2) Problems caused by differences in social and communicative processes in comparison to face-to-face communication

More statements made concerned technological problems compared to social/communicative problems. Out of a total of 72 comments, 14 participants stated

that there were no problems, 41 problems were related to technical issues and 17 comments related to social and communicative problems. It should be noted however that many of the social/communicative difficulties experienced in V.M.C are a direct result of technological constraints.

3.3.4A Technological constraints

All problems that fall into the category of technological constraints are listed below and include the number of times each difficulty was mentioned.

Audio delay (5)

Poor picture quality (5)

Lose signal (7)

Unreliable (2)

Compatibility (1)

Sound (11)

Establishing links (5)

Motion Problems (1)

Difficult to control cameras (1)

Mic problems (2)

Person not actual size (1)

With respect to technological difficulties it would seem that problems with sound are the major source of difficulty associated with video-mediated communication. Eleven participants noted problems with sound, stating that the sound in a videoconference can be too quiet or that the sound quality is not always of an acceptable standard. Five participants also noted that there is often a delay between the audio signal and the visual

signal, meaning that the picture and sound are not synchronised. As discussed, research by O'Malley et al (1996) reports more occurrences of interruptions and poorer task performance when a delay of 500msec is introduced. Other problems mentioned frequently included poor picture quality (5), signal loss (7) and difficulty in establishing links (5). Poor picture quality can be problematic, as users will be unable to pick up the subtleties of non-verbal communication, for example facial expressions, however, Daly-Jones et al (1998) have noted that the importance of facial expressions in most task contexts has been over-emphasised. It is clear that technological difficulties will result in many problems for users. Although users of V.M.C should be advised to purchase the best possible equipment, the focus of this research is to improve the use of V.M.C irrespective of the 'quality' of a particular system.

3.3.4B Social/communicative differences

All problems that fall into the category of social/communicative difficulties are listed below and include the number of times each difficulty was mentioned

No eye contact (3)

Interruptions (3)

Isolated (2)

Distracting (1)

Visible people dominate (1)

Turn taking difficult (1)

Overwhelming (1)

Too polite (1)

Difficult to get point across (1)

Aware of self presentation (2)

Silences run for longer (1)

Social and communicative problems caused by differing processes between face-to-face and video-mediated communication are mentioned less frequently than technological problems. Obviously many of these difficulties are also a direct result of technological problems. One problem in this area is a lack of eye contact, with three mentions. Even though one participant believes that a lack of eye contact is a major problem with videoconferencing, it was further noted that looking up at the camera to replicate a form of gazing helped to resolve this difficulty. Perhaps then it would be useful to advise users to look up into the camera during a videoconference. By looking up at the camera the user can give the impression that eye contact is taking place. This method may improve the perceived level of eye contact, however, it should be noted that this form of gazing is only one way and therefore cannot be regarded as mutual eye contact. Furthermore looking at the camera means that the gazer cannot simultaneously attend to the visual image of their partner, partly defeating the purpose of the video connection. Another justification for advising users to look up at the camera concerns social co-presence. A review of the literature and responses from the questionnaire data indicates that feelings of distance are more distinct in video-mediated communication than in face-to-face interactions. This finding can be explained in terms of equilibrium theory (Argyle and Dean, 1965). Essentially, there are many cues to intimacy, including touch, proximity and eye contact. When people communicate at a distance, many of these cues are restricted and therefore individuals find it more difficult to gain a level of intimacy with another person. Adding another cue to intimacy, in this case eye contact should therefore go some way to reducing feelings of distance.

Two participants noted feeling isolated during a videoconference. Again, this relates to the loss of social presence in V.M.C (Rice, 1993).

One participant noted that interaction alters during a videoconference and people become too polite. He stated that people find it more difficult "to let you know when you have screwed up and you would have more readily picked it up if you were there in person i.e., someone would have pulled you aside and explained." This statement is reflective of the formal nature of videoconferencing that has been found by researchers such O'Conaill et al (1993) and Sellen (1995).

Fourteen out of the 48 informants also perceived there to be no problems associated with videoconferencing. One particular participant stated, however that this is only the case "if the appropriate thought and planning is used in advance." Again, this comment indicates that videoconferencing is only 'problem free' if there is a certain level of structure. Other data, however, would lead us to believe that no matter how much thought and planning are put into a videoconference, the loss of personal contact inevitably implies that interaction will never be as successful as face-to-face communication. Although it is possible to have a successful videoconference in terms of content, and users may also have their need for a visual image satisfied, feelings of distance still result in lower levels of preference.

3.3.4C Discussion of question 4

The problems pointed out by users of video and audio communications seem to fall into two categories: technological difficulties and social problems caused by the fact that video communication does not replicate face-to-face communication in the sense that

personal contact cannot be achieved. This difference between face-to-face and remote communication technologies is further illuminated by the responses gathered for question 3 ("Do you believe that video or audio conferencing systems can replicate face to face communication? Why?"), as 30 informants replied with the answers "no" and only 10 informants replied with a "yes". Most of the answers given to this question focused on social interaction and demonstrated that video-mediated communication acts as a barrier to successful communication as one loses an element of personal contact. One participant even went as far as stating, "that would be like saying machines can replace people." Another participant replied "I think they get close but ultimately there is a difference in psychological space as there is some distance between the person and the other/s. There is not the same degree of social interaction as face-to-face communication for example passing someone a jug of water or borrowing a pen cannot happen." One particular buzzword was "representation," suggesting that video-mediated communication merely represents face-to-face contact, but is not equal to it. In essence this sums up all we know about video and therefore the purpose of this thesis is to determine how to work with an imperfect but potentially useful media. Those participants who replied "yes" seem to suggest that a replication of face-to-face communication is only possible under limited circumstances, for example if the users have met before, if the videoconference is conducted with a small group, or if the problem with eye-contact is resolved.

Videoconferencing does appear to have its advantages. Firstly it allows the implementation of visual images, which cannot be achieved when using the telephone. Furthermore videoconferencing appears to be associated with a lack of personal contact and therefore a loss of emotional/social interaction, this can be useful in certain

instances, for example where the meeting is of a sensitive nature. The main advantage of videoconferencing, however, seems to be that it can bring people together who would not normally get the chance to meet because of obstacles such as distance. Such examples include distance learning and interviews. The implementation of video-mediated technologies is further advantageous because it reduces costs and time. The findings of the current study would suggest a need for a suitable method for improving feelings of distance and unfamiliarity.

The limitations of remote communication technologies suggest that in order to improve these methods of communication, procedures should be followed to allow users to adapt more successfully to the technology. To gauge users responses question 5 was included ("If you could make any changes to conferencing systems for the purpose of more successful communication what would they be?") The answers to this question were varied but seemed to focus largely on technical alterations. The focus of this research, however, will stay clear of technical issues and rather focus on advising or training users in improving social/conversational elements of communication

3.3.5 Question 5: If you could make any changes to conferencing systems for the purpose of more successful communication what would they be?

Considering the difficulties that are encountered by videoconferencing users, question 5 was included to assess proposals to improve the use of such technology. The proposed solutions with the frequencies of each suggestion are included in the table on the following page.

Table 3.5: Frequency of suggestions for proposed solutions to videoconferencing problems

Proposed solution	Freq.	Proposed solution	Freq.
1) Good quality video	3	18) Broadcastable virtual presence	1
2) Large T.V screens	5	19) Less participants	2
3) Providers accountable for losses	1	20) Eliminate echo	1
4) Simplify Whiteboard	1	21) Used for right occasions	1
5) More controllable sound	1	22) More availability	2
6) Voice activated control	1	23) Common interfaces	1
7) Key points repeated	1	24) Better audio	2
8) Real time video	7	25) Training	1
9) Cheaper	3	26) Point out limitations	1
10) Use it over phone line	1	27) Wider visual coverage	3
11) Automatic muting	1	28) Crisper focusing	1
12) Better filters for audio	1	29) More cameras	1
13) More reliable	5	30) Replicate eye contact	3
14) More compatible	1	31) Monitor for each individual	2
15) Greater bandwidth	8	32) Time to familiarise with equipment	1
16) Mic for each participant	1	33) Assess self presentation	1
17) Create virtual commons	1	34) More flexible camera	1
		Total	**67**

Comments are very disparate, but primarily centre on improving visual access (32/67). Five participants note that large monitors would be a method for the improvement of communication. The following comments were made concerning this suggestion: "show person in true size," "to identify non-verbal cues," and "so it doesn't seem like you are watching a television and the subject is removed from reality." These suggestions reflect the literature in that users indicate a preference for large screens, up to the point where the image is actually life size (Inoue et. al, 1984). Larger screens would essentially increase access to non-verbal communication and potentially increase social co-presence.

One participant notes that an improvement would be to create a virtual commons "where all the users could participate and feel as if they were at the same location. We have attempted to do this in our point-to-point administrative meetings by placing the VC (videoconferencing) cameras and monitors at the end of our conference tables at each site. The illusion of a large table is created. We all feel this facilitates communication." This suggestion seems to illuminate the idea that there is a lack of social co-presence in video-mediated meetings. The suggestion that "placing VC (videoconferencing) cameras and monitors" at the end of conferencing tables helped reduce this lack of social co-presence indicates that there are working methods to improve this problem. This particular method, referred to as 'media spaces' (Angiolillo, 1997) provides a spatial layout to the communication akin to that found in same room meetings and is preferred by users over other types of set-up (Sellen, 1995).

One participant states that a technological improvement would be to broadcast a virtual presence. "Actually the model for what I am talking about was described in one of

Isaac Azimov's robot stories. People's virtual presence was broadcast to one another. All nuances of communications: sound, visual cues, smells and body language were broadcast."

One participant notes that more cameras are necessary. One participant states that this would "give more context (e.g. showing the room in which the other person is sitting)." This seems to be another suggestion for improving the lack of social co-presence in video-mediated communication

Finally one participant states that more flexible cameras are needed in order that the image of the users' full body can be picked up.

Although monitoring which system each participant used is difficult, direct conclusions can be made from the answers provided in these questionnaires. As far as video communication is concerned it would seem reasonable to suggest that certain technical standards should be set in order to allow for the highest possible communicative success. The system should be of the highest quality for optimum effectiveness. This would include a microphone that could pick up all speakers distinctively and clearly; no delay on the sound; and a large screen with crisp focusing and high quality resolution. Indeed, O'Malley et al (1996) express the need for V.M.C systems to work with no delay for more effective communication. As far as extras on the system, these could obviously be tailored to suit the number of people using the technology and also the purpose for which the technology is being used. One suggestion that arises from this group of users is that, if a system were being used for small groups to communicate it may be useful to have a separate monitor for each individual in the conference. In

essence this particular set-up is akin to the Hydra set-up, which consists of a number of screens (dependent upon group size) with a camera mounted below each monitor and a speaker mounted below each camera. Experimental research suggests that participants favour the Hydra set-up to voice-activated switching modes and Picture in a Picture set-ups. Participants list reasons such as Hydra supporting selective gaze and listening and simultaneous conversations, that more closely reflect real life conversation (Sellen, 1995). This preference for the Hydra set-up can also be explained in terms of co-presence. The Hydra set-up creates a spatial layout indicative to a same room meeting (media spaces) and therefore may therefore help to reduce feelings of isolation.

Targeting the systems to the users and having the highest possible quality of system could indeed be the perfect scenario, however it is also the case that cost factors and technical limitations would prevent this from being a working solution. In essence users need to work with the equipment that they have available, and therefore users need to adapt to their technology rather than adapting technology to suit the needs of the users. Many suggestions can be put forward for training users to adapt to such technology. Video feedback could be employed so users could be become more aware of how they interact during a videoconference with a view to improving non-verbal and verbal interaction. One could even allow the participants to manipulate the technology directly so feelings of involvement are increased and in turn higher levels of confidence. Participants could even be given an overview of the technology that is being used so users feel less intimidated when using the equipment. As one informant reported in the questionnaire, video communication is unlike face-to-face interaction in the respect that there is a difference in "psychological space." Perhaps then to improve the effectiveness of such communication it would be wise to counteract this lack of intimacy.

3.3.6 Question 6: Please add any further comments.

All participants were asked to add any further comments if they wished to. The following comments were made:

"VC (videoconferencing) is an excellent method of communicating especially in the NHS. People from all corners of the country can get together to make decisions on business issues literally at the touch of a button."

"The biggest restriction on the evolution of VC (videoconferencing) as a primary communications tool is attitudes to technology and a reluctance to change the culture of drive to meeting. Future generations exposed to the technology in the classroom will expect to be able to communicate this way."

"Communication must occur in some way, shape or form. Not face to face is better than not at all."

"I think video linking is here to stay, so we must improve technology to make it better. Until then we should not expect too much from it."

"Videoconferencing has been useful but it is by no means a perfect form of communication."

"Video communication has disadvantages too. It's more invasive of privacy than the telephone. It's less convenient, because you have to face a camera."

3.3.7 General discussion

It would be unwise to make comparisons between all participants as all the different participants that took part use or have used remote communication technologies for different purposes and also there is no standardisation with respect to the systems being used. Furthermore, participants from one particular group may report a problem with video-mediated communication that does not apply to participants from another group. Essentially, it is highly likely that the problems with video communication are not only related to the kind of system being used, but also the kind of application the system is being used for. For example, participants who use a system with high bandwidth are less likely to report delay as problem in comparison to participants who use systems with low bandwidth. Participants who use systems for distance learning applications may also be less likely to report problems with turn-taking as communication is essentially in one direction. Problems with turn-taking may be more prevalent to participants who use the technology for two-way communication (for example, business users), and even more so for group communications. For this reason, care must be taken with regards to the advice that is given in order to improve communication, as this advice may not be directly relevant to all types of users.

It should also be noted that 11 participants had only used video-mediated equipment on one occasion. Furthermore, these participants had taken part in an experiment, and therefore the communication situation was a contrived one. In this sense any problems reported by such users may not be relevant to individuals who use such technology in real world settings. This is also evident in the fact that the video equipment was of good quality (high bandwidth, with no delay and large images), which is not entirely

representative of the equipment which is being used in every day settings. We can, however, view the problems and even proposed solutions more generally as it is the intention of the research to develop a holistic approach to solving these difficulties irrespective of set-up, equipment or even the use that the equipment is being put to. Therefore it is intended that this thesis will focus on devising techniques which will improve communication over a number of different set-ups and for a variety of different applications.

Although it is clear that technical issues are important, social difficulties are also encountered in the use of video-mediated communication. If we view social problems in terms of equilibrium theory (Argyle and Dean, 1965) then the majority of problems are caused by a lack of personal contact, resulting in users not being able to touch or make eye contact. Two of the social problems associated with V.M.C by this sample of users are known to cause difficulties with communication. The first difficulty stems from a lack of social co-presence between users. Another is the lack of eye contact. Both of these problems are a direct result of a lack of personal contact in video-mediated communication. Perhaps therefore eye contact lowers intimacy distance and increases co-presence. The focus of this research will be to concentrate on these two difficulties.

One could view the lack of co-presence as a direct result of participants not having access to the more intimate elements of an exchange. In order to counterbalance this effect it may be advisable to allow users to meet face-to-face before any meetings took place over videoconference. It is hypothesised that this practice will result in higher levels of familiarity between users, which will help to close the intimacy distance

between participants. Certainly, this particular suggestion comes out of the current qualitative data (discussed fully in chapter 4).

It would be impossible to replicate eye-contact over videoconference without making technical alterations, for example a video-tunnel (Smith et, 1991), however it may be advisable to train users to look up at the camera at certain moments, thus giving the impression to the other participant that eye contact is taking place. Although mutual eye contact could not be achieved, a sense that an interlocutor is looking right at you may be beneficial in increasing feelings of intimacy and providing important conversational functions such as emphasis (discussed fully in chapter 5).

In some respects the findings from this study are limited, however in many ways they consolidate much of what we already know about video-mediated communication. This particular study was interested in gathering qualitative perceptual data with a view to taking these perceptions further through the use of empirical studies. It may have been beneficial to gauge user perceptions further, for example a quantitatively analysed multiple-choice questionnaire. Such a questionnaire may have been useful, for example, to test whether users of the technology regard the approaches taken on for the empirical studies in the following chapters as being useful. It was considered to be more important however to focus on an experimental approach where particular variables could be manipulated. Essentially this thesis focuses on methods which can be used to improve video-mediated communication, and for this reason a further questionnaire would be unable to answer the research question and in many respects would be surface to requirements.

An interview was also conducted with two experts in the area of video-mediated communication. This interview also aimed to consolidate what we already know about video communication, however it also offers an alternative view point to those expressed by the participants in the questionnaire.

3.4: Study 1b: Expert analysis of videoconferencing

An interview was conducted with two senior members of a videoconferencing supplier, considered 'expert' due to the years of experience they have had working with V.M.C equipment and their considerable knowledge of systems and applications. Questions were chosen to reflect problems associated with video communication as discussed in the literature. Once completed, the full interview was transcribed and themes were discussed from the data.

For a transcript of the full interview please see appendix 1.

Interview - Expert analysis of videoconferencing

Date - 29/07/1999

"Worldwide Videoconferencing" - Donald West and Bill Howell (names have been changed to ensure anonymity)

Seven key issues relating to the use of video-mediated communication were noted.

The key issues will be discussed individually under the following headings:

1) Experience / confidence.

2) Importance of audio.

3) Quality of images.

4) The importance of familiarity.

5) Advantages / disadvantages of video-mediated communication.

6) Body language / Non-verbal communication.

3.4.1 Experience / confidence

According to the experts there is a relationship between experience with video-mediated technologies and confidence. Essentially, the more one uses the technology, the more confident one becomes. Furthermore, the experts assert that higher levels of confidence result in higher levels of efficiency and improved levels of communicative success.

Bill Howell: "I think the good news is that videoconferencing tends to be one technology which once you've done it for fifteen minutes it's sort of much less intimidating than it was before, because again the purpose of what you are doing is actually on content and when people are articulate about their content they tend to forget that fact that they are on T.V."

Here Bill Howell seems to illustrate the idea that at first people find videoconferencing off-putting, perhaps due to its novelty. According to Bill Howell a short period of time using the equipment will result in the lower levels of intimidation, as the content of what is being discussed essentially becomes more important than the images on the screen. This sentiment, however, would lead one to question what the purpose of having these images is.

If the content is indeed the most important element in this type of communication then surely one could argue that using the telephone would suffice. Certainly one could argue that the audio aspect of communication is the most important: it would be difficult to communicate without it. What we must consider here, as discussed earlier, is that even though visual and personal elements are less important for getting your point across, the combination of all three elements seems crucial for communication to be as effective as possible. Indeed, the literature in this area indicates that non-verbal information not only portrays emotional information, but also communicates a large amount of propositional information (for example, Evans and Rubin, 1979). It could be argued that one would be able to get their point across more successfully if communicating audio-only as opposed to with access to visual images only. The literature in this field, however views successful communication as an integration of both linguistic and non-verbal elements (Beattie, 1980; Clark and Brennan, 1991; Goldin-Meadow, Wein and Chang, 1992; McNeil, 1985; Weiner, Shilkret and Devoe, 1980): a point reinforced by the questionnaire data. In other words video-mediated communication is preferable and more effective than audio-only communication.

Bill Howell: "Videoconferencing can be somewhat intimidating to start with which inhibits your natural way of doing things, but appearing on television...the more you use it....the less intimidating the whole thing becomes and the more you know confident you become."

Once more this quote illustrates that confidence is an important factor in effective video-mediated communication. The interviewee makes it clear that with experience one gains higher levels of confidence. Although confidence may be an important factor,

other issues also need consultation. Gaining higher levels of confidence may go some way to improve the use of video-mediated technologies, however whether confidence resolves the lack of social co-presence experienced by users is debatable.

Bill Howell: "One of the problems is that people are really self conscious about their own image....on the T.V. People are extremely self conscious about seeing themselves on T.V....The people who really like videoconferencing are the right show-offs who think they are really good on video telly and want to broadcast to everybody that they are real smart, that they look the you know erm the, they've sort of got a presence."

People are self-conscious about another person being able to view their image on a monitor, which is strange considering that one can also be seen in a face-to-face interaction. In one experiment in the current thesis a participant stated "I haven't done my hair." Video-mediated communication is an unusual method of communication and it seems that people become more aware of self-presentation than in face-to-face interaction. Perhaps there is an association with television or it may even be the case that users believe that they are being recorded. Perhaps this also relates to the loss of the social dimension to communication, as a result of a lack of co-presence. Indeed, Fish et al (1993) indicate that video-mediated interactions are more formal than face-to-face interactions. Perhaps this formality, much like an interview for example, leads one to become more aware of self-presentation.

Bill Howell: "I think you don't need to do too much training erm to allow people to use it properly. What you do is you've got to build a little bit of erm confidence if you like, and I don't mean through training it's probably confidence of use and part of that clearly

is training but part of it is that you are not going to look really, really stupid you know to the person at the other end because everything's OK"

According to this expert analysis training is not as important as allowing users to simply sit down and use the technology. Perhaps this is the case, however, a number of users from the questionnaire data had considerable experience with V.M.C, but still reported difficulties.

Bill Howell: "Making productive use of the video communication (is a problem) and people are a little bit wary whether on a live call or on air if you like of making a mistake and you draw a parallel again of the newsreaders, when they get it wrong it's so bloody obvious that it appears to be so embarrassing that you can't possibly believe how this bloke could end up showing his face again."

This quote seems to illustrate the earlier idea that the visual element of communication becomes more powerful in a videoconference. On television it is expected that newsreaders are perfect in their delivery, as they are communicating to millions of people. One makes mistakes all the time in face-to-face communication, however, unless one is communicating in front of large groups of people the mistake seems unimportant and may even go unnoticed. In a live broadcast on television this mistake is perhaps amplified because it is in view of so many people. It is possible then that this association is passed onto a videoconference irrespective of the number of viewers. Also, evidence from the literature indicates that movements can become more noticeable because of the lack of peripheral cues in V.M.C (Angiolillo, 1997).

It would seem that the more people use videoconferencing systems the more confident they become and therefore more efficient users of the technology. It would be a reasonable suggestion then that practice with videoconferencing equipment is one method for improving communication, however data from the questionnaire would suggest that more improvements are needed for such communication to be as effective as face-to-face interaction, especially considering that many experienced users still reported residual effects.

3.4.2 Importance of audio

According to the interviewee the audio content of a videoconference appears to be more important than the visual element. This however leaves the question of why bother going to the expense of having a videoconference when a telephone call could suffice? If people are investing in videoconferencing equipment one would have to assume that the visual element does indeed play a crucial role.

Bill Howell: "Most of the content of a video conference happens to be audio. Believe it or not, about 85% from our estimations: in a sense that you may convey a picture and as long as you are real and I sort of try and .....have some sort of personal relationship with you to judge whether you are someone who is genuine, maybe someone who I believe in erm, is not dissimilar from the...initial erm personal contact you make by meeting people and providing you get to a level which is good enough erm in terms of the visual and the audio quality...then suddenly the audio content or the or what I call the intelligent content takes over, because you know the picture is unintelligent, a picture is unintelligent, but the content of audio is actually erm the stuff that decisions are made on."

The interviewee here seems to be suggesting that video-mediated meetings are not dissimilar to face-to-face meetings, implying that one can gain a level of familiarity with someone as successfully over a videoconference as in a face-to-face interaction. It is also evident, however, that many users believe that such relationships cannot be formed over videoconference, but rather you need to have met the person before hand, or had some sort of personal contact with them.

Bill Howell: "What happens is, as you become more experienced as a videoconference user then the environment you are in becomes fundamentally more important and particularly the audio content because if you can have a conference and feel as though you know your emphasis perhaps in, in linguistic terms is accepted by the far end it really doesn't matter if you are laughing or crying."

Although offering verbal feedback on the telephone would suffice in terms of having one's content accepted, the difference with a videoconference is that visual feedback can be used, for example with head-nods. In cases where visual signals are unavailable, it therefore follows that individuals use more verbal backchannels (for example 'mm-hmm) and consequently speech is longer (Boyle et. al, 1994). Bill also indicates that as long as what someone says is accepted then the emotional state of the person is unimportant, essentially access to facial expressions is of little consequence. Surely, however, the emotional content of the dialogue is vitally important. Whether someone is laughing or crying changes the outlook of what is being discussed completely. Indeed, research would suggest that facial expressions play a crucial role in social interactions (Argyle, 1988), helping one to decipher speech (Summerfield, 1972), and

monitor understanding and attention (Ekman and Friesen, 1975; Whittaker and O'Conaill, 1997).

Bill Howell: "It's really the content that starts to become quite important regarding the audio level and the audio quality rather than necessarily the body language, but I also think that it's the focus to the extent that body language if it's missing starts to become very obvious rather than if it's there."

This quote seems to emphasise the idea that body language serves little purpose in video-mediated communication. According to the experts, the audio content seems to be the most important element of communication. Indeed, V.M.C literature suggests that good quality audio is crucial for successful communication (for example, Angiolillo, 1997; Watson and Sasse, 1996; Hollier and Voelcker, 1997). The expert also indicates, however, that if body language is missing then it becomes obvious. In a sense this implies that body language serves no other purpose in a videoconference than to illuminate that you are conversing with a real person. Considering that research indicates that non-verbal behaviours are important in human communication, perhaps it is a logical assumption that it is less of a case of body language being unimportant in V.M.C, but more of a case that technical difficulties (for example delay and poor quality images) making receiving body language very difficult. Also, it is clear that many psychologists believe that non-verbal communication isn't crucial for certain processes (for example Beattie (1981) indicated that non-verbal communication isn't important in regulating turn taking).

3.4.3 Quality of images

Although Bill indicates that the audio element to the exchange is the most important, evidence from the literature would indicate that having a high quality visual image is very important for communication (for example, De Greef and Ijsselsteijn, 2001). Visual images are obviously an essential element of communication. The advantage that video has over the telephone would be negated if the quality of the image is so poor that body language and subtle non-verbal cues could not be picked up. It is interesting then that this individual's beliefs go against the psychological research in this field. This is a clear indication that user perceptions do not necessarily match reality.

Bill Howell: "Because there is a benchmark, T.V, where audio and video is brilliant you know in comparison, if you don't get to that benchmark, if you don't get within to, close to that benchmark with videoconferencing, what ends up happening is people don't use it and therefore the impression of it is that it isn't a useful business tool, because they haven't reached a certain level of capability and that is almost exclusively in my opinion is down to you know the use of the product as well as the quality of the product in the first place."

As mentioned earlier, it seems clear that certain technical standards would be useful as far as having a successful and productive videoconference. Although it is clear that recommending users to purchase the highest quality equipment may be useful advice, what must be considered is that even if a certain technical standard is adhered to problems in communication will still exist. The highest quality equipment cannot, for example, reproduce personal contact. Furthermore, Bill seems to be saying that quality

is less of an issue than practice, although Bill does indicate that individuals fail to adopt the technology if visual images are poor, a clear indication that visual signals matter.

Bill Howell: "We are just looking face-to-face here so the frame rates that we are exchanging in other words the perception of jerkiness if you like of the picture is very small, because there's not a lot of movement. So if you actually compare it to a T.V image the image that we are actually erm exchanging is good enough for a videoconference; good enough for a business meeting because it isn't particularly movement orientated. Now there again lies, if you like, a technique of videoconferencing. If I actually want to get rid of the fact that you are moving your hands around which could deteriorate the picture I could just zoom in on your face and therefore techniques of using videoconferencing can actually improve, if you like, the perception of the video quality simply by using a few tricks of the trade."

Here Bill indicates that 'tricks of the trade' can be employed in V.M.C to improve communication. The aim of this thesis is to discover 'tricks of the trade' to help improve the use of video-mediated technologies. Therefore, as discussed, two new 'tricks' will be implemented in the two experimental chapters that follow.

3.4.4 Importance of familiarity

Data from the questionnaire and from the literature implies that being familiar with an individual reduces intimacy distance and therefore decreases feelings of distance between locations. To further understanding of whether this method could indeed help, the experts were asked to comment on this suggestion.

Bill Howell: "In general terms I know what Donald looks like; I sort of know what he's going to do if I ask him a difficult question and so in a way it's not a fact that the body language is there it's a fact that the body language is sort of expected."

This quote seems to suggest that communication between the two experts is improved because they have had personal contact, and are therefore more familiar with one another. Upon being asked if this improved understanding as a result of increased familiarity through face-to-face contact Bill stated "Yes if you like. It's the negative aspects. If he doesn't do it I want to know why."

To clarify this point further the experts were asked if face-to-face contact prior to a videoconference increases levels of familiarity and in turn facilitates communication. In response to this question Donald West stated, "Yes absolutely, absolutely yes."

It would seem then that the experts agree that being familiar with someone through personal contact actually improves communication, although it should also be noted that this decision was not a result of free reporting, but rather an answer to a direct question.

Further supporting the hierarchy of communication Bill stated "it is very easy on the telephone to say no because you are actually talking, you are talking to an abstract voice." As mentioned earlier in the data, it is easier to remain detached when using the telephone, as there are fewer elements of communication available. In this respect audio-only communication can be advantageous, however, the addition of a visual channel does give videoconferencing its advantages. In some respects, this contradicts Bill's earlier statement concerning the lack of a role for non-verbal communication.

Bill Howell: "it is very difficult for me to say no I'm not going to answer any more of your questions because we have sort of got a little bit of a personal relationship, however short, it's a lot better than sitting on a telephone doing it and because you don't see me you don't know me erm on the telephone you could walk past me in the street and not give me any aggravation."

In this sense videoconferencing offers a higher level of social co-presence than the telephone. Access to visual signals seems to improve feelings of togetherness. It adds to the social necessity of the situation, perhaps making it more likely that people will conform to social norms and etiquette.

Bill Howell: "I think the most important thing about videoconferencing is it's an unusual way of meeting people, but it actually does mimic pretty closely a personal meeting"

Bill Howell stated that video "mimics pretty closely a personal meeting," which would imply that video is close to face-to-face but not quite as effective.

Although it is clear that the interviewee believes that bonds can be formed over videoconference it is also evident that he believes videoconferencing to be more efficient when users have met through real contact. Perhaps then it is the case that people can form a level of familiarity after extensive use of video-mediated communication, however levels of familiarity would be higher if users had met for the same period of time face-to-face.

3.4.5 Advantages/ disadvantages of V.M.C

Upon being asked to discuss the advantages and disadvantages of video-mediated communication the following points were mentioned.

3.4.5A Advantages

Bill Howell: "What we are not trying to do is stop people flying or going and meeting face to face, we are not trying to do that. What videoconferencing does is allow people to meet more often, so in other words the guy who went to New York six times a month now only needs to go three times a month because the other three meetings they can meet over the video."

Donald West: "Basically it is saving them time, management time, money because of air-fare and accommodation etc., and it improves their quality of life because people are spending more time at home and they are not spending eight hours sitting on a jet twiddling their thumbs."

As mentioned in the questionnaire video-mediated technology cuts cost and time and allows people to meet when geographically dispersed.

Donald: "Probably more focused actually than meeting people face to face."

Perhaps this focus is a result of users having to pay to use the technology and also having a limited amount of time to use the equipment. Also, one could argue that this

effect may be a direct result of increased formality (in comparison to audio-only). One could assume then that users want to make the best possible use of their time. It may also be the case that the lower levels of social co-presence allow for higher levels of emotional detachment (in comparison to face-to-face). For certain types of meetings it may be argued that this is beneficial.

Donald West: "In a sense the more charismatic you are the better you are on T.V."

This implies a subtle difference between communication via videoconference and face-to-face communication. As far as videoconferencing is concerned it is more than being able to see the person. Effective communication will take place when a presence is projected. During face-to-face interaction you are able to convey your point and hold someone's attention through tactile means and shared sensory input within the shared environment. One does not have this advantage in a videoconference, therefore it is necessary to be able to 'work the camera' and hold someone's attention. Looking at the camera is one method that may be employed to project oneself more effectively.

Bill Howell: "I am looking at a very focused position you know, I am looking at you in a room on a clearly small flat T.V and therefore my focus is completely on what you are doing at that end and it's very difficult for me to be distracted by things at this end, but you know that potentially will take away, if you like, the focus of the meeting. So you find videoconferencing actually increases the meeting focus because you are actually looking, to coin a phrase, at you know a sort of porthole, it's the porthole effect you know."

This may increase focus to the extent that all other sensory input that would normally be available in face-to-face communication is not present, therefore focus is directed solely onto the image of the person speaking. However, this would not necessitate that surrounding stimuli are not going to distract from the interaction taking place. It would be more difficult to detect if someone were not paying attention during a videoconference, because eye contact is not possible and therefore looking away is no indication of whether or not information is being attended to. Those who take part in a videoconference do not have a shared environment, thus do not have access to what is going on around the other, and therefore will be unaware if the other is attending to something else.

In response to question of whether an interviewee would have been better off meeting face-to-face as opposed to a videoconference Donald West states: "No, because they were appreciative of the fact that he was a forward thinking erm consultant and he was saving himself time and money by meeting them over a videoconference link and it worked in his favour, so they were quite happy..."

As mentioned earlier, the audio content of a videoconference seems to be the most important element. This would lead one to question the validity of using a videoconferencing system over a telephone. We must assume that some advantages must exist or why else would people invest time and money into this equipment?

Bill Howell: "But I think what it is, is it is very clearly a group and a team, if you like, meeting technology or meeting business application which probably isn't reproduced anywhere else."

3.4.5B Disadvantages

Bill Howell: "You'd much rather sort of pin someone down face to face and say 'how much is this going to cost me' rather than tap in perhaps to erm to a computer you know talking to someone on video and tap in and they get the numbers wrong."

The expert seems to agree here that personal contact is favoured over a videoconference. People are more comfortable asking questions to another person face-to-face. Videoconferencing brings about a level of 'performance anxiety' because it creates a formal atmosphere, almost as if performance is being assessed and recorded.

Bill Howell: "In terms of their application and I certainly know from experience that the take-up of kiosks in particular a trial was done by a financial institution which shall remain nameless, they put a kiosk in each of their, in each of their branch offices in the Diss area and just about everybody went to a person at a desk. The predominant reason for that is that a very large proportion of the people that were going into them were old people with savings you know people taking a few pounds out erm from savings accounts and predominately the people who were interested in doing it were young. It would seem there's an obvious connection."

Bill Howell: "Would much rather talk to a real person even though they had to queue up because it was perceived as being quite difficult to interface to. I don't think necessarily that it was the video component that was difficult it's actually the buttons you had to press to make it happen."

Donald West: "Some of the, well we get a lot of phone calls in London from customers, in fact they usually ring Slough, but the technical centre's where Bill is, but if they can't get through to Slough they'll come through to us and I have to be blunt about this. A lot of the errors that we get that we deal with have got nothing to do with the equipment at all. It's either users erm inexperience or they don't understand the keypad because the supplier, not normally us, but a, or it could be someone who has come into the room who has not been trained wants to know which button to press, why has the picture disappeared."

Bill Howell: "So, so, the, the I guess the practical things here are erm the network isn't reliable and continues not to be reliable erm the video systems tend to be a little unfriendly in places and I can see that's getting better."

The main disadvantages of videoconferencing then seem to be that personal contact is not available and also technical limitations, for example the network going down.

3.4.6 Body language / Non-verbal communication

The interviewees were asked to comment on non-verbal communication in video-mediated communication. A suggestion is that videoconferencing has advantages over the telephone because visual images are available, however, it also seems to be the case that many forms of non-verbal communication, such as orientation, are in some way altered due to the fact that video-mediated communication, unlike face-to-face interaction, cannot allow for personal contact.

Bill Howell: "You know if I'm angry, he's been party to that image, well that's not necessarily obvious on an audio conference. In other words I couldn't judge, if you like, the whole of the content and the whole of the participants in the meeting if it wasn't on video."

It seems then that on a basic level video-mediated communication does allow some forms of non-verbal communication, in this case facial expressions. Although gestures can be performed over video-mediated technologies, it is also evident that problems exist when performing gestures during a videoconference. Many users of video-mediated technologies state that they find it unnatural to gesture to a screen. Bill Howell, however, states that gesturing in a videoconference becomes more natural and easier with experience.

Another problem with non-verbal communication concerns the lack of eye contact. Although many forms of non-verbal communication are possible when using video-mediated technologies, the favoured set-up of videoconference means that mutual eye contact is not possible. One particular participant in the questionnaire, however, stated that he believed looking up at the camera, thus producing a level of video-mediated gazing, helped to improve this situation. Upon being asked their opinion on this matter the experts also agreed that this would indeed help to improve communication. Once more this decision was the result of a direct question as opposed to free reporting.

Many users state that there are more interruptions in a videoconference. One suggestion is that a lack of eye contact could be the reason for more interruptions as it has been suggested that eye contact is used to regulate turns. Upon being asked as to whether a

lack of eye contact could result in problems in turn taking Donald West replied: "most people either make a signal that they want to interrupt a flow and that happens in videoconferencing you'd be surprised. You know even if people in London want to interrupt a guy in Chicago, the guy in London will go like that (Hand up) hang on guys and even though, he's saying to them I'm going to interrupt here and he'll put his hand up and say hang on John let me just come in there you know, so it's amazing how people adapt and, and just get used to it."

It seems then that people adapt to the technology and find other methods to make up for the shortcomings of video-mediated communication. If this were the case then it would follow that other methods may be employed to make up for many of the other difficulties that are caused by the lack of personal contact.

3.5: Chapter summary

Through the use of a questionnaire targeting current users of video-mediated technologies and an interview with two experts, a number of key issues have been found in the data. One key issue relates to problems with current technology. Participants in the questionnaire study and the experts seem to agree that technical limitations are a major factor affecting communication. It would seem that individuals compare video-mediated communication to television and therefore television becomes the benchmark with which users evaluate V.M.C. Audio quality and visual quality are not always as of an acceptable standard. Technical limitations are abundant, for example motion problems and problems with the microphone (competition, having to pass it about etc.). These technical limitations often have knock on effects for communication. Audio delay, for example can result in poorer task performance and a higher number of

interruptions than when no delay is present (O'Malley et. al, 1996). This research also suggests that video-mediated communication is often used out of necessity rather than as a useful tool for communication. One user states, "not face-to-face is better than not at all." Although many researchers have argued that technology should be of the highest quality for successful communication (for example no delay, high quality video), it is the focus of this research to improve communication irrespective of set-up or quality. Moreover, Doherty-Sneddon et al (1997) have indicated that even when video-mediated technology operates at optimum capacity (for example, no delay, high quality audio and video and eye contact) it still does not deliver the same benefits as face-to-face co-present interaction. This would necessitate feasible methods for improving the use of video-mediated communication, other than ensuring high technical standards.

This research also indicates that video-mediated communication is novel and off-putting. According to the experts, experience with the technology results in higher levels of confidence, which in turn results in improved communication. Irrespective of whether experience has positive effects, the fact that experienced users in the questionnaire still reported difficulties suggests that there is still room for improvement even with seasoned users. Also, experience does not help individuals who are likely to use the technology infrequently. One way to make V.M.C more like face-to-face interaction would be to the close intimacy distance between interlocutors. This suggestion is reinforced by the abundant references to decreased social co-presence in V.M.C. Indeed, the lack of co-presence in V.M.C seems to be a pertinent issue to the participants in the questionnaire and has obvious ramifications for the overall preference for face-to-face communication. A lack of social co-presence also seems to have

negative communicative effects. One method that could be employed to close intimacy distance would be to increase levels of familiarity between users. According to the equilibrium theory (Argyle and Dean, 1965) intimacy is expressed through a number of cues, for example eye contact, proximity and touch. As it is the case that many of these cues cannot be expressed in V.M.C, it may therefore be beneficial to give participants the opportunity to express intimacy. One way to achieve this is to supplement video-mediated communication with face-to-face meetings. Indeed, Abbott et al (1993) have shown the benefit of employing such a practice. The question is therefore, will such face-to-face meetings have knock effects for V.M.C and help close intimacy distance? In line with this theory, it is expected that higher levels of familiarity will result in higher feelings of social co-presence. If familiarisation increases levels of co-presence then we would expect higher levels of satisfaction and better levels of communication than those who are unfamiliar: indeed this comes from perceptions reported here and the psychological literature. For example, higher levels of social co-presence increase the speed of a person's task performance on a simple task (Levine et al 1993; Sana, 1992), and can even effect task outcomes in a distance learning environment (Paechter et al, 2000).

Video-mediated communication also does not support normal gazing behaviour and further to this, the participants in the questionnaire reported that is a problem. Most forms of non-verbal communication are possible with video-mediated communication. For example, gestures may be performed and received and facial expressions can be viewed and interpreted. It should be noted, however, that this is dependent upon a number of factors, for example camera angle, image size and image quality. It has been noted, however, that eye contact and gazing behaviour is not possible. The reason for

this is because the usual set-up has the camera placed above or to the side of the monitor. What ensues then is an image of a person, who when looking at the monitor (or the image of the person whom they are talking to), appears as if to be looking down or off to the side. Many users find this off-putting, as it almost appears as if gaze is being averted. In face-to-face communication people associate gaze aversion with a number of social effects, for example with someone who is not telling the truth (although in fact, highly motivated liars control this to cover up their deception) and low levels of gazing are associated with someone with little confidence. Irrespective of these associations gazing is an important element of human communication and has been linked with conversational turn taking (Kendon, 1967; Goodwin, 1981) and providing emphasis (Fry & Smith, 1975; Otteson & Otteson, 1980). Gaze can also serve as an indicator of understanding, knowledge and attention (Beattie, 1981; Clark & Marshall, 1981; Rutter, 1984), as well as a number of other communicative functions (often in conjunction with other parts of the face, for example in the expression of emotion). It therefore follows that an inability to use gazing behaviour will have a number of social and communicative drawbacks. Levels of gaze also relate to co-presence, as increasing levels of gaze have been shown to close intimacy distance (Argyle and Dean, 1965). Replicating gaze in a videoconferencing environment then should be beneficial for a number of reasons, including the possible increase of social co-presence. One way to accomplish this would be to follow the suggestion of one participant in the questionnaire. He suggested that looking up at the camera helps to resolve the lack of eye contact in V.M.C. This would give the impression that gazing, although not mutual, is taking place. It would therefore be of interest to test whether this method could go some to improve video-mediated communication.

Overall, evidence from the literature, questionnaire data and the expert analysis of videoconferencing has lead to the decision that increasing familiarity and replicating eye gaze in V.M.C may be beneficial. These practices should decrease social distance and therefore facilitate communication and task performance. These practices will be evaluated in the following two experimental chapters.

Chapter 4: Closing Intimacy Distance in V.M.C through Participation in Preliminary Face-to-Face Meetings.

4.1: Introduction

4.1.1 Aims and objectives

The literature reported in the previous three chapters indicates that in some ways V.M.C meets the requirements of face-to-face interaction. The availability of a visual channel results in a number of benefits. For example, access to visual information allows participants to monitor understanding and attention more effectively (for example Daly-Jones et al, 1998), and complete tasks more successfully when an element of negotiation is required (for example, Short, 1974; Short et al, 1976). Although the addition of a visual channel to communication results in a number of benefits, a review of the literature also suggests that face-to-face communication is superior to V.M.C. One reason for this is that video-mediated communication does not take place in a co-present environment, and therefore users often report a loss of social co-presence. This is partly a result of physical remoteness and partly a result of the attenuation of visual cues.

Data from study 1b (questionnaire) also suggests that familiarity can decrease this effect. In other words it was proposed that knowing someone better helps to reduce psychological distance. This suggestion makes sense in terms of 'equilibrium theory' (Argyle and Dean, 1965). In everyday interactions individuals try to establish a certain level of intimacy with another person. Furthermore, intimacy is expressed by factors such as proximity, eye contact, smiling and personal topics of conversation (Argyle, 1969). In situations where participants cannot express cues for intimacy it becomes increasingly difficult to establish a comfortable equilibrium. According to Argyle (1969), proximity cues are particularly important and therefore during a videoconference, where many of such cues are unavailable, it can be expected that it will be difficult to establish a comfortable level of intimacy. Indeed, research suggests that if certain cues are unavailable, then other cues are

used to compensate. Porter et al (1970), for example showed that individuals stand physically closer to people who have their eyes shut, and therefore use other cues to close 'intimacy distance' (Argyle, 1969). V.M.C research (for example Doherty-Sneddon et al, 1997) also indicates that over-gazing occurs during video-mediated interactions which allow eye contact to take place. One suggestion for this finding was that this behaviour compensated for the lack of proximity cues. Furthermore, Doherty-Sneddon and McAuley (2000) propose a link between equilibrium and co-presence. In situations where individuals find it difficult to establish a comfortable equilibrium, a lack of co-presence will also be noticeable. This suggestion is a result of the finding that video-mediated interactions allow children to be more resistant to leading questions (Doherty-Sneddon & McAuley, 2000). Essentially, feelings of distance resulted in higher levels of confidence and independence, a theory also reinforced by an increase in anti-social behaviour.

Such research would therefore suggest that closing intimacy distance would result in increased participant perceptions of co-presence. This can be achieved through reducing psychological distance between participants or by allowing participants to express more cues for intimacy. One method that may be employed to reduce psychological distance is the implementation of initial face-to-face meetings. Essentially, knowing someone better should reduce feelings of isolation. Furthermore, higher levels of co-presence should result in improved task performance, improved levels of satisfaction and improved evaluations of others using the media. The justifications behind these assumptions will be discussed later in the chapter.

In order to achieve this aim, two separate pools of participants completed two tasks: a 'collaborative map reading task,' and a task involving the evaluation of an individual's

suitability for a job as a member of bar staff. The 'collaborative map reading task' provides extended dialogues that are content-controlled, allowing for an analysis of conversational features of the dialogues, for example the number of turns employed, and the number of Checks and Aligns (Kowtko et al, 1991, previously discussed in chapter 1) used by participants. These measures are of particular importance as they allow for an analysis of task efficiency and communicative success. Participant evaluations will also be key to performance measures, and in both tasks user perceptions and evaluations will be analysed. The use of such a measure reflects the fact that user perceptions and preferences are key features in V.M.C literature.

4.1.2 Previous research on social co-presence and establishing equilibrium

In chapter 3 (questionnaire analysis and expert analysis of videoconferencing), it was asserted that media richness has a positive relationship with levels of social co-presence. In other words, media with a high capacity for multiple communication cues and rapid feedback are more likely to result in feelings of being in the same room as another person. Gale (1989) noted that as bandwidth increases, so do feelings of social presence. In particular he noted that higher perceived levels of co-presence were achieved with a shared whiteboard with audio and video capabilities, in comparison to lower levels of perceived co-presence with shared whiteboard only. One explanation for this finding is that individuals in the higher bandwidth conditions were granted the ability to express more cues for intimacy. For example, in comparison to audio-only communication video-mediated interactions allow participants to use non-verbal signals (for example facial expressions). According to Dennis and Kinney (1998), richer media will also result in improved and more satisfactory communication. This is borne out by Campbell (1998) who declares that, in an overview of the literature, there is general consent that as

bandwidth narrows (for example from face-to-face communication to audio-only communication), communication becomes less suitable for intricate social exchanges and interpersonal communication. Anderson (1999) also asserts that there is a general reported preference by users for media that are richer. Experimental evidence for the effects of media richness on task performance and user satisfaction comes from Anderson et al (1997b) who found a significant difference for task performance for face-to-face communication over video-mediated and audio-only communication, but no difference between video-mediated and audio-only communication. According to Anderson et al (1997b) this result informs us that the visual processes that facilitate communication in face-to-face interaction may not transfer to video-mediated communication. It would seem then that face-to-face communication results in better task performance than both audio and video interactions. A post task questionnaire, however, did indicate that users found it easier to perform the task over video-mediated than over audio-only communication and easier to perform the task over face-to-face communication than both audio-only and video-mediated communication. Users also indicated a preference for video-mediated communication over audio-only communication. Again, these findings can be explained in terms of equilibrium theory (Argyle and Dean, 1965). Evidence from the literature suggests that by allowing individuals to express more cues for intimacy, higher levels of social co-presence are formed (Doherty-Sneddon and McAuley, 2000). Furthermore, higher levels of co-presence are formed in richer media and richer media also result in improved task performance and improved participant perceptions. It can therefore be hypothesised that reducing intimacy distance will result in improved task performance and improved user perceptions.

Research also suggests that there are differences in user perceptions of others when presented over video in comparison to face-to-face presentation. In a report prepared on the first six court trials in New Zealand to use CCTV, Whitney and Cook (1990) noted that children were judged more on facts than personalities when they were presented to judges via videoconference. With the use of questionnaires, the researcher gauged opinions relating to the implementation of video technologies in courtrooms. Many participants noted that using CCTV lessened the 'realness' and impact of the testimony. This sentiment is borne out by Pipe and Henaghan (1996) who noted that many professionals perceived that distancing children from the courtroom lessened the emotional impact of the evidence. One explanation that has been proffered for this finding is that the children felt less intimidated and subsequently had a more composed manner (Whitney and Cook, 1990). Whitney and Cook did however also note that brief face-to-face 'meetings' had significant effects on later V.M.C perceptions. When met face-to-face prior to video-mediated presentation the children were judged more on personality than facts (Whitney & Cook, 1990). This would indicate that the influence of 'personal' factors would be greater when someone is met face-to-face. Similarly, Williams (1975a) indicated that preference ratings for people who met for the first time over an audio-only link were lower than those expressed for individuals meeting face-to-face or via a video link. This finding would seem to indicate that participants view other individuals more favourably when more cues are available. This is borne out in the literature, with an abundance of evidence, which suggests that individuals are evaluated on their use of non-verbal behaviour.

The questionnaire data study 1a also agrees with this literature. Users reported higher levels of social co-presence in video-mediated interactions than during audio-only communication. Co-presence is affected by available cues for intimacy. One participant

for example stated that videoconferencing is a more personal method of interaction than audio-only communication: "depending upon the setting, number of people, topic etc., it can capitalize upon the 'person' or human factors of the exchange." The human or person factors that this participant was referring to could indeed be cues for intimacy (for example non-verbal behaviours). On the other hand, the participants also noted higher levels of co-presence in face-to-face interactions than video-mediated communication. As discussed, many of the cues for intimacy cannot be expressed in a video-mediated environment (for example touch or natural eye contact). For videoconferencing users, a lack of personal contact can have negative communicative effects. As one users states: there is a "difference in psychological space...not the same degree of social interaction...can't pass someone a jug of water for example."

As previously discussed, there are a number of ways in which performance can be measured. Researchers may evaluate user perceptions, undertake measures of task efficiency or even analyse aspects of spoken delivery. As with researchers such as Olson, Olson, Carter and Storrosten (1992), and Strauss and McGrath (1994), this thesis will attempt a multi-dimensional approach to measuring performance. Therefore, in this chapter performance was measured in a number of different ways. User perceptions and evaluations were measured through the means of questionnaires; task performance measures were undertaken (for example, time to completion and the number of turns taken by participants); and dialogues were analysed using Conversational Games Analysis (Kowtko, 1991).

4.2: Study 2: Face-to-face familiarisation and person perceptions across a video link

4.2.1 Introduction and aims

This study aims to test the effects of face-to-face familiarisation on participant perceptions of a partner across a video link. Strangers were familiarised with one another through face-to-face contact and were then judged in a video-mediated context. This condition was compared with a video-mediated familiarisation condition. It was expected that face-to-face familiarisation would close intimacy distance more effectively than video-mediated familiarisation due to the available cues for intimacy in face-to-face interactions. Furthermore, the closing of intimacy distance would result in increased perceptions of social co-presence. It was also expected that participants in the face-to-face familiarisation condition would rate their partners more favourably on a questionnaire specifically designed to assess user perceptions of job suitability and personality attributes. Essentially, participants would be judged more on personalities than facts. This theory is line with research by Whitney and Cook (1990) which suggests that the influence of 'personal factors' is greater when someone is met face-to-face.

4.2.2 Participants

Thirty-two Open University students were opportunistically sampled from the University of Stirling. Fifteen male participants and 17 female participants took part in the study. All participants had no prior experience with V.M.C. No details of age were recorded. Two confederates were also used in the experiment. The confederates were both male, and aged 21 and 24 respectively. Confederate A had no prior experience working in a bar. Confederate B had worked at the University Union bar for approximately 2 years.

4.2.3 Design

A between subjects design was used, with user perceptions (ratings) the dependent variable, and type of medium used for initial meetings the independent variable. Half of the participants were assigned to the video-mediated meeting condition and half the subjects were assigned to the face-to-face meeting condition. Further to this, half of the video-mediated condition group observed the interview with confederate one, and half with confederate two. Half of the face-to-face meeting group observed the interview with confederate one, and half with confederate two.

4.2.4 Materials

In room 1, a colour monitor (JVC TM-14EK(B)) was mounted in a wooden box, with a video camcorder (Sony CCD-TR2200EPAL) placed directly above the monitor. A microphone was placed to the right of the monitor, and video and audio quality were as high as achievable in the lab. The monitor and camcorder in room 1 were connected to room 2, adjacent to room 1, and with the exact same set-up. Monitors in both rooms were 14 inches in size.

Each participant (and confederate) was distanced approximately one metre from the monitor and the scope included the participant's face and upper body.

A questionnaire assessing user perceptions was used, with ten questions in total and a rating scale from 1 to 5 (1 being the lowest possible score and 5 being the highest possible score). Therefore the ratings ranged from 'not at all' to 'very.' The questions were as follows:

1) On a scale of 1 to 5, how friendly was the interviewee?

2) On a scale of 1 to 5, how confident was the interviewee?

3) On a scale of 1 to 5, how well did the interviewee communicate?

4) On a scale of 1 to 5, how intelligent was the interviewee?

5) On a scale of 1 to 5, how honest was the interviewee?

6) On a scale of 1 to 5, how suitable was the interviewee for the job?

7) On a scale of 1 to 5, how likely would you be to employ the interviewee?

8) On a scale of 1 to 5, how prepared was the interviewee for the interview?

9) On a scale of 1 to 5, how close was the interview to face-to-face communication?

10) On a scale of 1 to 5, during the interview, to what extent did you feel that the interviewee was in the same room as you?

4.2.5 Procedure

Prior to the experiment, two interviewees were recorded being interviewed for a job as a member of bar staff at Stirling University Union. Interviewees faced the camera but did not look directly into it, rather slightly below. These confederates were offered no monetary incentive to take part in the experiment. The confederates were asked a series of questions ranging from prior experience to suggestions for improving nightlife at the bar. Both confederates were given the list of questions one-day prior to the interview, in order that they could prepare answers for each question. Confederates were allowed to draw upon actual experience or if they preferred could fabricate information. It was considered to be of little importance whether or not the confederates were indeed suitable for the job, but rather whether they were deemed more suitable after being met face-to-face. Participants in both conditions observed interviews that were pre-recorded as opposed to a live video link. This method was employed in order to maintain a level of consistency in the performance of the interviewee. Although this method may have implications for the

findings (discussed later in this section 4.2.7), it was assumed that the use of a pre-recorded interview has much in common with a live link. Although there are obvious differences, interviewees would still be judged in an environment where a lack of cues and a low level of social co-presence were evident. After the interviews had been recorded the second stage of experimentation could take place.

Subjects entered a room and were asked to sit down. Participants were told that the nature of the experiment was to test interview techniques and memory. Participants were assigned to one of two conditions. In the first condition subjects met the interviewee face-to-face prior to watching the interview. In the second condition the subjects met the interviewee over videoconference prior to watching the interview. Both confederates were used in face-to-face and video-mediated conditions. In order to have a suitable reason for meeting the interviewee prior to watching the interview, subjects were informed that a memory test for knowledge of bar prices would need to take place. Participants were asked to test the interviewee's knowledge of bar prices by simply asking a list of 15 questions (for example "how much does a vodka and coke cost?), placing a tick next to the question if answered correctly and a cross if answered incorrectly. Participants were told to ask the questions in the order that they appeared on the sheet and were also asked not to give any feedback to the interviewee. Both interviewees rehearsed the answers to the questions prior to the experiment, however, two of the questions were intentionally answered incorrectly in order to increase ecological validity. The interviewees were also asked to take their time when answering so that the situation would appear more realistic. Both confederates were naïve as to the aims of the experiment.

After this stage of the experiment had been completed participants were asked to observe a pre-recorded interview. All participants were made aware that the interview was pre-recorded. After watching the interview the participants were given a questionnaire to fill in relating to perceptions of job suitability, personal attributes and perceptions of closeness of the video-mediated interview to a face-to-face context. Participants were specifically asked to rate the interview performance, although it was expected that the initial meeting context would influence ratings. The experiment took approximately 20 minutes to run. After filling in the questionnaire, participants were debriefed as to the deception and given a general overview of the research.

4.2.6 Results

Using a two-tailed hypothesis, independent measures t-tests was carried out between the face-to-face meeting condition and the video meeting condition for all questions on the questionnaire.

Results indicate that participants view the interviewees as being more suitable for the job, more honest, better communicators, and more intelligent after being met face-to-face as opposed to via video. Results also indicate that the participants viewed the video-mediated interview as being closer to a co-present face-to-face context when the interviewees had been met face-to-face. No significant differences were found between mode of initial meeting for evaluations of friendliness, confidence, employability, preparedness, and perceptions of the interviewees being in the same room.

Table 4.1: Questionnaire data: Independent t-test scores for differences between video meeting and face-to-face meeting conditions, with mean scores and standard deviations in brackets.

Question	Mean Score (S.D)	Significance
1) Friendliness	Face 3.9375 (.85) Video 3.7500 (1.06)	$t(30) = 0.55$, $p > 0.05$ **NS**
2) Confidence	Face 3.8750 (.62) Video 4.1250 (.96)	$t(30) = 0.88$, $p > 0.05$ **NS**
3) Communicative ability	Face 4.4375 (.81) Video 3.5625 (1.09)	$t(30) = 2.57$, $p < 0.05$
4) Level of intelligence	Face 4.3750 (.50) Video 3.7500 (.68)	$t(30) = 2.96$, $P < 0.05$
5) Honesty	Face 4.5625 (.63) Video 3.8750 (.72)	$t(30) = 2.88$, $p < 0.05$
6) Job suitability	Face 4.6250 (.50) Video 4.0625 (.93)	$t(30) = 2.13$, $p < 0.05$
7) Employability	Face 4.3125 (.60) Video 3.8750 (.89)	$t(30) = 1.63$, $p > 0.05$ **NS**
8) Preparedness	Face 4.3750 (1.02) Video 4.5000 (.82)	$t(30) = 0.39$, $p > 0.05$ **NS**
9) Closeness to face-to-face interaction	Face 3.9375 (.68) Video 3.0000 (1.21)	$t(30) = 2.70$, $p < 0.05$
10) Level of social co-presence	Face 3.4375 (1.09) Video 2.8750 (1.36)	$t(30) = 1.29$, $p > 0.05$ **NS**

Questions 1 to 5 represent evaluations of personal attributes. Questions 6 to 8 represent evaluations of employability. Questions 9 and 10 represent evaluations of how similar the video-mediated interviewee was to a co-present face-to-face context. From the above table it can be seen that significant differences have been found on the questionnaire data between the conditions. Participants rate the interviewees as being significantly more suitable for the job ($t(30) = 2.13$, p< 0.05) when the initial meeting takes place face-to-face as opposed to taking place over videoconference. Participants also rate the interviewees as being more honest when the initial meeting takes place via the face-to-face modality ($t(30) = 2.88$, p<0.05). Results indicate that perceptions of communicative ability are affected by medium of initial meeting, with ratings being significantly higher for interviewees who were met face-to-face ($t(30) = 2.57$, p<0.05). Interviewees were regarded as being more

intelligent when met face-to-face before the interview (t(30) = 2.96, p<0.05), as opposed to being met over videoconference. Finally, participants rated the actual video-mediated interviewee as being significantly more like face-to-face interaction after meeting the interviewee face-to-face (t(30) = 2.70, p<0.05). In contrast, no significant differences were found in terms of perceptions of friendliness, confidence, employability, and perceptions of the interviewees being in the same room.

4.2.7 Discussion

Perceptions of personal attributes differed between the two conditions. After being met face-to-face, the interviewees were regarded as being significantly more intelligent, honest and better communicators. The interviewees were also regarded as significantly more suitable for the job after being met face-to-face. These findings would suggest that the impact of 'personal factors' is stronger in face-to-face communication than in video-mediated interactions. The findings from this study therefore add weight to the claims of Whitney and Cook (1990) that initial face-to-face meetings result in participants being judged more on personalities than facts in a proceeding video-mediated context. These findings are also in line with the non-verbal communication literature, which suggests that individuals are judged on their use of non-verbal behaviour. In this case confederates were evaluated in a face-to-face context, and this evaluation influenced perceptions in the proceeding video-mediated condition. For example, the interviewees were unable to employ gazing behaviour in the video-mediated condition. Indeed, the confederates were requested to focus their attention on the monitor, thus giving the participants the impression that they were looking in a downward direction. Research suggests that the use of gazing behaviour can influence perceptions of intelligence and honesty (for example Hemsley & Doob, 1978). In the video-mediated meeting condition participants would

have been unable to make eye contact with the interviewee, and therefore this may have influenced perceptions of honesty. In contrast, participants in the face-to-face meeting condition would have been able to make eye contact, and therefore this would have influenced perceptions of honesty in the proceeding video-mediated context. Tankard (1970) also revealed that direction of gaze influences perceptions of job suitability. Participants have been shown to rate people in photographs as more deserving of a job or a higher salary when they looked straight ahead as opposed to looking in a downward direction. This may go some way to explain the perceptions of increased job suitability and intelligence in the face-to-face meeting condition. Furthermore, Exline and Eldridge (1967) noted higher evaluations of frequent gazers. Essentially, having access to non-verbal behaviours in a face-to-face context influenced later video-mediated perceptions, especially if we consider that certain non-verbal signals cannot be transmitted over video.

The video interview was perceived as being closer face-to-face communication after the interviewees had been met face-to-face (how close was the interview to face-to-face communication?). This indicates that initial mode of meeting has an effect on perceptions of co-presence. The fact that participants regarded the interview as being closer to a face-to-face context implies that an initial face-to-face meeting is one way to reduce intimacy distance in video-mediated communication. These findings are in line with Abbott et al's (1993) claims that face-to-face meetings should be used to supplement videoconferencing. One problem with these findings however concerns the possibility that participants are rating the confederates in the initial pre-interview context and not the confederates' performance in the interviews themselves. Although this may indeed be the case, the fact that there were differences in participant perceptions between conditions would indicate that users are still judged differently when communicating over different media. In this

case participants were viewed more favourably when communicating face-to-face than when communicating over video-mediated technologies.

One problem with using multiple dependent variables concerns the possibility of a type I error. Indeed one is bound to find levels of significance by chance if a sufficient number of measures are used. Although this is a concern the fact that 8 out of a possible 10 questions (even when non-significant) were in favour of the face-to-face condition would seem to indicate that there is an effect here. Although a number of comparisons revealed no significant differences (perceptions of confidence, friendliness, employability, preparedness and perceptions of the confederate being in the same room), a noticeable trend is evident in the data. In terms of friendliness, employability, and perceptions of the confederate being in the same room, non-significant numerical differences were found in favour of the face-to-face condition. In this case then there is a possibility of a type II error as these questions have failed to achieve significance. One possible reason for this would be that there is insufficient power to discriminate between the two conditions. This study may have therefore benefited from more than one rating per construct. Non-significant numerical differences were however found in favour of the video condition for perceptions of preparedness and confidence.

The evidence from this study indicates that initial face-to-face meetings can alter user perceptions of an individual when judging them in the context of video. These findings call into question the validity of conducting interviews over a videoconference, a more and more frequent procedure. Problems with such a procedure may ensue if interviewees are judged in different contexts (for example one interviewee is interviewed face-to-face and another over video). One argument for this finding is that initial face-to-face meetings have resulted in intimacy distance being closed and therefore psychological distance has

been reduced. Another explanation for the findings is that face-to-face contact has resulted interviewees being judged more upon their personalities than by the facts at hand. What is important to note is that the same individual is rated more highly when in essence the only difference is the initial medium of presentation. Participants felt that they 'knew' the interviewees better when having met them face-to-face, and there were some criteria that this reduced the negative effects of video-mediated communication. Although these results are interesting, they do not inform us as to whether face-to-face familiarisation has an effect on task performance. In other words, can initial face-to-face meetings facilitate later transfer of information across a video link? In order to test this, the following study was undertaken.

4.3: Study 3: The effects of closing intimacy distance on collaborative task performance

4.3.1 Introduction

This study investigated whether face-to-face familiarisation prior to a videoconference would improve task performance and communication on a collaborative map reading task. Three conditions were investigated in which initial meeting context (prior to the task) was manipulated: face-to-face meeting condition, a video-mediated meeting condition, and a no-meeting condition. In study 2 (face-to-face familiarisation and person perception across a video link) the differences between face-to-face and video-mediated meetings were assessed. The results suggest that face-to-face meetings are superior with regards to measures of participant perceptions. In the following study a third condition (no meeting) was introduced in order to test whether video-mediated familiarisation is better no form of familiarisation. It was expected that meeting a partner before the task would reduce psychological distance and in turn improve communication between users. Specifically, it was hypothesised that face-to-face familiarisation would result in improved user

perceptions of others using the media and general levels of satisfaction, more efficient communication (for example fewer turns and less interruptions), and improved task performance (specifically levels of collaboration) in comparison to the other two conditions. It was also expected that video-mediated familiarisation would lead to some benefits over the no-meeting condition.

4.3.2 Participants

The sample comprised of an opportunity sample of 72 participants from Stirling University. All participants were split into pairs; therefore there were 36 pairs in total. Half the participants were male, and the other half were female. Participants were randomly split into three groups: face-to-face meeting condition, video-mediated meeting condition, and no-meeting condition. Twelve pairs were allocated to each of the three meeting conditions. All participants were unfamiliar with their partner, and all participants had no prior experience with video-mediated technologies. Written informed consent was gained from each participant. This was so the interactions could be recorded for later analysis.

4.3.3 Design

A between-subjects design was used. The independent variable was the method of original meeting (face-to-face, video-mediated and none), and the dependent variables were user perceptions, task performance and measures of task efficiency.

4.3.4 Materials

Same set-up as previous experiment.

4.3.5 Collaborative map reading task

A 'collaborative map reading task' was specifically designed to encourage negotiation between participants (please see appendix 2 for example of completed map and appendix 3 for accompanying list). The reason for this is that the literature indicates that negotiation style tasks are particularly sensitive to finding differences between media. The 'collaborative map reading task' involves both participants having to plan a route together (on a map of a town centre with a number of shops), picking up five items from a shopping list along the way. The participants, however, have two different priorities: participant one must complete as short a route as possible, whereas participant two must complete the route spending as little money as possible. The map was constructed in such a manner that in order to get the shortest possible route a lot of money would be spent and in order to spend the least possible amount of money a long route would be completed. This is possible because all the cheapest items are located on the far end of town, the mid-priced items in the middle of town and the most expensive items nearest to the starting point. A significant negative correlation was found between route length and money spent (r = - 0.442, P< 0.01). In other words, if a long route were favoured then the participants would spend less money and if a short route were favoured then participants would spend more money. Therefore, to successfully complete the task participants need to collaborate to find a route that suits both of their needs as best as possible.

4.3.6 Procedure

The first phase of the experiment was different for each condition. The face-to-face and video-mediated meeting conditions both completed a 'getting to know you' style task, however, the face-to-face meeting group completed this face-to-face, whereas the video-mediated meeting group completed this over the videoconference. Participants in the no-

meeting condition did not complete this initial task. The 'getting to know you' task took the following format. Participants were asked to firstly find out the name of their partner and write it down on a piece of paper. After this had been completed the subjects were asked to find a word for each letter of their partner's name that appropriately described their partner: so for example the word could be a hobby, character trait, or even favourite food. An example was given to the participants involving the experimenter's own name so that there would be no confusion over how the task was to be completed. The experimenter exited the room and gave the participants 10 minutes to complete this task. Irrespective of whether or not the task had been fully completed the experimenter returned to the experimental room after 10 minutes and began the next phase of the experiment. The first phase of the experiment was not assessed.

The second phase of the experiment was the same for all three conditions. Firstly the participants were told that they were about to complete the 'collaborative map reading task' and were given instruction sheets (1 for each participant in the pair) containing information on the nature of the task, how to complete the task, and their priorities (see appendix 4 for list of instructions).

Participants were then asked if they had any questions about the task and a few important points were clarified. Participant 1, who's objective was to get the shortest route was also informed that he/she should attempt to be as close to the starting point as possible at the last shop visited. Both participants were informed that they had as long as they needed to complete the task, but they should remember that if they wished to they could collaborate in order that the route suited both of their needs as best as possible. After the participants were happy with the instructions, they were told to begin and the dialogues were recorded

and timed from start to completion. After completion of the task all participants were asked to fill in a questionnaire, which was designed to specifically target perceptions of how well participants collaborated. The following questions were assessed on a scale of 1 to 5 (one being the lowest possible score and 5 being the highest possible score), so ratings ranged from 'not at all' to 'very much.'

1: Please rate how much you enjoyed using the videoconferencing equipment.

2: If it were your choice would you use this particular method of communication again?

3: How successfully would you rate the communication process between yourself and the other participant?

4: Please rate how well you achieved your aim.

5: Please rate how well you collaborated with the other participant.

Participants were also asked to add further comments if they wished to do so.

4.3.7 Dialogue coding

Transcribed dialogues were scored using Conversational Games Within Dialogue (Kowtko et al, 1991), specifically the number of Checks and Aligns used by participants. A Check "checks self understanding of a previous message or instruction by requesting confirmation directly or indirectly; makes sure that a complicated instruction is understood" (Kowtko et al, pg. 4, 1991). An example of a Check would include "so you want me to move up four inches?" An Align "checks the other participant's understanding or accomplishment of a goal; elicits a positive response which closes a larger game; checks alignment of both participants' plans or position in task with respect to goal; checks attention, agreement, or

readiness (Kowtko et al, pg.4, 1991). An example of an Align would include "yeah?" or "O.K?" meaning "are you following me?"

Transcribed dialogues were firstly coded into turns. A turn began at the moment a participant started speaking and was completed at the point at which the next participant began speaking. Further to this transcribed dialogues were coded into Moves and Games. Individual Moves are grouped into single dialogue units called Games. A Game can be defined simply by the goal that they serve within the interaction itself. All Check and Align Games were coded and then individual moves were counted. Therefore the total number of Checks and Aligns in each dialogue were tallied for further analysis.

Doherty-Sneddon et al (1997) found a significantly higher number of Align games in audio-only interactions than in video-mediated interactions. The researchers explained this finding by indicating that participants in the audio-only condition were unable to use visual cues to monitor understanding effectively (for example facial expressions to signal puzzlement). In this experiment it was expected that participants would be better able to monitor understanding and attention after face-to-face familiarisation. In other words, participants in the face-to-face familiarisation condition would use fewer Align and Check games. This expectation can be justified with a higher level of rapport being formed after face-to-face familiarisation. Consequently participants will feel less need to check that their messages have been understood or that they understand their partner's messages.

The reliability of the coding of the dialogues into Checks and Aligns was checked using Kappa inter-rater reliability with one experienced coder. A representative sub-sample of 50% of the dialogues were used. Therefore, 18 dialogues (6 from each condition) were

coded for reliability. The experienced coder coded the dialogues using the same procedure as the original coder. The number of agreements for each instance of a Check and Align were tallied as well as the number of times the coders disagreed. Kappa inter-rater reliability was then used to test the level of agreement. Agreement was at 84%, indicating a significant level of agreement ($p < 0.05$).

4.3.8 Scoring Collaborative Performance

The task was scored using a ranking system. The rank score was taken by ranking all of the items from 1 to 6 for price. For example, all of the prices for bread were listed from lowest to highest and ranked accordingly. All items on the participants' shopping list could then be attributed a rank and the mean rank could be found through averaging the 5 items. A low rank would indicate that the participants had favoured saving money and a high rank would indicate that the participants had favoured getting a short route. Therefore an average rank would indicate that the participants had collaborated more efficiently.

Negative scores on both the word count difference and difference in average words per turn between route and price objectives indicate participants with the route objective controlling the conversation. A positive score indicates that the participants with a price objective are controlling the conversation. A score closer to zero would indicate that greater levels of collaboration have taken place, as no particular participant has dominated the conversation.

4.3.9 Results

In addition to the questionnaire items, all dialogues were transcribed, and the following assessments were made for all three conditions:

1) Time to completion of task (measured in seconds).

2) Average route length (measured in centimetres).

3) Average amount of money spent (measured in pence).

4) Number of interruptions.

5) Number of times participants spoke to themselves.

6) Checks

7) Aligns

8) Total word count.

9) Word counts for participants with price and route objectives.

10) Total number of turns.

11) Total average number of words per turn (total turns divided by total word count).

12) Average number of words per turn for participants with price and route objectives.

13) Rank (for items bought).

14) Difference in number of words per turn between price and route objectives (positive scores were scored in favour of price and negative scores were in favour of route).

15) Difference in word count between price and route objectives (positive scores were scored in favour of price and negative scores were in favour of route).

Between groups analyses of variance (ANOVA) reveal that significant differences were found between the 3 conditions on a number of measures. These were time taken to complete the task, the number of aligns, the total word count, the total price word count, the total number of turns. No significant differences between the 3 conditions were found for the following measures: route length, money spent, number of checks, number of interruptions, number of occasions spent talking to self, the total route word count, total average number of words per turn, total average number of words per turn for route and

price objectives, rank, difference in words per turn between price and route objectives, and difference in word count between price and route objectives. A summary table of results can be found of the next page.

Table 4.2: Between-groups analyses of variance for video-meeting, face-to-face meeting and no-meeting conditions, with mean scores for all conditions for the above measures (standard deviations in parentheses).

Variable	Face Mean	Video Mean	None Mean	Significance
Time	539.8 (218.2)	686.4 (266.5)	1099.1 (410.8)	$F(2,45)=10.53$, $p<0.01$
Length	56.5 (14.8)	52.4 (16.2)	48.6 (13.1)	$F(2,45)=0.86$, $p>0.05$ NS
Money	1038.3 (76.9)	1055.4 (73.6)	1053.3 (75.5)	$F(2,45)=0.18$, $p>0.05$ NS
Interruptions	5.4 (4.2)	7.5 (6.9)	8.3 (8.6)	$F(2,45)=0.56$, $p>0.05$ NS
Talk to self	3.2 (3.5)	5.1 (4.2)	7.3 (5.9)	$F(2,45)=2.34$, $p>0.05$ NS
Checks	4.4 (3.6)	6.3 (2.7)	7.6 (3.6)	$F(2,45)=2.76$, $p>0.05$ NS
Aligns	2.8 (2.3)	4.6 (6.5)	9.8 (9.3)	$F(2,45)=3.64$, $p<0.05$
Total word count	749 (334.7)	1007.8 (430.2)	1344.6 (280.9)	$F(2,45)=8.54$, $p<0.01$
Price word count	352.5 (212.2)	477.3 (243.3)	717.9 (344.4)	$F(2,45)=5.58$, $p<0.01$
Route word count	396.5 (183.4)	530.6 (264.4)	626.7 (235)	$F(2,45)=3.03$, $p>0.05$ NS
Total turns	100.3 (43.9)	117 (46.6)	153.5 (47.2)	$F(2,45)=4.22$, $p<0.05$
Total average words per turn	7.5 (1.8)	8.6 (1.7)	9.2 (2.3)	$F(2,45)=2.21$, $p>0.05$ NS
Route average words per turn	8.3 (2.6)	8.9 (2.6)	8.3 (2.5)	$F(2,45)=0.25$, $p>0.05$ NS
Price average words per turn	6.9 (2.1)	8.2 (2.7)	10 (5.9)	$F(2,45)=1.92$, $p>0.05$ NS
Rank	2.1 (.9)	2.2 (.9)	2.4 (.9)	$F(2,45)=0.32$, $p>0.05$ NS
Difference in words per turn	-1.46 (3.3)	-0.75 (3.9)	1.69 (7.7)	$F(2,45)=1.14$, $p>0.05$ NS
Word count difference	-44 (212.9)	-53.3 (270.5)	91.3 (518.4)	$F(2,45)=0.61$, $p>0.05$ NS

Table 4.2 illustrates a similar pattern for all variables analysed. The face-to-face meeting condition scores best, followed by the video-mediated meeting condition and then the no-meeting condition scoring worst. To test whether these differences were significant between groups one-way analyses of variance were conducted.

Results indicate a significant difference between the conditions on the amount of time taken to complete the task ($F(2,45) = 10.534$, $p<0.05$), (mean scores: face-to-face meeting = 539.8; video meeting = 686.4; no meeting = 1099.1). So, richer mediums of initial meeting result in faster task completion rates. Using Tukey HSD post hoc tests, a significant difference between face-to-face and no-meeting conditions was found ($p<0.01$), a significant difference between video-mediated and no-meeting conditions ($p<0.01$), but no significant difference between video-mediated and face-to-face meetings. These results suggest that familiarisation in both video-mediated and face-to-face conditions result in more efficient communication (i.e., the task is completed more quickly).

Significant differences were also found between the three groups on total word count ($F(2,45) = 8.54$, $p<0.05$) (mean scores: face-to-face = 749; video-mediated = 1007.8; no-meeting = 1344.6). Once more this indicates that the richer the medium of initial meeting, the fewer words that are needed to complete the task. Post hoc Tukey HSD tests reveal a significant difference between face-to-face and no-meeting conditions ($p<0.01$), a significant difference between video-mediated and no-meeting conditions ($p<0.05$), and no significant difference between video-mediated and face-to-face meeting conditions. These results indicate that familiarisation in both video-mediated and face-to-face conditions results in more efficient use of language.

Results indicate a difference between the three conditions for word count for participants with a price objective (F(2,45) = 5.575, p<0.05) (mean scores: face-to-face = 325.5; video-mediated = 477.3; no-meeting = 717.9). This indicates that richer mediums of initial meeting result in a more efficient use of language for participants with the price objective. Post hoc Tukey HSD tests reveal a significant difference between face-to-face and no-meeting conditions (p<0.01), a significant difference between video-mediated and no-meeting conditions (p< 0.05), and no significant difference between face-to-face and video-mediated meeting conditions. These results once more indicate that as the medium of initial meeting gets richer, fewer words are needed in order to complete the task. Participants with the price objective completed the task more efficiently after familiarisation took place.

Results indicate a significant difference between the three conditions on the total number of turns in the dialogue (F(2,45) = 4 .22, p<0.05) (mean scores: face-to-face = 100.3; video-mediated = 117; no-meeting = 153.5). This result indicates that the richer the medium of initial meeting, the fewer turns that are needed in order to complete the task. Post hoc Tukey HSD tests reveal a significant difference between face-to-face and no-meeting conditions (p< 0.01), a significant difference between video-mediated and no-meeting conditions (p< 0.05), and no significant difference between video-mediated and face-to-face meeting conditions. These findings indicate a more efficient use of language in both familiarisation conditions.

Results also indicate that there is a significant difference between the three conditions with respect to the number of aligns used by the participants (F(2,45) = 3.644, p<0.05). Means scores for the three groups (face-to-face meeting = 2.8, video meeting = 4.6, no meeting = 9.8) once more indicates a relationship between richness of medium of initial meeting and

communication. Post hoc Tukey HSD tests indicate a significant difference between face-to-face and no-meeting conditions ($p<0.01$), no significant difference between video-mediated and no meeting conditions, and no significant difference between face-to-face and video-mediated meeting conditions. Therefore, there is some evidence here that face-to-face familiarisation allows participants to monitor understanding and attention more effectively in proceeding video-mediated meetings.

A further a priori t-test was carried out to test if there were any differences between conditions on the number of Checks used by participants. The justification behind this test is that Clark (1996) has indicated that familiar people have to work less hard to establish mutual understanding. The number of Checks is one way to measure mutual understanding. Essentially, participants with a higher level of mutual understanding would need to check less often that they have understood their partner's message. Significantly fewer Checks occurred in the face-to-face condition than the no-meeting condition ($t(22) = 2.16$, $p<0.05$). This would indicate that participants in the face-to-face meeting condition had higher levels of mutual understanding than participants in the no-meeting condition. No other significant differences were found.

A further a priori t-test was also carried out between the face-to-face meeting condition and the no-meeting condition on the number of times participants spoke to themselves. Doherty-Sneddon and McAuley (2000) propose a link between intimacy distance and social co-presence. Essentially, in conditions where individuals do not have the opportunity to express cues for intimacy, lower levels of perceived co-presence will be experienced. It was expected that face-to-face familiarisation would help to reduce psychological distance between participants, and one way to measure this is to look at the

number of times participants talk to themselves. Essentially, if a participant feels isolated from the communication, it is expected that they are more likely to exhibit anti-social behaviour (for example talking to themselves). Participants spoke to themselves on significantly more occasions in the no-meeting condition than the face-to-face condition ($t(22) = 2.07$, $p < 0.05$). This would indicate that levels of social co-presence were lower in the no-meeting conditions.

A further a priori t-test was also carried out between the face-to-face meeting condition and the no-meeting condition on the length of turns (in words). Research suggests that in comparison to face-to-face communication video-mediated interactions are more formal with participants using longer turns. This can be accredited to the increased difficulty in managing speaker exchanges. It was therefore of interest to test whether initial face-to-face familiarisation changes communication in the subsequent video-mediated context. The number of words per turn was on average significantly higher in the no-meeting condition than in the face-to-face meeting condition ($t(22) = 1.96$, $p < 0.05$). This indicates that face-to-face familiarisation results in communication closer to face-to-face in the proceeding video-mediated context.

Further to the analyses reported, analysis of the questionnaire data was also undertaken. Results are reported in table 4.3.

Table 4.3: Questionnaire data: Between-groups analyses of variance for video meeting, face-to-face meeting and no-meeting conditions, with mean scores for all three conditions (standard deviations in parentheses).

Measure	Face Mean	Video Mean	None Mean	Significance
Ques1	3.75 (.89)	3.88 (.68)	3.46 (.88)	$F(2,45)= 1.61, p> 0.05,$ NS
Ques2	3.33 (1.13)	3.54 (1.22)	3.25 (1.15)	$F(2,45)= 0.39, p> 0.05,$ NS
Ques3	4.21 (.72)	4.13 (.61)	3.67 (1.05)	$F(2,45)= 3.07, p> 0.05,$ NS
Ques4	4.04 (.55)	3.83 (.87)	3.70 (.81)	$F(2,45)= 1.19, p> 0.05,$ NS
Ques5	4.17 (.70)	4.17 (.64)	3.96 (.86)	$F(2,45)= 0.64, p> 0.05,$ NS

Key 1:

Question 1: Please rate how much you enjoyed using the videoconferencing equipment.

Question 2: If it were your choice would you use this particular method of communication again?

Question 3: How successfully would you rate the communication process between yourself and the other participant?

Question 4: Please rate how well you achieved your aim.

Question 5: Please rate how well you collaborated with the other participant.

Between groups analyses of variance revealed that there were no significant differences between the three groups. Further a priori t-tests were carried out comparing the face-to-face meeting condition with the other conditions on all 5 questions. Research reveals that participants are generally more satisfied communicating face-to-face than via any other medium. For this reason it was expected that participant evaluations would be higher after

face-to-face familiarisation. The only significant result was found on question 3 for face-to-face meeting against no-meeting (t(46) = 2.084, p<0.05). This result indicates that participants rate the communication process as being more successful when they have had a face-to-face meeting prior to the task taking place, compared to when no form of initial meeting takes place.

4.3.10 Discussion

Data from this study indicates that face-to-face familiarisation results in a number of benefits over no form of familiarisation. Video-mediated familiarisation also results in a number of benefits over no form of familiarisation. Meeting face-to-face prior to completing the task, however, produced no significant differences in comparison to the video-mediated meeting condition. There was however a noticeable trend in the data for all measures taken. The face-to-face familiarisation condition scored best, followed by the video-mediated familiarisation condition, and then the no-meeting condition. Although there were numerical differences between the face-to-face and video conditions, these differences were not significant and therefore there was insufficient power to discriminate between these two conditions.

One difference that was found between the face-to-face and no-meeting conditions was in the number of Checks and Aligns employed by the users. Anderson et al (1997b) noted that conversants in face-to-face dialogues check less often that their listener understands them (ALIGN) or that they have understood their partner (CHECK) when compared to interaction that takes place audio-only. A similar pattern has been shown by Doherty-Sneddon et al (1997), however participants used significantly more Aligns in audio-only interactions than in video-mediated interactions. These studies indicate that visual signals

can be used to monitor understanding and attention. Results from this experiment reveal that participants used significantly fewer Checks (42% less) and Aligns (72% less) after meeting face-to-face than when no initial meeting has taken place. It would appear then that this initial meeting has resulted in participants being better able to monitor understanding and attention in the proceeding video-mediated context. One possible explanation for this finding is that through face-to-face familiarisation there had been a decrease in psychological distance. In this respect, feelings of social co-presence would have been heightened. In turn higher levels of mutual understanding would have been formed between participants. Participants therefore would have felt less need to check that their partner understood them or that they had understood their partner. A comparison of video-mediated meeting and no-meeting conditions does not reveal this difference, meaning that face-to-face familiarisation produced one advantage over no-meeting that video-mediated familiarisation did not

Participants also spoke to themselves on significantly fewer occasions in the face-to-face meeting condition in comparison to the no-meeting condition. This may be a result of a reduction in psychological distance, as talking to oneself may be an indication that the user feels in some way removed or isolated from the communication. This indicates that higher levels of co-presence were formed between participants in the face-to-face meeting condition. Video-mediated familiarisation did not result in more examples of participants talking to themselves in comparison to the no-meeting condition. Once more, this indicates that face-to-face familiarisation produced one advantage over no-meeting that video-mediated familiarisation did not. Although it would seem that initial face-to-face familiarisation is beneficial, the fact that there were no differences with the video-mediated meeting condition would indicate that familiarisation rather than face-to-face contact per se

is important here, although numerically it appears that on all measures face-to-face is better than video.

There were also differences between the conditions with regards to how efficiently the end goal of communication was carried out. For example, face-to-face familiarisation resulted in significantly fewer words per turn in total. The video-mediated meeting and face-to-face meeting conditions also resulted in significantly shorter dialogues than the no-meeting condition. Both forms of familiarisation also resulted in the task being completed significantly quicker. This more efficient use of language is also reflected in participant evaluations of how successful the communication process was. Participants in the face-to-face meeting condition rated the communication process as significantly more successful than participants in the no-meeting condition. The longer dialogues in the no-meeting condition can be partly credited to the higher number of aligns used by participants. They can also be credited to participants taking a longer time getting to know one another. Many participants in the no-meeting condition introduced themselves and enquired as to how their partners were. For example:

P1 "How you doing? My name's Peter?"

P2 "I'm Laura, how are you?"

P1 "Not too bad"

Participants in other two conditions were given this opportunity in the initial meeting prior to the task, and therefore did not find it necessary to make introductions during the task. This finding adds some support Alexander et al's (1999) conjecture that initial warm-up sessions before video-mediated meetings with an emphasis on informal interaction (for example personal introductions or games to break the ice) are beneficial.

An increased occurrence of participant introductions and the number of aligns do not fully explain why the dialogues were so much longer in the no-meeting conditions. Another explanation for the increased length of the dialogues is that participants also spent longer amounts of time attempting to establish what the task was about, and what each of their priorities were. Essentially, participants in the no-meeting condition spent longer amounts of time attempting to establish mutual knowledge. One example from the transcribed dialogues includes:

P2 "Good. Right so what's your priority for your shopping list, what have you got to do?"

P1 "Right, what I'm, I'm on my lunch break just now and I have to try and get all these items, the bread, steak, wine light bulbs and dog food as quickly as possible"

P2 "As quickly as possible?"

P1 "As quickly as possible."

Another example includes:

P2 "Right so what is your objective, where are you to go?"

P1 "I have to erm buy these five items from a shopping list and draw like a, the shortest route possible around the shops"

P2 "What's the items?"

P1 "Erm, bread, steak, wine, light bulbs and dog food"

P2 "Right"

This finding is surprising considering that participants in all conditions were given the same information. This finding also supports Clark (1996) in that familiar people find it

easier to establish mutual understanding. In total participants in the face-to-face meeting condition spent on average 5.7% of the dialogue engaging in introductions and attempting to establish mutual knowledge. In the no-meeting conditions an average of 8.9% of the dialogues was spent engaging in such behaviours. Although this difference is not statistically significant, combined with the higher number of Aligns, this finding goes some way to explaining the difference between conditions in the length of dialogues.

All participants in this study had no prior experience with video-mediated technologies. It is likely then that many of these participants would have had a lack of confidence with using the equipment. Indeed researchers such as Armstrong-Stassen study (1998) point out that new inexperienced users of video-mediated technologies often experience levels of anxiety. Furthermore participants in the video meeting and the face-to-face meeting conditions differed in the amount of experience that they had with the equipment. The video-mediated meeting participants conducted the initial meeting over video whereas the face-to-face meeting participants conducted the initial meeting face-to-face. It is therefore likely that the video-mediated meeting participants were more confident than the face-to-face meeting participants accredited to higher levels of experience. In order to remove this confounding variable it is necessary to investigate the effects of face-to-face familiarisation in more applied longitudinal settings. If all participants have similar levels of experience with video-mediated technologies, will the effects of face-to-face familiarisation be more profound?

The questionnaire used in this study aimed to target user perceptions of collaborative performance. This questionnaire was included as it was expected that familiarisation would have an effect on collaboration, and one way to measure this is through subjective

perceptions. In retrospect it may have been beneficial to also include a number of the questions used in Study 2 (Face to face familiarisation and person perceptions across a video link). In the previous study face-to-face familiarisation resulted in more favourable perceptions of a salesman and an increase in terms of an illusion of social co-presence. It would have been useful to see if these same effects occurred in this investigation, especially considering that this study involved a true dialogue as opposed to the monologue used in Study 2. Considering that video-mediated technologies are used for a number of different purposes, it would have been beneficial to test user perceptions of co-presence and of others using the media in a number of different communication situations. The social dynamics of a monologue are vastly different from a true dialogue. During a monologue communication is one-way and therefore the listener has little opportunity to affect the communication process. Indeed, in the previous study the sales recitations were pre-recorded and therefore the participants had no opportunity to affect the communication process. The true dialogue on the other hand involves two-way communication and therefore both conversational participants can directly affect the process of communication.

Overall it would seem that video-mediated meetings are statistically like face-to-face meetings. Face-to-face familiarisation has advantages over video in respect to its comparison with no meeting on some measures, however, video was also statistically better than no-meeting for many measures. Therefore, having a face-to-face meeting does bring additional benefits, but a video meeting can fulfil some and is better than nothing at all. The exact number, timing or length of such face-to-face meetings for successful communication is unknown, and will vary according to the task that needs to be completed and the type of people who are using the media. For example, it would be expected that

established friends or individuals who have regular face-to-face contact would not benefit from such an intervention. Indeed, such a method would be most suitable for individuals who have never met before or who do not meet regularly. Many questions have been raised from this research. One area of interest in particular concerns the duration and frequency of familiarisation meetings. Would there be any extra benefit to having face-to-face familiarisation meetings for longer than 10 minutes? Future research should aim to compare task performance between varying levels of meeting duration. It also of interest to discover how many face-to-face meetings would be required for individuals communicating at a distance at regular intervals. A future study could compare varying frequencies of face-to-face meetings in order to test the most appropriate level.

4.4: Chapter summary

Evidence from both studies would seem to suggest that face-to-face meetings prior to a videoconference can be beneficial in a number of respects, and would therefore validate Abbott et al's (1993) claims that such meetings need to placed alongside remote conferencing for more successful communication. However, the results also indicate that many benefits can be realised from video-mediated familiarisation and therefore if face-to-face meetings are impossible, such a method will be better than nothing at all. It should be noted here that the tasks were performed immediately after the familiarisation process had taken place. Would the benefits of face-to-face familiarisation be more profound if tasks over videoconference are completed a number of days or even weeks after participants had originally met?

One explanation for the difference in findings between the two studies lies in the types of tasks that were employed. Results from study 2 revealed many differences between

conditions. Perhaps the number of differences here are a result of the fact that the task was more social in nature and involved judging another person. Indeed a review of the literature indicates that many differences have been found between media when a more social element to tasks is used (for example Short, Williams and Christie, 1976; Williams, 1977; Whittaker, 1995; Gale, 1989). Study 3 on the other hand was more of a problem-solving task (although involving an element of negotiation). Studies on the differences between media on problem solving tasks have been shown to reveal few differences (for example, Ochsman & Chapanis, 1974, Chapanis, 1975; Davies, 1971b; Reid, 1970). This would therefore indicate that the precise importance of initial meeting is determined by the task in hand. It would seem then that tasks that involve a more social dimension would necessitate the use of face-to-face familiarisation, however problem-solving tasks would not need such an intervention.

It was also proposed in chapter 3 that attenuation of visual cues (especially eye gaze) was another explanation for the feelings of distance experienced in the use of video-mediated technologies. Consequently, it was also suggested that looking up at the camera might go some way to resolve this problem. Therefore, the following chapter will investigate the effects of video-mediated gazing.

Chapter 5: Investigations into the Effects of Video-Mediated Gazing.

5.1: Introduction

5.1.1 Aims and objectives

Evidence from chapter 4 indicates that face-to-face familiarisation goes some way to reduce the negative impacts of video-mediated communication. One finding in particular was that participants were perceived more favourably (in a video-mediated context) after being met face-to-face. There was also some evidence to suggest that face-to-face familiarisation reduces psychological distance. Familiarisation (in both video-mediated and face-to-face conditions) also revealed benefits in terms of task outcome and communicative efficiency. These findings indicate that the precise importance of initial meeting is determined by the task in hand. One reason why there were clearer face-to-face benefits in study 2 was because it was more of a social task, which involved judging another person. Study 3, however was a problem solving task. These findings therefore reflect the task difference literature.

It has been proposed that video-mediated communication does not deliver the same level of social co-presence as face-to-face interaction. This is partly credited to physical remoteness (which was addressed through face-to-face familiarisation), but also credited to the attenuation of visual cues (in particular eye gaze). Gazing behaviour can not be used effectively in video-mediated communication. Therefore, another suggestion to improve video-mediated communication that comes out of the literature and study 1b (questionnaire) is to employ the practice of video-mediated gazing. Video-mediated gazing involves users looking directly into the camera at important points in speech. This particular practice will give the other participant the impression that eye contact is being made. The purpose of this chapter is to discover if such a practice has an effect on task performance and user satisfaction. The successful implementation of video-mediated

gazing will be measured using three tasks with three separate pools of participants. Measures of success will include participant evaluations and perceptions of others and the media, and measures of task performance (accuracy of recall). It is expected that the use of video-mediated gazing will have an impact on user satisfaction and task performance.

5.1.2 Previous research on gaze

The implementation of videoconferencing systems is largely based upon the assumption that visual signals improve human interaction. Research suggests that access to visual signals improve the quality of communication, although does not necessarily have an effect on task performance. One particular type of non-verbal behaviour that has an impact on communication is eye gaze. Gazing behaviour has been shown to be important in human communication for a number of reasons: gaze can serve as an indicator of understanding, knowledge and attention (Beattie, 1981; Clark & Marshall, 1981; Rutter, 1984); gaze plays a role in regulating conversation (Kendon, 1962; Goodwin, 1981), as well as a number of other communicative functions (often in conjunction with other parts of the face, for example in the expression of emotion). Further to this, gazing behaviour has been noted to have an effect on personal perceptions, for example, gaze can affect evaluations of liking and attraction (Exline and Eldridge, 1967; Thayer and Schiff, 1974), honesty (Hemsley and Doob, 1978), and intelligence (Wheeler et al, 1979).

Gaze has also been noted to have a number of cognitive effects. Research has shown that students remembered more instructions from a teacher who looked at them more frequently (Fry & Smith, 1975). Similarly, primary school students remembered more of a story when their teacher gazed more frequently (Otteson & Otteson, 1980). One interpretation for these findings is that gazing results in increased physiological arousal. The effects of

gazing on information recall can be explained using the Yerkes-Dodson law (1908) which predicts an inverted 'u' shaped relationship between arousal and cognitive performance. According to Beattie (1981) high levels of inappropriate gazing result in high levels of arousal, which in turn disrupt cognitive processing. A relationship has also been found between low levels of arousal and poor task performance, as low arousal levels do not result in significant motivation to complete a task. Moderate levels of gazing, however, will result in moderate levels of arousal, which may help to provide the motivation to learn (Yerkes and Dodson, 1908). It therefore follows that gazing behaviour has an effect on recall as it increases arousal and therefore improves cognitive processing. Also, it may be the case that gazing serves as an indicator that something important is going to happen, therefore serving to reinforce salient points in conversation, much in the same way that eyebrow movements are used to reinforce important aspects of speech (Ekman, 1979; Whittaker and O'Conaill, 1997).

Although it is clear that V.M.C systems allow users access to non-verbal signals, one problem is that video-mediated communication does not replicate the complexities of human gazing behaviour. During a videoconference information from the eyes is limited. Due to the manner in which normal video systems are set up, mutual eye contact is impossible. The camera is usually placed above the monitor and not inside of it, and therefore if one looks at the eyes of another person, it will appear to the other user that he/she is looking in a downward direction. Joint attention is also impossible as users do not share a common environment, and therefore there are no shared objects that can be mutually oriented to. It is also evident, however, that some level of gazing behaviour can be replicated in videoconferencing technologies through looking directly into the camera: this will give the viewer the impression that the other participant is looking them in the

eyes. Most users of this technology, however, focus attention at the monitor (displaying the image of the other conversational participant) and not at the camera. The result of this practice is that users appear to be looking off camera, which in turn means that they do not have eye contact that gives conviction to what they are saying (Tiffin and Rajasingham, 1995). Even when videoconferencing systems are set up in manner in which natural eye contact can take place, for example videotunnels (Smith et al, 1991), the results are far from pleasing. Research by Doherty-Sneddon et al (1997), for example, indicates that in such a set-up users tend to over-gaze, which in turn affects cognitive processing and consequently task performance. It is therefore suggested that users adopt video-mediated gazing behaviours (look directly into the camera). It is important that users get the most out of the video link by using the projected images effectively (i.e., look at the monitor), but at the same time adapt to the medium and project themselves appropriately. In this case deliberately sending gaze signals to the camera. The implications of using such a technique are investigated in the following experiments.

5.2: Study 4: The effects of video-mediated gazing on recall and user perceptions of a salesman and product

5.2.1 Introduction and aims

This study was conducted in order to investigate whether video-mediated gazing behaviour improves video-mediated communication. Gazing behaviour serves many functions in human interaction, however this study focuses on how video-mediated gazing affects the remembering of information and participant evaluations of social co-presence, confidence, friendliness, honesty and communicative ability. Abele (1981) for example noted that people tend to favour, when engaged in social interaction, those who employ moderate levels of gaze as opposed to low levels of gaze. Exline and Eldridge (1967) also noted that

confederates who gazed frequently received higher evaluations from participants than confederates who gazed infrequently. These studies would seem to indicate that moderate levels of gazing behaviour in a videoconferencing environment, as opposed to no gazing behaviour would be perceived as friendlier and more favourable. Also it was expected that video-mediated gazing would be viewed as more friendly and akin to face-to-face communication than in situations where gazing behaviour was not present. Research conducted by Fry and Smith (1975) and Otteson and Otteson (1980) would also suggest that video-mediated gazing would result in increased recall of information.

In order to test the benefits of video-mediated gazing behaviour, recorded video recitations were used which displayed a confederate attempting to sell a soap product. In one condition the participants viewed the confederate gazing at the camera at predefined points in the speech. In the other condition the confederate adopted the usual response when communicating by videoconference, focusing his attention on the monitor and therefore no gazing behaviour was present. Measures of task performance included a recall test, and a questionnaire, which aimed to test whether gazing behaviour in a videoconferencing environment increases evaluations of friendliness, confidence, honesty and communicative ability.

5.2.2 Participants

An opportunity sample consisting of 32 students (both undergraduate and postgraduate) were taken from the University of Stirling. Sixteen participants were male, and 16 participants were female. All participants had no prior experience with video-mediated technologies, or limited experience with video-mediated technologies (for example had used such equipment on a few occasions in the past only). No details of age were taken.

The salesman (confederate) was male and 35 years of age. Only one salesman was used in order that differences in selling techniques could be eliminated.

An opportunity sample of a further 28 participants from the University of Stirling were also taken. Fourteen of these participants were male and 14 were female. All participants had no prior experience with video-mediated technologies, or limited experience with video-mediated technologies (for example had used such equipment on a few occasions in the past only). No details of age were taken. These participants were played the same sales pitches, however the recitations were played audio only.

5.2.3 Design

A within-subjects related-samples design was used. Participants were assigned to watch one of four videos; therefore 8 participants were assigned to each video. In order to counterbalance order effects and product effects, two products and four videos were used. The videos were as follows:

Video 1 – Product 1 (gaze) and Product 2 (no gaze)

Video 2 – Product 2 (no gaze) and Product 1 (gaze)

Video 3 – Product 1 (no gaze) and Product 2 (gaze)

Video 4 – Product 2 (gaze) and Product 1 (no gaze)

5.2.4 Materials

In room 1, a colour monitor (JVC TM-14EK(B)) was mounted in a wooden box, with a video camcorder (Sony CCD-TR2200EPAL) placed directly above the monitor. A microphone was placed to the right of the monitor, and video and audio quality were as

high as achievable in the lab. The monitor and camcorder in room 1 were connected to room 2, adjacent to room 1, in which a video camcorder (Sony CCD-TR2200EPAL) was used to play the pre-recorded recitations to the participants. Both monitors were 14 inches in size.

Each participant was distanced approximately one metre from the monitor and the scope of the view included the participant's face and upper body.

For product 1, the gazing condition was accompanied by 30% gazing (of total speech time), and 32 % gazing (of total speech time) for product 2. As well as being instructed to gaze at the camera during certain points in the monologue, the actor was also instructed to look at the monitor when not looking at the camera. The actor was given time to rehearse the sales pitches and time to practice the recitations until perfected. During the recording of the sales recitations, the experimenter ensured the actor looked up at the camera at the appropriate points and also ensured the actor gazed at the monitor when not looking into the camera. If at any time the actor failed to follow the instructions appropriately, the experimenter stopped recording and the actor was instructed to begin anew. Each recorded recitation was later viewed over video in order to that the experimenter was happy that the instructions had been followed appropriately.

The recitations were memorised by the salesman and took the following format (Underlined parts indicate when gazing behaviour occurred)

<u>Product 1</u>

"Fresh Face" is a brand new soap that has been developed by scientists in **Italy**. This product is a triple purpose soap and provides an all new cleansing experience. Firstly, it leaves your skin smelling like your **favourite after-shave or perfume**. Secondly it doubles as an **air-freshener**, leaving your bathroom smelling perfumed and odour free thanks to our unique **odour eating ingredients**. Thirdly this product helps to fight the build up of spots through the active ingredient **blemish-buster**. This product is on sale for **£1.89** or alternatively you can purchase a special three pack for only **£4.50**. This soap is available at all good retailers but thanks to a special promotion can be purchased at **Johnson's Chemists** where if you buy one soap you receive a free box of **toothpaste**. Our soap makes an **ideal gift for your partner** and is sold in **Five** different varieties – For women there's **Moonlight** for the romantic, the sweet smelling **Daphne** and the refreshing **Atlantis**. For men there's **Rhino and Brutus**. We hope that our product will reach you **in perfect condition but we have a money back guarantee if you are not entirely satisfied**. For further information on this product phone **Glasgow** on our free 24-hour line. The number is **(0141) 446619** and you can talk to our customer services manager **Lynne Thomas**. So please remember – **"Fresh Face" is the feel good soap.**

Product 2

"Smooth-Skin" is a completely new soap produced in **Portugal** that leaves you feeling invigorated and refreshed. This product is a triple action soap, which contains **apricot and honey**. This soap **exfoliates your skin** ridding you of the build up of dead skin cells. Also it contains **active moisturising ingredients** which leave your face feeling smooth and silky. Our product, however, also contains a **special gel** which prevents the bar of soap from leaving a messy, sticky residue on your bathroom sink. Our soap costs **£1.25** for a single bar or alternatively you could pick up a triple pack for **£3.90**. This product can be

bought at all good shops and due to a special promotion with **Wilson's toiletries** can be purchased with a free bottle of **Shampoo**. This product makes an ideal present for **Mothers** and is sold in **4 different colours** – sensuous **ocean blue**, exotic **coconut brown**, exciting **fox red** and relaxing **shamrock green**. **We guarantee 100% satisfaction with our product** but if you wish to return this item you can **swap it for another product of equal or less value**. For more information on our exciting range of gifts please ring toll free on **Birmingham (0121) 559937** and speak directly with our product designer **Billy Wood**. Please remember **"Smooth Skin" is the fuss free soap**.

A questionnaire containing 10 questions was circulated to all participants. Each question was rated on a 5-point scale; with one being the lowest possible score and five being the highest possible score. So answers ranged from 'not at all' to 'very.'

The questions were as follows:

1) How well did the salesman communicate?

2) How friendly was the salesman?

3) How honest was the salesman?

4) How confident was the salesman?

5) How likely would you be to purchase this product?

6) How unique was the product?

7) How close was the video interaction to face-to-face communication?

8) Did you feel that the salesman was talking to you?

9) How much did you enjoy using the equipment?

10) Would you use the equipment again?

5.2.5 Procedure

Participants were seated in a video-conferencing room and were asked to face the monitor. Participants were informed that the study was investigating the marketability of two separate products and also whether videoconferencing equipment could be used to successfully sell products. This measure was taken to distract the participant's attention away from the real aim of the study. Participants were informed that they were about to see someone who would describe two different products to them. Participants were led to believe that the salesman was communicating real time over a video link. In order to prevent participants from talking to the salesman and thus discovering that it was actually a pre-recorded video, participants were requested to refrain from conversing with the salesman. Participants were informed that the experimenter wished to ensure standardisation with all participants, needing all participants to receive the same information in roughly the same period of time. Once these instructions had been relayed the experimenter left the room and played the tape (from room 2). Once the participant had viewed the salesman describe the first product, a questionnaire relating to their evaluations of the salesman, product and equipment was completed. The participant then saw the second product described and filled out a second questionnaire. Finally, all participants were given an unexpected recall test for both products (in the order that they had seen them). The number of correct answers and intrusions were noted. An intrusion was classed as an instance when an answer was recalled incorrectly and contained information that belonged to the other product. After completing the experiment, participants were debriefed as to the deception and the true nature of the research.

Further to this, a separate sample of participants were played the sales pitches audio-only. These participants followed the same procedure and were given the same instructions as

the previous participants, however due to time restrictions these participants were only given the questionnaires to complete and no form of recall test was administered.

5.2.6 Scoring

Participants were given a recall test for information contained in the recitations for both products. For both recall tests a highest possible score of 21 points was achievable. The questions were as follows, with the number of points for each question in brackets.

Product 1

1) What was the name of the product? (1)

2) In what country has the product been developed? (1)

3) What three purposes does this product serve? (3)

4) What is the name of the spot-fighting ingredient? (1)

5) How much does it cost to purchase one bar? (1)

6) How much does it cost to purchase a three pack? (1)

7) With which retailer is there a special promotion? (1)

8) What do you get free when you buy one bar of soap? (1)

9) For whom does this product make an ideal gift? (1)

10) How many different varieties are there? (1)

11) Name the women's varieties (3)

12) Name the men's varieties (2)

13) What guarantee comes with this product? (1)

14) What is the free-phone number that you must call for more information? (1)

15) What is the name of the customer services manager? (1)

16) Complete the slogan for this product: "Fresh Face is the......." (1)

Product 2

1) What was the name of the product? (1)

2) In what country is the product produced? (1)

3) What two ingredients are contained in this product? (2)

4) What three purposes does this product serve? (3)

5) How much does it cost for a single bar? (1)

6) How much does it cost for a three pack? (1)

7) With which company is there a special promotion? (1)

8) What do you get free when you buy a bar of soap? (1)

9) How many different colours does the product come in? (1)

10) Name the colours (full name with exact shade required) (4)

11) What guarantee comes with this product? (1)

12) What is the free-phone number that you must call for more information? (1)

13) Which city is this service based in? (1)

14) What is the name of the product designer? (1)

15) Complete the slogan for this product: "Smooth skin is the…." (1)

5.2.7 Results

Results indicate that participants recalled significantly more information in the camera gaze condition than in the control condition. Participants in the camera gaze condition also rated the salesman as being significantly more friendly, more honest, more confident and a better communicator. Perceptions of co-presence were also affected, with higher ratings in the camera gaze condition for the question "how close was the video interaction to face-to-face communication?" Higher ratings were also found in the camera gaze condition for the

question "did you feel that salesman was talking to you?" Participants also indicated significantly higher levels of enjoyment in camera gaze condition.

The number of intrusions was unaffected by gaze access, with no difference between the conditions. Participant perceptions of the product were also unaffected, with no difference in ratings for likeliness to purchase the product and product uniqueness. Also, participants did not indicate a preference between the two conditions for likeliness of using video-mediated technologies in the future.

Table 5.1: Independent t-test scores for correct answers, intrusions and for questionnaire questions 1-10 for camera gaze and control conditions, with mean scores (standard deviations in parentheses).

Measure	Gaze Mean	No gaze mean	Significance
Correct Answers (recall test)	7.02 (3.25)	5.70 (2.45)	$t(31) = 2.30$, $p<0.05$
Number of intrusions	1.06 (1.11)	1.22 (1.18)	$t(31) = -0.76$, $p > 0.05$, **NS**
Question 1: How well did the salesman communicate?	3.03 (0.86)	2.19 (0.90)	$t(31) = 5.00$, $p< 0.01$
Question 2: How friendly was the salesman?	2.72 (0.99)	2.19 (1.06)	$t(31) = 2.96$, $p< 0.01$
Question 3: How honest was the salesman?	2.78 (1.04)	2.38 (0.98)	$t(31) = 1.71$, $p< 0.05$
Question 4: How confident was the salesman?	2.75 (0.98)	2.16 (0.88)	$t(31) = 2.89$, $p< 0.05$
Question 5: Would you purchase this product?	2.50 (1.08)	2.10 (1.09)	$t(31) = 1.68$, $p> 0.05$, **NS**
Question 6: How unique was the product?	2.63 (1.29)	2.44 (1.13)	$t(31) = 0.73$, $p> 0.05$, **NS**
Question 7: How close was the video to face-to-face interaction?	2.94 (0.88)	1.91 (0.82)	$t(31) = 6.25$, $p< 0.01$
Question 8: Was the salesman talking to you?	2.72 (0.85)	1.56 (0.80)	$t(31) = 5.24$, $p< 0.01$
Question 9: How much did you enjoy using the equipment?	2.56 (0.72)	2.31 (0.74)	$t(31) = 2.10$, $p< 0.05$
Question 10: Would you use the equipment again?	2.34 (1.10)	2.16 (0.99)	$t(31) = 1.53$, $p> 0.05$, **NS**

Using paired-samples t-tests with a two-tailed hypothesis, differences between the two conditions were analysed. From table 5.1 it can be noted that participants in the camera gaze condition scored significantly better on the recall test than participants in the control condition ($t(31) = 2.30$, $p < 0.05$). Participants can therefore remember more information from speech that is accompanied by video-mediated gazing. For all questions that evaluated the salesman significant differences in favour of the camera gaze condition were found. Participants rated the salesman as a better communicator in the camera gaze condition ($t(31) = 5.00$, $p < 0.01$). Participants rated the salesman as being significantly friendlier in the camera gaze condition ($t(31) = 2.96$, $p < 0.01$). Participant ratings concerning perceptions of honesty were significantly higher for the camera gaze condition ($t(31) = 1.71$, $p < 0.05$), indicating the salesman was considered more honest when access to gazing behaviour was available. The salesman was also considered significantly more confident in the camera gaze condition ($t(31) = 2.89$; $p < 0.05$). Perceptions of social co-presence were also affected by access to gazing behaviour. Participants rated the video feed as being significantly closer to face-to-face interaction in the camera gaze condition ($t(31) = 6.25$, $p < 0.01$). Participants also gave a significantly higher rating for the camera gaze gazing condition on the question "did you fell that the salesman was talking to you?" ($t(31) = 5.24$, $p < 0.01$). This indicates that access to video-mediated gazing reduces psychological distance. Participants also indicated that they enjoyed using the equipment more when access to eye gaze was available ($t(31) = 2.10$, $p < 0.05$).

No differences were found between the conditions on the number of intrusions, indicating that participants were equally likely to substitute items from one product to the other in both conditions. Participant ratings concerning the product also did not differ between

conditions. Participants rated likeliness to purchase product and product uniqueness equally between conditions. Also, participant evaluations concerning likelihood of using the equipment in future was unaffected by gaze access.

An analysis was also undertaken using the same sales pitches, however they were played audio-only to a further 28 participants. This analysis was conducted in order to rule out the possibility that gaze instruction was affecting verbal behaviour. No significant differences were found between the camera gaze condition and the control condition on any of the 10 questions in the questionnaire.

5.2.8 Discussion

Results from this experiment lend support to Fry and Smith (1975) and Otteson & Otteson (1980) in that a speaker's gazing behaviour can influence a listener's mental processing of information and subsequent memory for it. Furthermore, these results inform us that this function of gaze can also be completed successfully over videoconference. Although gazing behaviour cannot be replicated to the fullest extent in a normal videoconferencing set-up, using video-mediated gazing can duplicate one particular function of eye gaze in face-to-face interaction: to facilitate the remembering of information. This finding can be explained in terms of the Yerkes-Dodson law (1908). This law indicates that moderate levels of arousal are needed in order to motivate oneself to complete a task. High and low levels of arousal are both counterproductive. Considering that gazing behaviour causes one to become aroused and that over-gazing, for example using videotunnels (Doherty-Sneddon et al, 1997) produces negative effects in task performance, it could be argued that the moderate levels of gaze used in this study have led to sufficient arousal to complete the task. Indeed, levels of gazing in this study were set at approximately 30%, which differs

considerably from the high levels of gazing found by Doherty-Sneddon et al (1997). Doherty-Sneddon et al (1997) found more than double the amount of eye contact in the use of videotunnels when compared with face-to-face interactions. These results may also be explained in terms of eye gaze acting as a cue to important information. Through looking up at the camera in speech one would assume that this would bring attention to what the speaker is saying, perhaps in the same way that we raise our eyebrows, nod our head, or change the tone of our voice to signal the importance of any given element of a speech (Ekman, 1979; Whittaker and O'Conaill, 1997). Gaze in this instance is informing the listener that something important has been uttered.

Perceptions of the salesman were also altered as a direct result of gazing. Wheeler et al (1979) noted that levels of gaze influence perceptions of intelligence. In this study, the same salesman was considered a better communicator when he used gazing behaviour. The findings also support Helmsley and Doob (1978) in that the salesman was considered more honest when he used gazing behaviour. It is a common perception that gaze aversion signals dishonesty, and therefore it is not surprising that the salesman was considered less honest when no form of gazing behaviour was used. Participants also rated the salesman as more confident and friendlier when gazing was used, findings that are once more borne out in the literature. The fact that there was no difference between the conditions with regards to perceptions of the product also bears weight to the argument that the gazing is having a social effect. The product itself remained constant between the conditions, however the salesman changed (in his use of his gazing behaviour).

In order to rule out the possibility that gaze instruction was affecting verbal behaviour, the same sales recitations were played audio-only to a separate pool of participants. This

analysis reveals no significant differences between the conditions for all 10 questions on the questionnaire, which would indicate that video-mediated gazing is responsible for the increased user perceptions of the salesman and the increased perceptions of social co-presence.

It could also be argued that a comparison can be made between video-mediated communication and television. On news broadcasts for example, newsreaders spend a vast amount of time looking directly into the camera. Therefore, it may be the case that we are used to information being presented in this way, and this may carry over into videoconferencing, as television would be the closest medium to judge video-mediated communication by. The experts in the interview in chapter 3 (study 1b) have even remarked that television is the benchmark for how videoconferencing is judged. It may be the case then that no gazing in a videoconference will be viewed as unusual and will be regarded with suspicion, as it does not fit expectations, especially for people who have never used video-mediated technologies before.

The use of video-mediated gazing results in a reduction of psychological distance. This is evident through higher ratings for two questions. Firstly the participants rated the experience as being closer to face-to-face interaction and secondly it was noted that the participants felt as though the salesman was more likely to be talking to them. As has been discussed in chapter 3, it is evident that one of the major factors preventing V.M.C being as successful as face-to-face interaction is the lack of personal contact. It has also been shown that social co-presence is an important outcome caused by this lack of personal contact. Users feel that there is a noticeable sense of distance. According to Doherty-Sneddon and McAuley (2000) social co-presence is affected by available cues for

intimacy. By replicating some sort of gazing behaviour participants are given another available cue to intimacy, and therefore psychological distance has been reduced. Many other methods could be employed to help reduce feelings of distance. In study 2 (chapter 4) it was shown that higher levels of familiarity improved feelings of social co-presence. In chapter 3 one user suggested that constructing a video set-up in a manner similar to how a meeting would be conducted face-to-face (around a table with monitors for each person) also helps. This current research, however, suggests that through employing the technique of video-mediated gazing, the remembering of information is facilitated, users are regarded more favourably, and levels of social co-presence are increased. So, some drawbacks associated with V.M.C can be overcome using deliberate gazing.

5.3: Study 5: The implementation of video-mediated gazing with normal participants

5.3.1 Introduction and aims

The previous study shows that one method for improving video-mediated communication is the practice of video-mediated gazing behaviour. This research, however, was conducted in such a way that the confederate was given time to memorise the sales pitch and practice it until he got it right. Although the results indicate the video-mediated gazing is beneficial in so far as it helps to improve memory and alters users perceptions, it is necessary to investigate whether video-mediated gazing is a practice that can be adopted readily by users of the technology. In other words, will video-mediated gazing be as useful when users are merely advised to adopt this practice, and are not given the opportunity to practice presentations until perfected? Furthermore it will be necessary to discover whether purely advising users to adopt such behaviour is suitable, or rather if more in depth training is required. In order to test this, participants were either asked to complete a

series of tasks with no advice at all, or were advised to adopt video-mediated gazing behaviours. As it is clear that video-mediated gazing has benefits, if the advice condition results in improved performance, then one could conclude that advice is sufficient. If there is no difference between the groups then training may be required. Based upon the previous study, it was expected that advising users to adopt video-mediated gazing would result in improved recall, improved perceptions of the salesperson, and improved perceptions of social co-presence.

This study consisted of two separate tasks:

1) Recall of story on 'Vernon the very old wolf'

2) Questionnaire on perceptions of salesman and product

In the second task additional participants were used as raters and therefore the methods section will be split into two separate parts.

5.3.2 Methods section for recall task on 'Vernon the very old wolf'

5.3.2a Participants

A sample consisting of twenty pairs (40 participants in total) of Open University students from the University of Stirling was taken. Seven of these participants were male and 33 were female. All pairs from this sample were completing a Psychology course and therefore had some degree of familiarity with one another. Participants volunteered to take part in the study and were sampled via a sign-up sheet. Participants were therefore assigned to pairs on the basis of which slot they had signed up for. No details of age were taken. Written informed consent was obtained so recordings of dialogues could be made for further analysis.

5.3.2b Design

A between groups independent measures design was used. Pairs were randomly assigned to one of two conditions: advice or no advice. Participant pairs completed two tasks: task 1 (a recall task), and task 2 (involving attempting to sell a soap product, and using product 1 – 'Fresh Face' from previous study). The sales recitations from task 2 were recorded for later analysis.

5.3.2c Materials

Participant pairs were separated by a screen. On one side of the screen, a colour monitor (JVC TM-14EK(B)) was mounted in a wooden box, with a video camcorder (Sony CCD-TR2200EPAL) placed directly above the monitor. A microphone was placed to the right of the monitor, and video and audio quality were as high as achievable in the lab. The monitor and camcorder on one side of the screen were connected to the other side of the screen and with the exact same set-up. The monitors were 14 inches in size.

Each participant was distanced approximately one metre from the monitor and the scope of the view included the participant's face and upper body.

One participant in the pair was given a story on 'Vernon the very old wolf' with an accompanying sheet containing all the important points of the story in order of occurrence (see appendix 5 for story and accompanying list). The other participant in the pair was given a list containing information on a soap product ('Fresh Face' from previous experiment) and an accompanying list of key points (same key points that were used on the recall test in the previous study).

5.3.2d Procedure

A story was devised about an old wolf called Vernon. One participant from each pair in both conditions (advice and no advice) was given 5 minutes to familiarise themselves with the story. However, in condition 1, the subjects were advised to use video-mediated gazing behaviours and in condition 2 the subjects were given no advice at all (see appendix 6 for advice given to participants).

After familiarising themselves with the story the participants were given a sheet of paper containing all the important points of the story in order of occurrence and were asked to relay the story to the other participant. The sheet containing the important points in the story was placed on a clipboard in front of the participant so that they could view it any time whilst conversing over the videoconference. The important points in the story were also the points in which the other participant would be tested for recall. Participants were asked to look up at the camera when relaying these important points. The participant telling the story was informed to build the story around these points of reference, as opposed to just listing them as they appeared.

After the story had been told the other participant in the pair was given a task of selling a product to the other participant, so that he/she could be distracted momentarily from the story. These participants were either advised to use video-mediated gazing behaviours or were given no advice at all. Participants were given information concerning a soap product and were given 5 minutes to familiarise themselves with this information. Again, key points were placed on a clipboard in front of the participant. The participants were also requested to look up at the camera at these key points. After this time participants were asked to sell the product to the other subject. Sales recitations were recorded for later

analysis. After this task had taken place, the participant who listened to the story was given a recall test for the information contained in the story. This distracter task was used as in the previous study participants completed the recall test after filling out the questionnaire (and therefore after a delay). It was therefore important to test long term memory for information in both studies.

5.3.2e Scoring

Participants were asked to recall parts of the story on 'Vernon the very old wolf.' The following questions were asked (the answers for each question are provided in brackets). Unless otherwise indicated, one point was available for each question. A total of 25 points were achievable.

1: What was the name of the wolf? (Vernon)

2: Where did he live? (Forest)

3: What was his problem? (Too old to catch food)

4: What was his favourite meal? (Deer)

5: What colour was the castle? (Black)

6: What colour were the flags? (Blue)

7: What symbol was on the flags? (Golden Lion with Red Cross – one point Golden Lion and one point for Red Cross. Half a point if golden, lion, red or cross indicated on own.)

8: What meal did the wolf steal first? (Chicken)

9: What kind of table did he steal it from? (Brown oak – one point for brown and one point for oak)

10: How did the queen react? (Angry and went to bed – one point for angry and one point for bed)

11: What was the queen's favourite instrument? (Harp)

12: What meal did the wolf steal second? (Pot roast)

13: What was the queen's solution this time? (Lay a trap)

14: What was the name of the huntsman? (John)

15: How did the wolf know there was a trap? (Saw big black boot sticking out from behind the table – one point for boot, and one point for sticking out from behind table)

16: What did the wolf hide behind? (Purple velvet curtain – one point for purple, one point for velvet and one point for curtain)

17: What food did he steal third time? (Fish)

18: How did the queen react this time? (Not angry)

19: What did she decree? (Food left outside castle for Wolf)

5.3.3 Methods section for task on perceptions of salesman and product

5.3.3a Participants

An opportunity sample of 20 undergraduate psychology students from the University of Stirling was taken. These participants were unfamiliar with the Open University participants used in the previous task. Three of these participants were male and 17 were female. All participants had no prior experience with video-mediated technologies, or limited experience with video-mediated technologies (for example had used such equipment on a few occasions in the past only). No details of age were taken.

5.3.3b Design

A between groups independent measures design was used. Participants viewed the recorded sales recitations from the previous experiment and were therefore randomly assigned to view either a salesman who had been given advice on video-mediated gazing, or a salesman who had been given no advice. Each participant viewed one sales pitch only. These participants were asked to rate the sales recitations as participants in the original pair had a level of familiarity with one another and therefore any ratings from these pairs may have been biased.

5.3.3c Materials

The same set-up as in the first task, except that in room 2 the video camcorder was used to play back the pre-recorded sales pitches.

A questionnaire was circulated to all participants. The questionnaire contained 8 questions, all of which were the same as in the previous study, however two questions were omitted. The questions 'how close was the interaction to face-to-face communication?' and 'how likely would you be to use the equipment again?' were dropped. The question 'how close was the interaction to face-to-face communication?' was omitted as it was believed that the question 'did you feel the salesperson was talking to you?' would be sufficient to gauge user perceptions of social co-presence. The question 'how likely would you be to use the equipment again?' was omitted as it was believed that the question 'how much did you enjoy using the equipment?' would be sufficient to gauge user perceptions of satisfaction with the technology. Furthermore, the previous study has already indicated that video-mediated gazing affects co-presence and has no effect on the assessments of the technology. This study is primarily interested in how well users adapt video-mediated gazing behaviours, and therefore measures of performance (including recall and participant evaluations of the salesperson) are more important. All questions were assessed on a 5-point scale, with one being the lowest possible score, and five being the highest possible score. So scores ranged from 'not at all' to 'very.'

5.3.3d Procedure

Participants were shown the pre-recorded sales recitations and were asked to rate the salesperson's performance on a questionnaire. Like the previous experiment, participants were informed that this was a live video feed. After the participants had seen the video,

they were requested to complete a questionnaire. After finishing the questionnaire participants were debriefed as to the deception and were given a brief overview of the aims of the study.

5.3.4 Results

Participants were rated as being significantly more friendly in the advice condition compared to the no advice condition. Participants were more inclined to believe that the salesman was talking to them in the advice condition compared to the no advice condition. Participants also enjoyed using the equipment more in the advice condition than in the no advice condition.

No significant differences were found between the conditions on the number of correct answers for the recall test. No significant differences were found for ratings of communicative ability, honesty, and confidence. The level of advice also did not affect perceptions of the product, with no significant differences for ratings of likeliness to purchase product and product uniqueness.

Table 5.2: Independent t-test scores for correct answers on recall test and additional participants raters' evaluations of salesperson, product, equipment and social co-presence between advice and no advice conditions, with mean scores and standard deviations in parentheses.

Measure	Advice Mean	No Advice Mean	Significance
Correct answers (Recall on Vernon story)	9.10 (3.57)	8.80 (3.75)	t(18) = 0.18, p > 0.05, **NS**
Question 1) How well did the salesperson communicate?	3.30 (1.34)	3.00 (1.25)	t(18) = 0.52, p > 0.05, **NS**
Question 2) How friendly was the salesperson?	4.10 (0.57)	3.10 (1.19)	t(18) = 2.39, p < 0.05
Question 3) How honest was the salesperson?	3.40 (1.08)	3.00 (1.15)	t(18) = 0.80, p > 0.05, **NS**
Question 4) How confident was the salesperson?	3.30 (1.49)	2.50 (1.35)	t(18) = 1.26, p > 0.05, **NS**
Question 5) How likely would you be to buy the product?	2.90 (1.45)	2.10 (1.10)	t(18) = 1.39, p > 0.05, **NS**
Question 6) How unique was the product?	2.70 (1.26)	2.60 (0.84)	t(18) = 0.21, p > 0.05, **NS**
Question 7) Did you feel the salesperson was talking to you?	3.90 (1.29)	2.30 (1.42)	t(18) = 2.64, p < 0.01
Question 8) How much did you enjoy using the equipment?	3.30 (0.67)	2.70 (0.82)	t(18) = 1.78, p < 0.05

Using independent-samples t-tests with a two-tailed hypothesis, differences between the conditions were tested. A significant difference was found between the conditions on perceptions of how friendly the salesperson was (t(18) = 2.39, p < 0.05). This indicates that the salesman was perceived as being friendlier when users were advised to employ video-mediated gazing, as opposed to when no advice was given. Participants also indicated a significant difference between the two conditions for the question 'did you feel the salesperson was talking to you?' (t(18) = 2.64, p < 0.01). This indicates that

participants felt that the salesperson was more likely to be speaking to them in the advice condition compared to the no advice condition. A significant difference was also found between the conditions on how much the participants enjoyed using the equipment $(t(18) = 1.78, p < 0.05)$. This indicates that participants enjoyed using the equipment more in the advice condition compared to the no advice condition.

No significant differences were found between the conditions in the number of correct answers on the recall test. Perceptions of honesty, confidence and communicative ability did not differ significantly between conditions. Perceptions of the product also did not differ between conditions, with no significant differences for the ratings of likeliness to purchase product and product uniqueness. All scores in the advice condition were however numerically higher than scores in the no-advice condition. Possible explanations for these findings will be discussed later.

Assessments were also made concerning the length of sales pitch in each condition, the total amount of time spent gazing at the camera in each condition, the total number of gazes at camera in each condition, and the proportion of time spent gazing at camera in each condition. These measurements were recorded in order to test whether participants had successfully adopted the advice. The length of the sales pitch was recorded from the point at which the salesperson began speaking, to the point at which speech concluded. Gazing was said to take place when the participant looked in a forward direction (directly at the camera). In the video recording, this behaviour was detected when, from the coder's point of view, it appeared as if eye contact was being made. A gaze was measured from the point at which gaze at the camera occurred and concluded at the point when gaze was averted from the camera. Due to the difficulty of coding such gazing behaviour an inter-

judge reliability test was conducted using one other coder. A sub-sample of 8 participant recordings was used. Agreement for total gaze count was at 98.5%.

Participants employed significantly more gazes in the advice condition than in the no advice condition. Participants also spent a significantly higher proportion of the sales pitch gazing in the advice condition than in the no advice condition. No significant differences were found between the conditions in the total time of the sales pitch and the amount of time spent gazing in seconds.

Table 5.3: Independent t-test scores for total time spent speaking, total time spent gazing, number of gazes in speech, and proportion of total speech spent gazing for advice and no advice conditions, with mean scores and standard deviations in parentheses

Measure	Advice Mean	No Advice Mean	Significance
Total time of speech	78.80 (20.73)	92.80 (36.07)	$t(18) = 1.06$, $p > 0.05$ NS
Total time spent gazing	20.10 (16.55)	9.70 (28.28)	$t(18) = 1.00$, $p > 0.05$ NS
Number of gazes in speech	18.80 (12.41)	5.90 (13.52)	$t(18) = 2.22$, $p < 0.05$
Proportion of speech gazing	31.20 (20.89)	6.60 (17.00)	$t(18) = 2.89$, $p < 0.01$

Independent measures t-tests were carried out in order to test the difference between the two conditions. A significant difference was found between conditions for the mean total number of gazes in sales pitch ($t(18) = 2.22$, $p < 0.05$). This indicates that participants in the advice condition used more gazes in total than participants in the no advice condition. A significant difference between conditions was also found in the proportion of the sales pitch spent gazing ($t(18) = 2.89$, $p < 0.01$). This reveals that participants who were advised spent a higher proportion of the sales pitch engaged in gazing behaviour than participants who were not advised. No significant difference between the conditions was found in the

total length of the sales in pitch in seconds. No significant difference was also found in the total number of seconds spent gazing, although the advice condition was numerically higher on average.

5.3.5 Discussion

Overall, there is some evidence that merely advising users to adopt video-mediated gazing behaviour is beneficial. Like the previous study, participants who are advised to gaze are regarded as being friendlier. Results from this study are also akin to the previous study in that a higher level of co-presence was formed: participant ratings on the question 'did you feel the salesperson was talking to you?' were significantly higher in the advice condition compared to the no advice condition. Participants also rated the experience with using the video-mediated equipment as significantly more enjoyable in the advice condition than the no advice condition. Once more, this result reflects the results of the previous study.

There is, however, evidence to suggest that advising users to adopt video-mediated gazing behaviours is not as successful as the previous experiment. For example, 3 of the 4 measures relating to the perception of the salesperson showed no differences between the conditions: no differences were found for ratings of communicative ability, honesty and confidence. This contrasts with the previous experiment in which the salesman received significantly higher ratings for these measures. Also, in this study there was no difference between conditions on the recall task, however in the previous study participants remembered significantly more information in the video-mediated gazing condition than the no-gazing condition. A noticeable trend in the data is also noticeable here however, as for each question and for the recall test (even when not significant) the scores are always in

higher for the advice condition. This indicates that there is insufficient power here to discriminate between conditions.

In order to test whether these lack of effects were a result of participants failing to adopt the advice effectively, the differences between the groups in gazing behaviour were assessed using post hoc t-tests. Furthermore, it was necessary to test whether these lack of effects were due to the no advice condition adopting video-mediated gazing behaviours irrespective of the fact that they were not advised to. Results, however, indicate that participants in the advice condition gazed at the camera more frequently, and spent a higher proportion of the sales pitch gazing at the camera in the advice condition compared to the no advice condition. Although the no advice condition did adopt some sort of gazing behaviour, it was significantly less frequent than the advice condition. Results also reveal that there was no difference between the conditions in the amount of time spent gazing at the camera in seconds, although numerically the advice condition spent twice as long gazing than the no advice condition. Participants in the advice condition spent on average 31.2% of the sales pitch gazing at the camera. This is similar to the previous study in which 30% of the sales pitch for product 1, and 32% of the sales pitch for product 2 were accompanied by gaze. At first glance then it would seem that the participants in this study have adopted video-mediated gazing behaviours effectively.

A number of confounding variables are evident in this study. Firstly, participants in the no advice condition did use some level of video-mediated gazing. In the previous study, the comparison for the video-mediated condition consisted of a condition in which no gazing occurred at all. Therefore, it is likely that reduced amounts of gazing are better than no gazing. Secondly, although adopting a significant amount of gazing behaviour,

participants gazed at inappropriate times. In the previous study, due to memorisation and practice, it was ensured that the salesperson looked at the camera at the exact points in the sales pitch that needed to be recalled. In this study however, the lack of time allocated to memorise and practice meant that the participants, although using a high level of gazing behaviour, failed to look up at the camera at points in the sales pitch that needed to be remembered. An analysis of a sub-sample of 5 of the recordings indicates this to be the case: only 52.63% of the items on the recall test were accompanied by gaze. In the trained study (study 4) 100% of the items on the recall test were accompanied by gaze.

Overall, this study indicates that merely advising participants to adopt video-mediated gazing behaviours leads to some level of success, however training (or indeed a higher level of preparation) may be required in order to realise the benefits found in study 4. Furthermore, the fact that some participants naturally use some gaze to the camera would indicate that this behaviour is not necessarily an unnatural response to using the media.

5.4: Study 6: Participant evaluation of most suitable level of video-mediated gazing.

5.4.1 Introduction and aims

Findings from the previous two studies suggest that employing video-mediated gazing can be beneficial. Results indicate that video-mediated gazing behaviour alters user perceptions of friendliness, confidence, communicative ability, and honesty. Video-mediated gazing increases feelings of togetherness, levels of satisfaction and also facilitates the remembering of information. It is also evident, however, that a high level of preparation is more successful than merely advising users to adopt this strategy. What should be noted, however, is that in the first study gazing frequency was set to approximately 30% of the overall speech, and in the second study, where users were advised to gaze at comfortable

levels, once more the speech was accompanied by approximately 30% gazing. Although this would indicate that this frequency of gazing is what users adopt naturally, it is still questionable as to whether this is the most suitable frequency of video-mediated gazing. It is clear that in face-to-face interaction different frequencies of gazing behaviour result in varied perceptions and cognitive effects. It is therefore of interest to note the effects of employing different frequencies of video-mediated gazing. In other words, what level can we push gaze to in order to increase performance before detrimental effects set in?

It is expected that lower frequencies of gazing and higher frequencies of gazing would be viewed less favourably than intermediate frequencies of gazing. It is also expected that any frequency of video-mediated gazing would be viewed more favourably than none at all. To test these hypotheses 5 different frequencies of gazing were used (0%, 25%, 50%, 75% and 100%). Therefore it was expected that there would be a normal distribution of the data, in that lower and higher frequencies (0% and 100%) would be viewed least favourably and intermediate frequencies (25%, 50%, and 75%) would be viewed more favourably. These expectations can be justified by research on gazing behaviour in face-to-face interactions. Abele (1981) for example reported that participants favoured people in social interactions when they employed moderate levels of gaze as opposed to low or high levels of gaze. Klienke and Nuessle (1968) have also indicated that people who look at their conversational partner infrequently are judged as 'defensive' or 'evasive.' On the other hand, those who look more frequently at their conversational partner may be judged as being 'friendly,' 'mature' and 'sincere' (Kleck and Nuessle, 1968). Doherty-Sneddon et al (1997) also indicate that high levels of gazing interfere with task performance in V.M.C, and therefore this would imply a preference for intermediate levels of gazing as opposed to high levels of gazing.

5.4.2 Participants

Forty-five participants were opportunistically sampled from the University of Stirling. Thirteen male subjects and 32 female subjects took part in the study. Participants were undergraduate and postgraduate students. All subjects were naïve as to the nature of the experiment and had no or limited videoconferencing experience. No details of age were taken. A confederate acted as salesman. The confederate was 29 years of age and male.

5.4.3 Design

An independent, between subjects design was used. Participants were split into five groups. Nine participants were randomly allocated to each condition, and therefore an equal amount of participants observed the salesman for each gaze frequency (0%, 25%, 50%, 75%, and 100%). In each condition, the frequency of gazing was the proportion of time spent gazing in the sales pitch. For each condition a pre-recorded video was used with a salesman attempting to sell a soap product ('Fresh Face' from study 4).

As well as being instructed to gaze at the camera during certain points in the monologue, the actor was also instructed to look at the monitor when not looking at the camera. The actor was given time to rehearse the sales pitches and time to practice the recitations until perfected. During the recording of the sales recitations, the experimenter ensured the actor looked up at the camera for the appropriate amount of time and also ensured the actor gazed at the monitor when not looking into the camera. If at any time the actor failed to follow the instructions appropriately, the experimenter stopped recording and the actor was instructed to begin anew. Each recorded recitation was later viewed over video in order to that the experimenter was happy that the instructions had been followed appropriately.

5.4.4 Materials

In room 1, a colour monitor (JVC TM-14EK(B)) was mounted in a wooden box, with a video camcorder (Sony CCD-TR2200EPAL) placed directly above the monitor. A microphone was placed to the right of the monitor, and video and audio quality were as high as achievable in the lab. The monitor and camcorder in room 1 were connected to room 2, where a video camcorder (Sony CCD-TR2200EPAL) was used to play the sales pitches. The monitors were 14 inches in size.

Each participant was distanced approximately one metre from the monitor and the scope of the view included the participant's face and upper body.

A questionnaire containing 9 questions was circulated to all participants. Each question was rated on a 5-point scale; with one being the lowest possible score and five being the highest possible score. So scores ranged from 'not at all' to 'very.'

The questions were as follows:

1) How well did the salesman communicate?

2) How friendly was the salesman?

3) How honest was the salesman?

4) How likely would you be to buy this product?

5) Did you feel that the salesman was talking to you?

6) How close was the video to face-to-face communication?

7) How much did you enjoy using the equipment?

8) How comfortable were you when taking part in this study?

9) How suitable was the salesman's level of gaze?

5.4.5 Procedure

Participants entered the videoconferencing room and sat down in front of a monitor. Participants were informed that they were about to witness a live presentation of someone attempting to sell them a product. They were also told that all that was required of them was that they listen to the information and fill out a questionnaire after the presentation relating to their perceptions of the salesman and the product. In order to prevent participants from talking to the salesman and thus discovering that it was actually a pre-recorded video, participants were informed that they must not converse with the salesman. Participants were informed that the experimenter wanted to ensure standardisation, needing all participants to receive the same information in roughly the same period of time. Once these instructions were relayed pre-recorded videos were played to participants (from room 2). Once the pre-recorded tapes were finished, participants filled in a questionnaire. After completion of the questionnaire, participants were debriefed as to the deception and were informed as to the nature of the experiment.

5.5.5 Results

Results indicate that participant ratings differ significantly between one or more of the gaze frequencies for the question 'how close was the experience to face-to-face interaction?' Participant ratings also differed significantly between one or more of the conditions for the question 'how suitable was the level of gaze?' No overall significant differences were found between conditions for ratings of communicative ability, friendliness, honesty, and confidence. Participant ratings also did not differ significantly overall for ratings of likelihood of purchasing the product. Ratings for 'did you feel the salesman was talking to you?' were unaffected by condition. Ratings also did not differ

significantly overall for levels of comfort with using the technology and ratings for enjoyment of using the technology.

Table 5.4: Between groups analysis of variance for questions 1-10 on questionnaire for different frequencies of gaze (0%, 25%, 50%, 75%, and 100%), with mean scores and standard deviations in parentheses.

Measure	0% Mean	25% Mean	50% Mean	75% Mean	100% Mean	Significance
Communicative ability	2.22 (.97)	3.0 (1.0)	2.33 (.87)	2.66 (1.22)	2.66 (.87)	$F_{(4,40)}$ = 0.87, $p > 0.05$ NS
Friendliness	3.11 (1.17)	3.44 (.88)	2.44 (1.13)	2.56 (1.01)	2.56 (1.01)	$F_{(4,40)}$ = 1.55, $p > 0.05$ NS
Honesty	2.11 (1.05)	3.11 (1.05)	2.56 (.88)	2.89 (1.05)	2.11 (1.17)	$F_{(4,40)}$ = 1.68, $p > 0.05$ NS
Confidence	2.22 (1.30)	3.0 (1.12)	2.0 (.71)	2.33 (.87)	2.67 (1.12)	$F_{(4,40)}$ = 1.28, $p > 0.05$ NS
Likelihood to buy product	1.89 (1.05)	2.44 (1.13)	2.44 (1.24)	2.0 (.71)	1.56 (.53)	$F_{(4,40)}$ = 1.39, $p > 0.05$ NS
Talking to you	1.56 (.73)	2.44 (.88)	2.67 (1.12)	2.44 (1.42)	2.67 (1.0)	$F_{(4,40)}$ = 1.71, $p > 0.05$ NS
Closeness to face-to-face	1.33 (.50)	2.67 (.87)	2.44 (.88)	2.22 (.97)	2.67 (.87)	$F_{(4,40)}$ = 3.97, $p < 0.05$
Enjoyment	3.11 (.60)	2.67 (.71)	2.78 (.44)	2.67 (.71)	2.89 (1.05)	$F_{(4,40)}$ = 0.58, $p > 0.05$ NS
Comfort	3.11 (1.17)	3.56 (1.01)	3.22 (.97)	2.56 (1.01)	3.0 (1.41)	$F_{(4,40)}$ = 0.93, $p > 0.05$ NS
Gaze frequency suitability	1.0 (.00)	1.78 (.44)	2.0 (1.32)	1.67 (1.0)	2.44 (1.13)	$F_{(4,40)}$ = 2.96, $p < 0.05$

Between groups analyses of variance were conducted between the 5 levels (0%, 25%, 50%, 75%, and 100%) for all questions on the questionnaire. A significant difference was found between one or more of the conditions for participant ratings for the question 'how close was the experience to face-to-face communication?' ($F_{(4,40)}$ = 3.97, $p < 0.05$), indicating

that frequency of video-mediated gazing has an effect on feelings of co-presence. A significant difference was also found for participant ratings for suitability of gaze frequency ($F(4,40) = 2.96$, $p < 0.05$), indicating that certain frequencies of video-mediated gazing are seen as more suitable than others. No overall significant differences were found for ratings of communicative ability, confidence, friendliness and honesty. No overall significant differences were found for ratings of likeliness to purchase the product. No overall significant differences were found for the question 'did you feel the salesman was talking to you?' Also, no significant differences were found for ratings of comfort with using the equipment and enjoyment of use.

Post hoc Tukey HSD tests between the different levels of gaze were carried out to test which frequency of gaze was perceived as being closer to face-to-face interaction (question 7), and which frequency of gaze was perceived as most suitable (question 10).

Participants rated the 25% condition as being a significantly more suitable level of gazing than the 0% condition ($p < 0.01$). Participants rated the 50% condition as being a significantly more suitable level of gazing than the 0% condition ($p < 0.05$). Participants rated the 75% condition as being a significantly more suitable level of gazing than the 0% condition ($p < 0.05$). Participants also rated the 100% condition as being a significantly more suitable level of gazing than the 0% condition ($p < 0.01$). These tests indicate that all levels of gazing are regarded as being more suitable than none at all. Furthermore, participants rated the 100% condition as being a significantly more suitable level of gazing than the 25% condition ($p < 0.05$). This indicates that participants seem to prefer higher levels of gazing than lower levels of gazing in video-mediated communication. No other differences were found.

Participants rated the 25% condition as being significantly closer to face-to-face communication than the 0% condition (p< 0.05). Participants also rated the 100% condition as being significantly closer to face-to-face than the 0% condition (p< 0.05). No other differences were found, although numerically all levels of gazing had higher ratings than the 0% condition.

5.5.6 Discussion

Statistical analysis indicates that all levels of video-mediated gazing (25%, 50%, 75%, and 100%) are considered to be more appropriate than none at all. In comparison to 0%, 25% and 100% levels are considered to be closer to face-to-face interaction, however numerically the 25% and 75% conditions were also higher than the 0% condition. In comparisons of the different levels of gazing (25%, 50%, 75% and 100%) however, the only difference found was that the 100% level was considered more suitable than the 25% level. It would seem then that users do not have a level of gaze preference when interaction is mediated. This contrasts with face-to-face communication, where preferences for different levels of gazing are noticeable. For instance, Argyle (1988) found that in conversation of an emotionally neutral topic, approximately 30% of the dialogue is accompanied by eye contact. In this situation participants therefore preferred moderate levels of gazing. Abele (1981) reported that moderate levels of gaze were preferred over low or high levels of gaze in social interactions. There are many possible explanations for this preference in face-to-face communication. Klienke and Nuessle (1968) for example have indicated that infrequent gazing is associated with someone who is 'defensive' or 'evasive.' Frequent gazing on the other hand is associated with someone who is 'friendly,' 'mature' or 'sincere.' Therefore, people may regard others who gaze infrequently with suspicion. This level of suspicion can be a result of the belief that the person is trying to

hide something. For example, gaze affects perceptions of credibility. Hemsley and Doob (1978) have indicated that an attorney is more likely to judge a witness as credible if he/she does not avert gaze during questioning. Higher levels of gaze are preferred in more intimate situations, in which case they indicate that someone is interested in what another is saying or is attracted to another person (Argyle, 1988). However, high levels of gazing can also be uncomfortable in some situations, and can result in feelings of being observed or being an object for another (Argyle and Williams, 1969). On the whole then it would seem that in face-to-face communication moderate levels of gazing are preferable. High levels of gazing are preferable in certain contexts (i.e. where people share a level of intimacy), but are considered intimidating and uncomfortable in the majority of contexts.

There are a number of possible explanations for the lack of differences in preference between varying levels of gazing in video-mediated communication. Doherty-Sneddon and McAuley (2000) have indicated that child witnesses are more resistant to leading questions in video-mediated communication. The researchers go on to explain this finding by saying that the children were more confident and more comfortable. This finding makes sense in terms of social co-presence. Essentially, because there was a distance between the participants (they were not in the same room), the influence of person factors was weaker. This finding was also reinforced by an increase in anti-social behaviour. Essentially, the participants felt they could get away with more because nobody was in the room with them. In this case, it can be reasoned that higher levels of gazing (which may be considered uncomfortable in certain contexts), will not have the same impact in a video-mediated environment.

Considering that a number of researchers (for example, Nichols and Champness, 1971; Strom and Buck) have found that high levels of gaze result in increased physiological arousal in a face-to-face context, it would also be interesting to see if the same effect occurs in video-mediated communication. Does the fact that communication is mediated result in gaze not having the same physiological effect? It is also interesting to note that although no preference was given between varying levels of video-mediated gazing, lower levels would probably be more useful than higher levels. This can be quantified by the findings of study 4, where a 30% gaze level aided participants to remember more information. If indeed video-mediated gazing acted as a cue to important information, theoretically a 100% gaze level would not be as effective. If gaze is constant, then individual gazes can not act as cues to important information. Furthermore, if the arousal hypothesis (Beattie, 1981) is correct, then constant levels of gaze will interfere with cognitive processing, as found by Doherty-Sneddon et al (1997). Therefore, although participants do not indicate a gaze frequency preference, it is also clear that certain levels may be more beneficial than others. In this case user perceptions do not match the psychological literature.

5.5 Chapter summary

The evidence from this chapter indicates that video-mediated gazing adds value to communication in a number of ways. Video-mediated gazing can be used to emphasise important aspects of speech, and therefore helps to facilitate the remembering of information. This effect clearly demonstrates that gaze can serve the same communicative function in video-mediated interactions as in face-to-face communication. Video-mediated gazing behaviour is also beneficial in that it helps to reduce psychological distance. In

addition to this participants who employ video-mediated gazing are also perceived more favourably than participants who do not employ gazing behaviour. Participants are perceived as friendlier, more honest, better communicators and more confident. This finding indicates that individuals are assessed on their use of non-verbal behaviour in video-mediated contexts, much in the same way that people are judged in face-to-face contexts. Participants can therefore use video-mediated gazing to project their images more effectively. This may be useful for a number of applications, for example interviews or presentations.

One criticism that may be levelled at the use of video-mediated gazing concerns the idea that when using this technique fewer opportunities will be available to pick up non-verbal information from conversational partners. If participants are looking up at the camera, then they will not be able to see the image of their partner on the monitor. If we consider normal face-to-face communication, however, it is clear that participants do not stare continuously at the person with whom they are conversing. In fact, it has been estimated that listeners spend approximately 75% of conversational time looking at speakers, and speakers spend approximately 40% of conversational time looking at listeners (Argyle, 1988). Therefore, in normal face-to-face interactions conversational participants speakers will have limited opportunity to pick up non-verbal information. Research also suggests that over-gazing can be detrimental, for example Beattie (1982) indicates that individuals can become distracted by partners faces, and this effects cognitive performance. Doherty-Sneddon et al (1997) also point out that over-gazing in a video-mediated context can affect task performance. Video-mediated gazing will therefore give participants ample opportunity to look away from their partner's face. Furthermore research from study 6 indicates that participants have no gaze preference when communication is mediated. This

means that levels of gazing would be as effective as higher levels of gazing. Participants could therefore use lower levels of gazing, which would give them ample opportunity to look at the monitor (and thus picking up crucial non-verbal information).

Chapter 6: Thesis Conclusions

6.1: Introduction

The aim of this thesis was to investigate the psychological and communicative implications of using video-mediated technologies, in order to determine whether certain practices could improve user use of current technology. The effects of two practices in particular were examined: face-to-face familiarisation and video-mediated gazing. These practices were devised after a review of the literature and analyses of questionnaire and interview data. It was expected that both practices would lead to benefits in communication.

Perceptions of video-mediated communication were assessed through the use of a questionnaire and an interview with two experts. This investigation was conducted in order gauge user perceptions concerning whether V.M.C is a viable alternative to face-to-face communication. The motivations behind these analyses were to discover the problems that exist in the use of the technology and methods that can be used to improve user use of V.M.C. These investigations were also conducted in order to validate the experimental focus of this thesis.

The effects of face-to-face familiarisation were analysed using a multi-level approach. Participant perceptions were key to the analysis. In particular, perceptions of social co-presence, conversational partners, and the media were investigated. Task outcome was investigated using the 'collaborative map reading task,' in which participants had to work together in order to complete the task successfully. Verbal aspects of process were also investigated through making global linguistic assessments (for example turn length, dialogue length, and the number of interruptions). Verbal aspects of process were also measured using Conversational Games analysis, where the functions of participants' utterances were assessed. In particular, the number of Align and Check games used by

participants were investigated. The motivations behind these analyses were to see if face-to-face familiarisation could be used to close intimacy distance and therefore improve communication and alleviate many of the negative effects associated with a loss of co-presence.

The effects of video-mediated gazing were analysed using a multi-level approach. Once more participant perceptions were key to the analysis. In particular perceptions of another individual's job suitability and personal attributes (for example intelligence, confidence, and friendliness). Perceptions of social co-presence and the media were also investigated. Task performance was investigated using recall tests, in order to investigate whether video-mediated gazing could facilitate the remembering of information. Non-verbal aspects of process were measured through an analysis of gazing behaviour, in particular the number of gazes employed and the length of time spent gazing. The motivations behind these analyses were to see if video-mediated gazing could be used to close intimacy distance and allow participants to project themselves more effectively. All three investigations are now discussed.

6.2: Investigations into user perceptions of V.M.C

Looking at user perceptions of video-mediated communication gives useful insights into the problems experienced by users of the technology as it exists currently. Such insights may also provide useful suggestions for the improvement of video-mediated communication. This approach has been used by a number of researchers. O'Conaill et al (1993) for example asked participants to evaluate video-mediated communication in comparison to audio-only interactions. One resounding finding was that participants preferred the use of V.M.C to audio-only communication, and this was mainly credited to

access to visual information. Many participants however noted that the quality of the system was also an important factor, with poor quality signals making it difficult to discriminate between speakers. Furthermore, a number of participants noted that in group interactions it was difficult to detect whom certain remarks were being directed to. Armstrong-Stassen et al (1998) noted that students in a distance learning experiment reported a number of problems. These participants noted that it was difficult to connect with the lecturer, the environment was distracting and impersonal, and there was a noticeable loss of classroom exchange. Many of these problems relate to a loss of social co-presence. Many researchers (for example Daly-Jones et al, 1998) have also noted that participants are concerned with the loss of natural eye contact in a video-mediated setting.

An evaluation of the questionnaire data (study 1a) revealed a number of interesting findings. Participants agreed that video-mediated communication has many advantages over audio-only interactions. The main advantage concerned the ability to portray visual signals, in particular non-verbal behaviours such as facial expressions and body language. In this sense participant perceptions agreed with the literature in that non-verbal behaviours are considered important aspects of human communication. The participants in the questionnaire also indicated however that video-mediated communication is less satisfactory than face-to-face communication. These perceptions once more reflect the communication literature, in that there is a hierarchy of preference in relation to richness of media. Participants also made a number of suggestions to explain why video-mediated communication was less satisfactory than face-to-face communication. The main reason for this preference concerned the idea that video-mediated interactions result in lower levels of social co-presence. In other words, participants felt isolated or distanced from their conversational participants. Furthermore, a number of participants suggested that

face-to-face contact was essential at some point and for certain types of meetings (for example negotiation) face-to-face contact was preferable. This finding agrees with Abbott et al (1993) in that face-to-face meetings should be used to supplement video-mediated communication. In addition to this participants also noted that V.M.C does not take place in a co-present environment. This results in a loss of social exchange, for example one participant noted that it was impossible to go for a drink or pass someone a jug of water.

A number of participants in the questionnaire study (1a) also indicated that V.M.C does not replicate natural eye contact. Further to this, it is suggested that natural eye contact is essential for a number of applications (for example business meetings or negotiation). It would therefore seem that one of the major barriers preventing V.M.C from being as successful as face-to-face communication is a loss of co-presence. This is partly credited to physical remoteness and partly a result of the attenuation of visual cues.

A number of suggestions have been made in order to deal with the problems associated with video-mediated communications. Participant perceptions have therefore often lead to suggestions for how best to optimise the use of video-mediated technologies. Sellen (1992) for example noted that video systems could be modified in order to deal with many of the problems pointed out by participants in O'Conaill et al's (1993) study. Sellen suggested that the implementation of directional audio and video might help participants detect who comments are aimed at. This can be achieved through the use of Media Spaces, which are V.M.C systems that are specifically designed to mimic the sensation of being in a meeting room by using multiple monitors positioned around a table. The inability of video-mediated technologies to replicate natural eye contact has also lead to a number of novel solutions. One example includes the use of specially placed mirrors ('videotunnels,'

Smith et al, 1991) which provide participants with the ability to use natural eye contact in video-mediated settings. One problem with many of these suggestions is that they focus on technological alterations or advancements. In such cases many users may be either unwilling or unable to purchase such technology. Many systems are also not readily available on the market. This would suggest that methods should be devised which improve communication over the types of systems that already readily available. Furthermore, such methods should not find their basis in technology. Many researchers have found that, through participant perceptions and evaluations, a number of non-technological practices can be employed for more successful communication. Alexander et al (1999) and Hackman et al (1990) for example noted that employing practices such as personal introductions, games to break the ice, or panning the camera around the room could be beneficial.

Although the participants in study 1b proposed a number of solutions to the problems associated with video-mediated communications, many of the solutions were technologically based. For example, a number of participants proposed the use of larger monitors, and therefore taking advantage of non-verbal behaviours. The purpose of this thesis however was to devise methods to improve video-mediated communication which did not rely on technological advancements or alterations. It was regarded more important that users were trained or advised to improve the use of the technology, as opposed to changing the technology to suit the user. A number of suggestions that came out the data were considered useful, as they did not focus on technological advancements, and they addressed the loss of co-presence in V.M.C. One suggestion to help with the loss of co-presence involved initial face-to-face meetings. Indeed, it was suggested that this practice would alleviate many of the negative effects of V.M.C, and this would be achieved through

reducing psychological distance. Another participant indicated that some form of gazing behaviour could be made by looking directly into the camera. This particular practice was considered useful, as potentially it would help to reduce psychological distance. Furthermore, gaze serves a number of important communication functions, and therefore allowing users to gaze in V.M.C could potentially improve communication. These ideas were therefore taken forward into the experimental chapters. Consequently the effects of video-mediated gazing and face-to-face familiarisation were investigated.

Care must be taken when drawing conclusions from this research. Participants were drawn from a number of different backgrounds, and therefore there was no standardisation with respect to the systems being used and the type of applications the systems were being used for. Previous research has indicated that differing systems result in varying levels of communicative success. O'Conaill et al (1993) for example have shown that better quality video systems (LIVENET) produced more similar communication to face-to-face interaction than poor quality systems (ISDN). The task difference literature also indicates benefits in using video-mediated technologies for socially orientated tasks, but little benefit for problem-solving tasks. In this respect problems reported by one user may not be relevant to another user. For example, problems with turn-taking are going to be more prevalent to individuals who use such systems for two-way communication and less so for one-way communication (e.g a lecture). Further to this a number of participants in this study had only used this type of technology on one occasion. Furthermore, these participants had taken part in a videoconferencing study and therefore the task they had performed was not indicative of the kinds of applications this technology is used for in the real world. Also, these participants used a high quality system (large image with no delay) and this is not reflective of the kinds of systems used in every day settings.

6.3: Investigations into the effects of face-to-face familiarisation

Chapter 4 aimed to test the effects of face-to-face familiarisation on video-mediated communication. It was expected that face-to-face familiarisation prior to video-mediated meetings would reduce psychological distance between participants, and furthermore this would lead to a number of psychological and communicative benefits. In order to test this, two investigations took place. Study 2 tested the effects of face-to-face familiarisation on user perceptions across a video link. Study 3 reported findings for the effects of face-to-face familiarisation on task performance using a 'collaborative map reading task.'

Evidence from study 2 indicates that face-to-face familiarisation results in a number of psychological and communicative benefits. Participants are evaluated more favourably in a video-mediated context after being met face-to-face. For example, they are perceived as being better communicators, more suitable for a job as a member of bar staff, more intelligent and more honest. This finding would suggest that the influence of personal factors is stronger in a face-to-face context, adding weight to the claims of Whitney and Cook (1990). Participants were asked to rate the interviewee on their performance in the video-mediated interview, however it is clear that initial meeting context affected later V.M.C perceptions. Interviewees in a video-mediated environment can be judged upon a number of factors, for example their use of the English language. In a face-to-face setting, however, individuals can also be judged upon their use of non-verbal behaviour. An individual's use of non-verbal behaviour may influence another's perception of them. This has been indicated in a number of studies on gazing behaviour, where gaze seems to influence evaluations of honesty, intelligence, and a number of other personal attributes. This finding can also be explained in terms of co-presence. Face-to-face familiarisation

resulted in a reduction in psychological distance, and therefore the video-mediated interview was perceived as being closer to a face-to-face conversational context. This increased level of co-presence between the participant and the interviewee may have therefore fed into the more favourable evaluations.

Face-to-face familiarisation therefore helps to reduce psychological distance, resulting in higher levels of co-presence and a more satisfactory communicative process. This finding can be explained in terms of Argyle and Dean's equilibrium theory (1965). This theory suggests that individuals attempt to maintain a comfortable intimacy distance in everyday interactions. Intimacy distance depends upon a number of factors, for example the relationship between the interlocutors. Furthermore, there are many cues to intimacy, for example eye contact, facial expressions and touch. In communications that do not allow participants to express these cues to intimacy, lower levels of co-presence will be experienced (Doherty-Sneddon and McAuley, 2000). It therefore follows that through allowing participants to express cues for intimacy in the face-to-face meetings, psychological distance was reduced and perceived levels of co-presence were improved. This increased social co-presence lead to an increase in satisfaction, which may have also fed into the increased participant evaluations of the interviewee.

Study 3 aimed to test the effects of face-to-face familiarisation on task performance. The results indicate that face-to-face meetings result in a number of communicative benefits in comparison to when no form of initial meeting takes place. In particular, face-to-face familiarisation results in a more efficient use of language, for example the end goal of communication is carried out in a shorter period time and in fewer words. Face-to-face familiarisation also results in participants having to check less often that their partner

understands them and check less often that they have understood their partner. This finding suggests that familiarisation increases mutual understanding, and therefore allows participants to be better monitor understanding and attention. This finding is line with Clarke (1996) who asserted that familiar people have to work less hard to achieve levels of mutual understanding. This indicates that a meeting as short as 10 minutes can improve levels of mutual understanding. Would there be any added benefit in conducting familiarisation meetings for longer periods of time? Familiarisation did not however affect task performance, as there were no differences between conditions with regards to how efficiently participants collaborated.

Although study 3 revealed that face-to-face familiarisation results in a number of benefits, the fact that there were no differences on any of the measures between the face-to-face and video meeting conditions, would suggest that familiarisation rather than face-to-face contact per se is important here. Although it should be noted that the face-to-face meeting condition scored numerically better on all measures in comparison with the video meeting conditions. This would indicate that there was insufficient power of the dependent variable to discriminate between conditions.

Overall, the evidence from studies 2 and 3 suggest that face-to-face familiarisation results in a number of psychological and communicative benefits. Therefore these findings add weight to Abbott et al's (1993) claim that video-mediated communications should be supplemented by face-to-face meetings for more effective communication. It is further suggested that the precise importance of face-to-face familiarisation is dictated by factors such as the level of relationship between participants and the type of task that is being employed. It is expected that face-to-face familiarisation would be of more benefit to

strangers (or indeed those who have not met face-to-face), and with tasks that involve a more social dimension, for example negotiation or tasks that involve judging another person.

6.4: Investigations into the effects of video-mediated gazing

Chapter 5 aimed to test the effects of video-mediated gazing on communication. It was expected that video-mediated gazing would reduce psychological distance, and would allow participants to project their images more effectively. In order to test this, study 4 investigated the effects of video-mediated gazing on recall and perceptions of a salesman and product. Study 5 tested whether users of the technology could adopt video-mediated gazing behaviours readily. Study 6 tested levels of gaze preference in video-mediated interactions.

Evidence from study 4 suggests that video-mediated gazing results in a number of communicative and psychological benefits. For example, participants who employ video-mediated gazing behaviours are perceived more favourably than participants who adopt no form of gazing behaviour. For example, perceptions of honesty, friendliness, confidence and communicative ability were affected by gaze access. This finding suggests that gaze serves similar functions in video-mediated interactions as it does in face-to-face communications. An abundance of research suggests that participants are viewed more favourably in face-to-face contexts when the employ frequent gazing behaviours. Results from study 4 also indicate that video-mediated gazing aids participants to recall information more effectively. Once again, this finding demonstrates that gaze can serve the same function in video-mediated communication as it does in face-to-face interactions. Fry & Smith (1975) and Otteson and Otteson (1980) for example have indicated that

individuals remembered more information from others who gazed at them frequently. This finding can be explained in a number of ways. Eye gaze may have acted as cue to important information, much in the same that eyebrows are used to emphasise important aspects of speech (Ekman, 1979; Whittaker and O'Conaill, 1997). This finding may also be explained in terms of an increased level of physiological arousal on the part of the participant. Levels of gaze can lead to physiological arousal (Bettie, 1981). There is also a relationship between physiological arousal and cognitive performance. Both high and low levels of arousal are counterproductive with regards to cognitive performance, however moderate levels of arousal are beneficial (Yerkes and Dodson, 1908). It therefore follows that moderate levels of gazing have lead to moderate levels of arousal, which in turn affected cognitive performance. Results from this study also indicate that participants perceive video-mediated meetings where gaze is employed as having higher levels of co-presence, than in video-mediated contexts where no gaze is evident. This finding once more indicates that through increasing participant access to cues to intimacy, psychological distance is reduced and consequently higher levels of co-presence are formed

Study 5 was conducted in order to test whether participants could adopt video-mediated gazing behaviours readily. Results indicate that while merely advising users to adopt such behaviours is beneficial (for example participants are perceived as friendlier and perceptions of co-presence are improved), a number of benefits found in study 4 were not realised. This would imply that some form of training would be required in order to get the best use out of this technique. Indeed, further analysis indicates that even though participants gazed significantly more at the camera after being advised, this gazing behaviour did not accompany the appropriate material (hence why there was no difference in the amount of information recalled by participants).

Study 5 aimed to test levels of gaze preference in video-mediated communications. The results from this study show that while all levels of gaze are preferred over none, there seems to be no gaze preference when communication is mediated. These findings differ from research into face-to-face interactions, which indicate that individuals prefer moderate levels of gaze over high and low levels of gaze (Abele, 1981). This preference makes sense if we consider that low levels of gaze are associated with someone who is defensive and evasive (Klineke & Nuessle, 1968), and that high levels of gaze result in feelings of being observed or being an object for another (Argyle and Williams, 1969). This is not always the case, however, and may depend upon the relationship between those who are communicating. For example, in more intimate situations, high levels of gaze are considered normal (Argyle, 1988). Although participants did not indicate a level of gaze preference, it is also suggested that some levels of gaze may be more useful than others. For example, it is suggested that high and low levels of gaze would not help individuals to remember more information.

6.5: Conclusion

Video-mediated gazing and face-to-face familiarisation are practices that can be used to improve video-mediated communication. Furthermore, these practices are not grounded in the technology and therefore it is feasible that, in most cases, these practices will not require users to purchase new technology, or alter existing technology. Evidence from this thesis indicates that video-mediated gazing results in participants projecting their image more effectively and consequently results in improved evaluations. Video-mediated gazing can be used to perform face-to-face gaze functions, in particular to emphasise important information. In addition to this video-mediated gazing helps to reduce psychological distance and therefore results in a more satisfactory communicative

experience (on behalf of the participant). Face-to-face familiarisation also helps to reduce psychological distance, resulting in improved user perceptions of social co-presence. Face-to-face familiarisation also results in improved participant perceptions of others, and can even influence perceptions of job suitability. This thesis therefore illustrates that the use of simple practices can improve video-mediated communication.

Although these practices have been shown to result in a number of communicative benefits, it is difficult to generalise these findings to all video-mediated systems. For instance, the types of tasks used in this thesis, although varied, do not fully represent the types of applications for which these types of systems are used for in the real world. It should also be noted that the participants in these investigations were in most cases students from the University of Stirling. In this case it is clear that these participants had limited or no experience with using such technology and also did not use such communication systems on a regular basis. For this reason, these participants are not representative of the kinds of individuals who do use such conferencing systems. The video configuration was also not representative of all the types of systems used in every day situations. As has been discussed in the literature, there are a number of different systems available on the market, and these not only differ with regards to quality (for example audio and video quality), but also with regards to set-up (for example the manner in which individuals are presented on screen). In these investigations the scope included the head and body; participants were usually distanced approximately one metre from the screen, and audio and video quality were as high as achievable in the lab. It should be noted that this type of configuration is only one of many possibilities, and therefore the practices used in this thesis should be tested with a number of other configurations in order for these findings to be generalised.

References

Abbott, L., Dallat, J., Livingston, R., & Robinson, A. (1993). The Application of Videoconferencing to the Advancement of Independent Group Learning for Professional Development. *ETTI*, 31 (2).

Abel, A. (1981). Acquaintance and Visual Behaviour between Two Interactants: Their Communicative Function for the Impression Formation of an Observer. *European Journal of Social Psychology*, 11, 409-425.

Abel, M. (1990). Experiences in an Exploratory Distributed Organisation. **In** J. Galegher, R. Kraut, & C. Egido (Eds). *Intellectual Teamwork* (pp 189-510). Hillsdale, N.J: Lawrence Erlbaum Associates.

Alexander, W., Higgison, C., & Mogey, N. (Eds). (1999). *Videoconferencing in Teaching and Learning: Case Studies*. LTDI and TALiSMAN, Institute for Computer Based Learning, Heriot-Watt University, Edinburgh.

Altmann, S.A. (1967). *Social Communication Among Primates*. University of Chicago Press: Chicago.

Anderson, A.H., Bard, E., Sotillo, C., Doherty-Sneddon, G., & Newlands, A. (1997). The Effects of Face-to-Face Communication on the Intelligibility of Speech. *Perception and Psychophysics*, 59, 580-592.

Anderson, A.H., O'Malley, C., Doherty-Sneddon, G., Langton, S., Newlands, A., Mullin, J., Fleming, A.M., & Van der Velden, J. (1997b). *The Impact of VMC on Collaborative Problem Solving: An Analysis of Task Performance, Communicative Process, and User Satisfaction.* **In** K.E. Finn, A.J. Sellen, S.B. Wilbur (1997). *Video Mediated Communication.* Lawrence Erlbaum Associates, New Jersey.

Anderson, A.H., Smallwood, L., MacDonald, R., Mullin, J., Fleming, A., & O'Malley, C. (1999). *Video Data and Video Links in Mediated Communication: What Do Users Value?* Under review for International Journal of Human Computer Studies.

Angiolillo, J.S., Blanchard, H.E., Israelski, E.W., & Mane, A. (1997). *Technology Constraints of Video-Mediated Communication.* **In** K.E. Finn, A.J. Sellen, S.B.Wilbur (1997). *Video Mediated Communication.* Lawrence Erlbaum Associates, New Jersey.

Argyle, M. (1969). *Social Interaction.* London: Methuen.

Argyle, M. (1988). *Bodily Communication (second edition).* Routledge: London.

Argyle, M., & Dean , J. (1965). Eye contact, Distance, and Affiliation. *Sociometry*, 28, 289-304.

Argyle, M., & Graham, J. (1977). The Central European Experiment: Looking at Persons and Looking at Things. *Journal of Environmental Psychology and Nonverbal Behaviour*, 1, 6-16.

Argyle, M., & Kendon, A. (1967). The Experimental Analysis of Social Performance. *Advances in Experimental Social Psychology*, 3, 55-98.

Argyle, M., Lalljee, M., & Cook, M. (1968). The Effects of Visibility on Interaction in a Dyad. *Human Relations*, 21 (1), 3-17.

Argyle, M. & Williams, M. (1969). Observer or Observed? A Reversible Perspective in Person Perception. *Sociometry*, 32, 369-412.

Armstrong-Stassen, M., Landstrom, M., & Lumpkin, R. (1998). Students' Reactions to the Introduction of Videoconferencing for Classroom Instruction. *The Information Society*, 14, 153-164.

Backhouse, A., & Drew, P. (1992). The Design Implications of Social Interaction in a Workplace Setting. *Environment and Planning B: Planning and Design*, 19, 573-584.

Bales, R.F. (1950). *Interaction Process Analysis*. Addison-Wesley: Cambridge, Mass.

Barber, P., & Laws, J. (1994). Image Quality and Video Communication. **In** R. Damper, W. Hall, & J. Richards (Eds). *Proceedings of the IEEE International Symposium on Multimedia Technologies and their Future Applications* (pp 163-178), London: Pentech Press.

Baron, R.A., & Byrne, D. (1984). *Social Psychology* (4th Edition). Allyn and Bacon Inc.

Beattie, G.W. (1980). The Role of Language Production in the Organisation of Behaviour in Face-to-Face Interaction. In B.Butterworth (Ed), *Language Production* (pp 69-107). London: Academic Press.

Beattie, G.W. (1981). The Regulation of Speaker Turns in Face-to-Face Conversation: Some Implications for Conversation. *Semiotica*, 34, 55-70.

Beattie, G.W. (1981b). A Further Investigation of the Cognitive Interference Hypothesis of Gaze Patterns During Conversation. *British Journal of Social Psychology*, 20, 213-248.

Birdwhistell, R. (1970). *Kinesics and Context: Essays in Body Motion Communication*. Harmondsworth: Penguin.

Bly, S., Harrison, S., & Irwin, S. (1993). Media Spaces: Bringing People Together in a Video, Audio and Computing Environment. *Communications of the ACM*, 36, 28-45.

Boyle, E.A., Anderson, A.H., & Newlands, A. (1994). The Effects of Visibility on Dialogue and Performance in a Co-operative Problem Solving Task. *Language and Speech*, 1994, 37(1), 1-20.

Brown, G., Anderson, A.H., Yule, G., & Shillcock, R. (1984). *Teaching Talk*. Cambridge University Press: Cambridge.

Bull, P. (1978). The Interpretation of Posture Through an Alternative Method to Role Play. *British Journal of Social and Clinical Psychology*, 17, 1-6.

Bull, P.E. (1983). *Body Movement and Interpersonal Communication*. Chichester: Wiley.

Burns, J.A., & Kintz, B.L. (1976). Eye Contact While Lying During an Interview. *Bulletin of the Psychometrics Society*, 7, 87-89.

Burton, A., McGregor, H., & Berry, P. (1979). The Effects of Social Reinforcement on Dominant and Dependent Mildly Intellectually Handicapped School-Leavers. *British Journal of Social and Clinical Psychology*, 18, 129-133.

Campbell, J.A. (1998). Participation in Videoconferenced Meetings: User Disposition and Meeting Context. *Information and Management*, 34, 329-338.

Chapanis, A., Ochsman, R.B., Parrish, R.N., & Weeks, G.D. (1972). Studies in Interactive Communication: The Effects of Four Communication Modes on the Behaviour of Teams during Cooperative Problem Solving. *Human Factors*, 14(6), 487-509.

Chapanis, A. (1975). Interactive Human Communication. *Scientific American*, 232, 3-42.

Chapanis, A. (1988). Interactive Human Communication. **In** I.Grief (Ed) *Computer-Supported Cooperative Work: A Book of Readings*, Morgan Kauffman Publishers, San Mateo, California, 127-142

Chomsky, N. (1957). *Syntactic Structures*. S'Gravenhage, Netherlands: Mouton.

Church, R.B., & Goldin-Meadow, S. (1986). The Mismatch Between Gesture and Speech as an Index of Transitional Knowledge. *Cognition*, 23, 47-71.

Clark, H.H. (1996). *Using Language*. Cambridge: CUP.

Clark, H.H., & Brennan, S.E. (1991). *Grounding in Communication*. **In** L.B Resnick, J. Levine & S.D. Teasley (Ed's). *Perspectives on Socially Shared Cognition*. Washington, D.C: American Psychological Association.

Clark, H., & Marshall, C. (1981). Definite Reference and Mutual Knowledge. **In** A. Joshi, B. Webber & I. Sag (Eds), *Elements of Discourse Understanding* (pp 10-63). Cambridge: Cambridge University Press.

Clark, H., & Schaefer, E. (1989). Contributing to Discourse. *Cognitive Science*, 13, 259-292.

Clark, H., & Wilkes-Gibbs, D. (1986). Referring as a Collaborative Process. *Cognition*, 22, 1-39.

Cohen, K. (1982). Speaker Interaction: Video Teleconferences Versus Face-to-Face Meetings. *Proc. Teleconferencing and Electronics Communications*, 189-199.

Cook, M., & Lalljee, M. (1972). Verbal Substitutes for Visual Signals in Interaction. *Semiotica*, 3, 212-221.

Daft, R.L, & Lengel, R.H. (1986). A Proposed Integration Among Organisational Information Requirements, Media Richness, and Structural Design. *Management Science*, 32(5), 554-571.

Dallat, J., Fraser, G., Livingston, R., & Robinson, A. (1992). *Videoconferencing and the Adult Learner*. University of Ulster.

Daly-Jones, O., Monk, A., & Watts, L. (1998). Some Advantages of Video Conferencing Over High-quality Audio Conferencing: Fluency and Awareness of Attentional Focus. *Journal of Human-Computer Studies*, 49, 21-58.

Davies, M.A. (1971). *Communication Effectiveness as a Function of Mode.* Unpublished master's thesis, University of Waterloo, Ontario, Canada.

Davies, M.F. (1971b). Cooperative Problem Solving: An Exploratory Study. Communications studies group paper No E/71159/DV.

Davies, G. & Noon, E. (1991). *An Evaluation of the Live Link for Child Witnesses.* Home Office report.

De Greef, P., & Ijsselsteijn, W.A. (2001). Social Presence in a Home Tele-application. *Cyber-Psychology and Behaviour*, 4(2), 307-315.

Dennis, A.R, & Kinney, S.T. (1998). Testing Media Richness Theory in the New Media: The Effects of Cues, Feedback, and Task Equivocality. *Information Systems Research*, 9(3), 256-273.

Doherty-Sneddon, G., Anderson, A.H., O'Malley, C., Langton, S., Garrod, S., & Bruce, V. (1997). Face-to-face and Video-mediated Communication: A Comparison of Dialogue Structure and Task Performance. *Journal of Experimental Psychology*, 3(2), 105-125.

Doherty-Sneddon, G., & McCauley, S. (2000). Influence of Video-mediation on Adult-Child Interviews: Implications for the Use of the Live Link with Child Witnesses. *Applied Cognitive Psychology*, 14, 379-392.

Dublin, R., & Hedley, R.A. (1969). *The Medium may be Related to the Message: College Instruction by TV.* Eugene: University of Oregon Press.

Duncan, S. (1972). Some Signals and Rules for Taking Speaker Turns in Conversation. Journal of Personality and Social Psychology, 23, 283-292.

Duncan, S. & Fiske, D.W. (1977). *Face-to-Face Interaction: Research, Methods and Theory.* Hillsdale, NJ: Lawrence Erlbaum.

Duncan, S. & Niederehe, G. (1974). On Signalling that it's Your Turn to Speak. *Journal of Experimental Social Psychology*, 10, 234-247.

Efron, D. (1972). *Gesture, Race and Culture.* The Hague: Mouton.

Egido, C. (1990). Teleconferencing as a Technology to Support Co-operative Work: A Review of its Failures. **In** J. Galegher, R. Kraut, & C. Egido (Eds), *Intellectual teamwork* (pp 351-372). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ekman, P., & Friesen, W.V. (1969). The Repertoire of Nonverbal Behaviour: Categories, Origins, Usage and Coding. *Semiotica*, 1, 49-98.

Ekman, P. (1979). About Brows: Emotional and Conversational Signals. **In** M. von Cranach, K. Foppa, W. Lepenies and D. Ploog (Eds), *Human Ethology*, Cambridge: Cambridge University Press.

Ekman, P., & Friesen, W.V. (1975). *Unmasking the Face*. Englewood Cliffs: Prentice Hall.

Ellis, A., & Beattie G. (1986). *The Psychology of Language and Communication*. Hillsdale, N.J: Erlbaum.

Ellsworth, P.C. (1975). *Direct Gazing as a Social Stimulus: The Example of Aggression*. **In** P.Pliner, L. Krames, & T. Alloway (Eds). *Nonverbal Communication of Aggression* (pp 53-76). New York: Plenum Press.

Evans, M., and Rubin, K. (1979). Hand Gestures as a Communicative Mode in School-Aged Children. *Journal of Genetic Psychology*, 135, 189-196.

Exline, R.V., Gray, D., & Schuette, D. (1965). Visual Behaviour in a Dyad as Affected by Interview Content and Sex of Respondent. *Journal of Personal Social Psychology*, I, 201-209.

Exline, R.V., & Eldridge, C. (1967, April). *Effects of Two Patterns of a Speaker's Visual Behaviour upon the Perception of the Authenticity of his Verbal Message.* Paper presented at the meeting of the Eastern Psychological Association, Boston.

Exline, R.V. (1971). *Visual Interaction: The Glances of Power and Preference.* In J.K. Cole (Ed), *Nebraska Symposium on Motivation* (19, pp 162-205). Lincoln: University of Nebraska Press.

Finn, K.E. (1997). *Introduction: An Overview of Video-mediated Communication Literature.* **In** K.E. Finn, A.J. Sellen, & S.B. Wilbur. *Video Mediated Communication.* Lawrence Erlbaum Associates, New Jersey.

Fish, R., Kraut, R., & Chalfonte, B. (1990). The Videowindow System in Informal Communications. *Proceedings of Conference on Computer Supported Cooperative Work* (pp 1-12). New York, ACM.

Fish, R., Kraut, R., Root, R., & Rice, R. (1993). Video as a Technology for Informal Communication. *Communications of the ACM*, 36, 48-61.

Frick, R.W. (1985). Communicating Emotion: The Role of Prosodic Features. *Psychological Bulletin*, 97, 412-429.

Frolich, D. (1995). Requirements for Interpersonal Information Management. **In** P.J. Thomas (Ed). *Personal Information Systems: Business Applications* (pp 133-153), Cheltenham: Stanley Thornes and Unicorn Seminars.

Fry, R., & Smith, G.F. (1975). The Effects of Feedback and Eye Contact on Performance of a Digit-Coding Task. *Journal of Social Psychology*, 96, 145-146.

Fussell, S.R., & Benimoff, N.I. (1995). Social and Cognitive Processes in Interpersonal Communication: Implications for Advanced Telecommunications Technologies. *Human Factors*, 37, 228-250.

Gale, S. (1989). Human Aspects of Interactive Multimedia Communication. *Interacting with Computers,* 2(2).

Gale, C., & Monk, A. (2000). Where am I Looking? The Accuracy of Video-mediated Gaze Awareness. *Perceptions and Psychophysics*, 62, 586-595.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.

Gaver, W., Moran, T., Maclean, A., Lovstrand, L., Doursih, P., Carter, K., & Buxton, W. (1992). Realizing a Video Environment: EuroPARC's RAVE System. **In** P. Bauersfeld, J. Bennett, & G. Lynch (Eds). *CHI' 92 Conference Proceedings* (pp 27-35), New York: ACM Press.

Gaver, W., Sellen, A., Heath, C., & Luff, P. (1993). One is Not Enough: Multiple Views in a Media Space. Proceedings of CHI 93 *Human Factors in Computing Systems* (pp 335-341), ACM: New York.

Glenberg, A.M., Schroeder, L., & Robertson, D.A. (1998). Averting the Gaze Disengages the Environment and Facilitates Remembering. *Memory and Cognition,* 26 (4), 651-668.

Glucksberg, S., & Krauss, R.M. (1967). What Do People Say After They Have Learned to Talk? Studies of the Developmental of Referential Communication. *Merril-Palmer Quarterly,* 13, 309-316.

Goffman, E. (1971). *Relations in Public.* London: Allen Lane/Penguin Press.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing Knowledge through Gesture: Using Children's Hands to Read Their Minds. *Cognition and Instruction,* 9(3), 201-219.

Goldman-Eislerm F. (1968). *Psychololinguistics: Experiments in Spontaneous Speech.* London: Academic Press.

Goodman, G.S., Tobey, A.E., Batterman, F.Y.M., Orcutt, H., Thomas, S., Shapiro, C., & Sachsenmaier, T. (1998). Face-to-face Confrontation: Effects of Closed-circuit Technology on Children's Eyewitness Testimony and Juror's Decisions. *Law and Human Behaviour,* 22, 165-203.

Goodwin, C. (1979). The Interactive Construction of a Sentence in Natural Conversation. **In** G. Psathas (Ed). *Everyday Language: Studies in Ethnomethodology.* New York: Irvington Publishers, Inc.

Goodwin, C. (1981). *Conversational Organisation: Interaction between Speakers and Hearers.* New York: Academic Press.

Graham, J.A. & Argyle, M. (1975). A Cross-cultural Study of the Communication of Extra Verbal Meaning by Gestures. *International Journal of Psychology,* 10, 56-67.

Greenbaum, P.F., & Rosenfeld, H.M. (1980). Varieties of Touching in Greeting: Sequential Structure of Sex-related Differences. *Journal of Nonverbal Behaviour,* 5, 13-25.

Grice, H.P. (1969). Utterer's Meaning and Intentions. *Philosophical Review,* 147-177.

Habash, T.F. (1999). The Impact of Video Mediated Interaction and Group Decision Support Tools on Group Perception and Satisfaction: An Experimental Investigation in Distributed Meeting Environment. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 60(1a).

Hackman, M., Walker, Z., & Kim, B. (1990). Instructional Communication in the Televised Classroom: The Effects of System Design and Teacher Immediacy on Student Learning and Satisfaction. Communication Education, 39(3), 196-206.

Harding, G. (1983). Setting the Stage for Language Acquisition: Communication Development in the First Year. **In** R.M. Golinkoff (Ed), The Transition from Prelinguistic to Linguistic Communication. Hillsdale, N.J, Erlbaum.

Hayes, C. (1951). *The Ape in our House*. New York, Harper.

Heath, C. (1984). Talk and Recipiency: Sequential Organisation in Speech and Body Movement. **In** J.M. Atkinson, & J.C Heritage (Eds). *Structures of Social Action* (pp 247-265), Cambridge: Cambridge University Press.

Heath, C., & Luff, P. (1991). Disembodied Conduct: Communication through Video in a Multimedia Environment. *Proceedings of CHI '91 Human Factors in Computing Systems* (pp 99-103), New York: ACM.

Heath, C., & Luff, P. (1992). Media Space and Communicative Asymmetries: Preliminary Observations of Video-mediated Interaction. *Human Computer Interaction*, 7, 315-346.

Hemsley, G.D., & Doob, A.N. (1978). The Effect of Looking Behavior on Perceptions of a Communicator's Credibility. *Journal of Applied Social Psychology*, 8, 136-144.

Heritage, J.C. & Watson, D.R. (1979). Formulations as Conversational Objects. **In** G. Psathas (Ed). *Everyday Language: Studies in Ethnomethodology*, New York: Irvington Publishers Inc.

Hollier, M.P., & Voelcker, R. (1997). Towards a Multimodal Perceptual Model. *BT Technological Journal*, 15(4), 162-316.

Hutchins, E., & Klausen, T. (1991). Distributed Cognition in an Airline Cockpit. **In** Y. Engerstrom, & D. Middleton (Eds). *Cognition in Communication at Work*. Cambridge: Cambridge University Press.

Inoue, M., Yoroizawa, I., & Okubo, S. (1984). **In** *Proceedings International Teleconference Symposium* (pgs 66-73). International Telecommunications Satellite Organisation. Reprinted in D. Bodson & R. Schaphorst (1989). *Teleconferencing* (pp 211-218), New York: IEEE.

Isaacs, E.A., & Tang, J.C. (1993). What Video Can and Can't Do For Collaboration: A Case Study. *Proceedings of ACM Multimedia 93* (pp 199-206), New York: ACM Press.

Ishii, H., & Kobayahsi, M. (1992). Clearboard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact. **In** *Proceedings of ACM CHI '92 Conference on Human Factors in Computing Systems* (pp 525-532), New York, ACM Press.

Jaffe, J., & Feldstein, S. (1970). *Rhythms of Dialogue.* New York: Academic Press

Jarvella, R.J. & Klein, W. (1982). *Speech, Place and Action: Studies in Deixis and Related Topics.* Chichester: Wiley.

Jefferson, G. (1972). Side Sequences. **In** Sudnow, D. (Ed). *Studies in Social Interaction.* New York: Free Press.

Kendon, A. (1967). Some Functions of Gaze Direction in Social Interaction. *Acta Psychologica,* 26, 22-63.

Kendon, A. (1967a) *Gesticulation and Speech: Two Aspects of the Process of Utterance.* **In** M.A. Key (Ed). (1980) *The Relationship of Verbal and Nonverbal Communication,* The Hague, Paris.

Kendon, A., & Ferber, A. (1973). A Description of Some Human Greetings. **In** R. Michael & J. Crook (Eds). *Comparative Ecology and Behaviour of Primates* (pp 591-668), London: Academic Press.

Kenyon, N.D., White, T.A., & Reid, G.M. (1985). Behavioral and User Needs for Teleconferencing. *Proceedings of the IEEE*, 73 (pp 689-699). Reprinted in D.Bodson & R.Schaphorst (1989). *Teleconferencing* (pp 219-229), New York: IEEE.

King, R.C., & Xia, W. (1999). *Media Appropriateness: Effects of Experience on Communication Media Choice.* **In** K. Kendall (Ed). *Emerging Information Technologies: Improving Decisions, Co-operation, and Infrastructure* (pp 143-175), Thousand Oaks, Sage Publications, CA.

Kitawaki, N., Kurita, T., & Itoh, K. (1991). Effects of Delay on Speech Quality. *NTT Review*, 3, 88-94.

Kleck, R., & Nuessle, W. (1968). Congruence Between the Indicative and Communicative Functions of Eye Contact in Interpersonal Relations. *British Journal of Social and Clinical Psychology*, 7, 241-246.

Kleinke, C.L. (1986). Gaze and Eye Contact: A Research Review. *Psychological Bulletin*, 100(1), 78-100.

Kleinke, C.L., & Pohlen, P.D. (1971). Affective and Emotional Responses as a Function of Other Person's Gaze and Cooperativeness in a Two-Person Game. *Journal of Personality and Social Psychology*, 17, 308-313.

Kleinke, C.L., Staneski, R.A., & Berger, D.E. (1975). Evaluation of an Interviewer as a Function of Interviewer Gaze, Reinforcement of Subject Gaze, and Interviewer Attractiveness. *Journal of Personality and Social Psychology*, 6, 7-12.

Kowtko, J.C., Isard, S.D., & Doherty, G.M. (1991). *Conversational Games Within Dialogue*. Research paper, Human Communication Research Centre. *Journal of the Acoustic Society of America*, 41, 286-292.

Krauss, R.M. & Bricker, P.D. (1967). Effects of Transmission Delay and Access Delay on the Efficiency of Verbal Communication.

Kraut, R.E. & Higgins, E.T. (1984). Communication and Social Action. **In** R.S. Wyer, & T.K. Srull (Eds). *Handbook of Social Cognition* (pp 87-127), London: Lawrence Erlbaum.

Kraut, R.E., Lewis, S.H. & Swezey, L.W. (1982). Listener Responsiveness and the Coordination of Conversation. *Journal of Personality and Social Psychology*, 43, 718-731.

Kraut, R.E., & Poe, D. (1980). Behavioral Roots of Person Perception: The Deception Judgements of Customs Inspectors and Laymen. *Journal of Personality and Social Psychology*, 39, 784-798.

Kraut, R., Fish, R., Root, B., & Chalfonte, B. (1993). Informal Communication in Organisations. **In** R. Baecker (Ed.). *Groupware and Computer Supported Cooperative Work* (pp 287-314), San Mateo, CA: Morgan Kaufman.

Leidner, D.E., & Jarvenpaa, S.L. (1993). The Information Age Confronts Education: Case Studies On Electronic Classrooms. *Information Systems Research*, 4, 24-54.

Lennard, H.L., & Bernstein, A. (1960). *The Anatomy of Psychotherapy*. Columbia U.P.

Levine, M.H., & Sutton-Smith, B. (1973). Effects of Age, Sex and Task on Visual Behaviour During Dyadic Interaction. *Developmental Psychology*, 9, 400-405.

Levine, J.M., Resnick, L.B., and Higgins, E.T. (1993). Social Foundations of Cognition, *Ann.Rev. Psychology*, 44, 585-603.

Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Lochman, J.E., & Allen, G. (1981). Nonverbal Communication of Couples in Conflict. *Journal of Research in Personality*, 15, 253-269.

Lyons, J. (1968). *Introduction to Theoretical Linguistics*. London, Cambridge University Press.

Machida (1986). Teacher Accuracy in Decoding Nonverbal Indicants of Comprehension and Noncomprehension in Anglo- and Mexican- American Children. *Journal of Educational Psychology*, 6, 454-464.

McCarthy, J. & Monk, A.F. (1994). Channels, Conversation, Cooperation and Relevance: All You Wanted to Know About Communication but Were Afraid to Ask. *Cooperative Computing*, 1, 35-60.

McGrath, J.E. (1984). *Groups, Interaction and Performance*. Prentice Hall, Englewood Cliffs, N.J.

McGurk , H., & MacDonald, J. (1976). Hearing Lips and Seeing Voices. *Nature*, 264, 126-130.

McNeill, D. (1985). So You Think Gestures are Nonverbal? *Psychological Review*, 92(3), 350-371.

McNeill, D., & Levy, E. (1982). *Conceptual Representations in Language Activity and Gesture.* **In** R. Jarvella & W. Klein (Eds.). *Speech, Place, and Action: Studies in Deixis and Related Topics* (pp 271-295), Chichester, England: Wiley.

McNeil, D. (1992). *Hand and Mind: What Gestures Reveal About Thought.* Chicago: University of Chicago Press.

Mehrabrian, A., & Williams, M. (1969). Nonverbal Concomitants of Perceived and Intended Persuasiveness. *Journal of Personality and Social Psychology*, 13, 37-58.

Monk, A.F. and Gale, C. (2002). A Look is Worth a Thousand Words: Full Gaze Awareness in Video-Mediated Conversation. *Discourse Processes*, 33(3), 257-278.

Monk, A., McCarthy, J., Watts, L., & Daly-Jones, O. (1996). *Measures of Process.* **In** M.Macleod and D.Murray (Eds). *Evaluation for CSCW* (pp 125-139), Berlin: Springer Verlag.

Morley, I.E. & Stephenson, G.M. (1969). Interpersonal and Interparty Exchange: A Laboratory Simulation of an Industrial Negotiation at the Plant Level. *British Journal of Psychology*, 60, 543-545.

Nardi, B., Schwartz, H., Kuchinsky, A., Leichner, R., Whittaker, S., & Sclabassi, R. (1993). Turning Away From Talking Heads: An Analysis of 'Video-as-Data.'

Proceedings of CHI '93 Human Factors in Computing Systems (pp 327-334), New York: ACM.

Nardi, B., Kuchinsky, A., Whittaker, S., Leichner, R., & Schwartz, H. (1996). 'Video-as-Data': Technical and Social Aspects of a Collaborative Multimedia Application. *Computer Supported Co-operative Work*, 4, 73-400.

Nichols, K.A., & Champness, B.G. (1971). Eye Gaze and the GSR. *Journal of Experimental Social Psychology*, 7, 623-626.

Ochsman, R.B., & Chapanis, A. (1974). The Effects of 10 Communication Modes on the Behaviour of Teams During Co-operative Problem-Solving. *International Journal of Man-Machine Studies*, 6, 579-619.

O'Conaill, B., Whittaker, S., & Wilbur, S. (1993). Conversations over Video Conferences: An Evaluation of the Spoken Aspects of Video Mediated Communication, *Human-Computer Interaction*, 8, 389-428.

Olson, G.M., Olson, J.S., Carter, M., & Storrosten, M. (1992). Small Group Design Meetings: An Analysis of Collaboration. *Human-Computer Interaction*, 7, 347-374.

Olson, J.S., Olson, G.M., & Meader, D.K. (1995). What Mix of Video and Audio is Useful for Remote Real-Time Work. *Proceedings of the Conference on Human Factors in Computing Systems* (pp 33-45), Denver, CO: Academic Press.

Olson, J.S., Olson, G.M., & Meader, D. (1997*). Face to Face Group Work Compared to Remote Group Work With and Without Video.* **In** K.E. Finn, A.J. Sellen & S.B. Wilbur (Eds). *Video Mediated Communication.* Lawrence Erlbaum Associates, New Jersey.

O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G., & Bruce, V. (1996). Comparison of Face-to-Face and Video-Mediated Interaction. *Interacting with Computers,* 8(2), 177-192.

Otteson, J.D., & Otteson, C.R. (1980). Effect of Teacher's Gaze on Children's Story Recall. *Perceptual and Motor Skills,* 50, 35-42.

Paechter, M., Schweizer, K., & Weidenmann, B. (2000). Parasocial Relations of Students to a Tutor in Internet-Based Instruction. *Medienpsychologie: Zeitschrift fuer Individual and Massenkommunikation,* 12(4), 242-259.

Parry, J. (1970). *The Psychology of Human Communication.* University of London Press, London.

Parsons, C.J., & Hughes, S.J. (1970). *Written Communication for Business Students.*

Edward Arnold, London.

Patterson, M.L. (1976). An Arousal Model of Intrerpersonal Intimacy. *Psychological Review*, 83, 235-245.

Pausch, R. (1993). Three Views of Virtual Reality. *IEEE Computer*, Feb, 79-80.

Peery, J.C., & Crane, P.M. (1980). Personal Space Regulation: Approach-Withdrawal-Approach Proxemic Behaviour during Adult-Preschooler Interaction at Close Range. *Journal of Psychology*, 106, 63-75.

Pellegrini, R.J., Hicks, R.A., & Gordon, L. (1970). The Effects of an Approval-Seeking Induction on Eye-Contact in Dyads. *British Journal of Social and Clinical Psychology*, 9, 373-374.

Peterson, C.L., Danner, F.W., & Flavell, J.H. (1972). Developmental Changes in Children's Response to Three Indications of Communication Failure. *Child Development*, 43, 1463-1468.

Pipe, M.E.& Henaghan, M. (1996). Accommodating Children's Testimony: Legal Reforms in New Zealand. *Criminal Justice and Behavior,* 23, 377-401.

Porter, L.W., Argyle, M., & Salter, V. (1970). What is Signalled by Proximity? *Perceptual and Motor Skills*, 30, 39-42.

Prussog, A., Muhlbach, L., & Bocker, M. (1994). Telepresence in Videocommunications. *Proceedings of the Human Factors of Ergonomics Society 38th Annual Meeting*, 1, (pp 180-184), Santa Monica, CA: Human Factors and Ergonomics Society.

Reid, A.A.L. (1970). Electronic Person-Person Communications. Communication studies group paper No. P/70244/RD. British Post Office.

Rice, R.E. (1993). Media Appropriateness: Using Social Presence Theory to Compare Traditional and New Organizational Media. *Human Communication*, Res, 19, 451-484.

Rime, B. (1982). The Elimination of Visible Behaviour from Social Interactions; Effects on Verbal, Non-verbal and Interpersonal Variables. *European Journal of Social Psychology*, 12, 113-129.

Robinson, A. (1995). Delivering Continuing Professional Development by Cross-Campus Videoconferencing. *International Journal of University Adult Education*, XXXIV (2): 33-43, 1995.

Robinson, A., & Jones, S. (1985). *Teaching by Teleconferencing*. University of Ulster.

Rogers, E.M. (1986). *Communication Technology: The New Media In Society.* The Free Press, New York, Ny.

Rosenfeld, H.M. (1987). *Conversational Control Functions on Non-verbal Behaviour.* **In** A.Seigman & S.Feldstein (Eds), *Non-verbal Behaviour and Communication* (pp 563-601), Hillsdale, NJ: Erlbaum.

Russo, N.F. (1975). Eye Contact, Interpersonal Distance, and the Equilibrium Theory. *Journal of Personality and Social Psychology*, 31, 497-502.

Rutter, D. (1984). *Looking and Seeing: The Role of Visual Communication in Social Interaction.* New York: Wiley.

Rutter, D. (1987). *Communicating by Telephone.* Pergamon.

Rutter, D., & Stephenson, G. (1977). The Role of Visual Communication in Synchronizing Conversation. *European Journal of Social Psychology,* 20, 41-52.

Rutter, D.R., Stephenson, G.M. & Dewey, M.E. (1981). Visual Communication and the Content and Style of Conversation. *British Journal of Social Psychology*, 20, 41-52.

Sacks, H., Schegloff, E.A. & Jefferson, G. (1974). A Simplest Systematics for the Organisation of Turn Taking in Conversation. *Language*, 50, 696-735.

Sasse, M.A. (1996). Noddy's Guide to Internet Videoconferencing. *Interfaces*, 31, 6-9.

Schegloff, E.A. (1968). Sequencing in Conversational Openings. *American Anthropologist*, 70, 1075-1095.

Schegloff, E.A. (1987). Recycled Turn Beginnings: A Precise Repair Mechanism in Conversation's Turn-Taking Organisation. **In** G. Button & J.R.E. Lee (Eds). *Talk and Social Organisation*, Clevedon: Multilingual Matters.

Schegloff, E.A. (1991). Conversation Analysis and Socially Shared Cognition. **In** L.B. Resnick, J.M. Levine & S.D. Teasley (Eds). *Perspectives on Socially Shared Cognition*, (pp 150-171). Washington: American Psychological Association.

Schegloff, E.A., & Sacks, H. (1973). Opening up Closings. *Semiotica*, 8, 289-327.

Scherer, K.R. (1986). Vocal Affect Expression: A Review and Model for Further Research. *Psychological Bulletin*, 99, 143-165.

Schiffrin, D. (1974). Handwork as Ceremony: The Case of the Handshake. *Semiotica*, 12, 189-202.

Schweizer, K., Paechter, M., & Weidenmann, B. (2001). A Field Study on Distance Education and Communication: Experiences of a Virtual Tutor. *Journal of Computer Mediated Communication*, 6(2), np.

Sellen, A.J. (1992). *Speech Patterns in Video-Mediated Conversations. In CHI'92 Conference on Human Factors in Computing Systems*, Monterey. California: ACM.

Sellen, A. (1995). Remote Conversations: The Effects of Mediating Talk with Technology. *Human-Computer Interaction*, 10(4), 401-444.

Sellen, A. (1997). *Assessing Video Mediated Conduct: A Discussion of Different Analytic Approaches.* **In** K.E. Finn, A.J. Sellen & S.B. Wilbur (Eds). *Video Mediated Communication.* Lawrence Erlbaum Associates, New Jersey.

Short, J.A. (1974). Effects of Medium of Communication on Experimental Negotiation. *Human Relations*, 27(3), 225-234.

Short, J., Williams, E., & Christie, B. (1976). *The Social Psychology of Telecommunications.* London, Wiley.

Siegel, J., Dubovsky, V., Kiesler, S., and Mcguire, T.W. (1983). Group Processes in Computer Mediated Communication, *Organic Behaviour and Human Processes*, 37, 157-187.

Smith, R., O'Shea, T., O'Malley, C., Scanlon, E., & Taylor, J. (1991). *Preliminary Experiments with a Distributed, Multi-Media, Problem-Solving Environment.* **In** J.Bowers and S. Benford (Eds), Studies in Computer-Supported Co-operative Work: Theory, Practice and Design (pp 31-48), Amsterdam: Elsevier.

Smith, H.R., & Nagel, T.S. (1972). *Instructional Media in the Learning Process.* Columbus, Merrill.

Sperber, D. & Wilson, D. (1986). *Relevance: Communication and Cognition.* Oxford: Blackwell.

Sproull, L.S., and Kiesler, S. (1986). Reducing Social Context Cues: Electronic Mail in Organisational Communication, *Management Science,* 32, 1492-1512.

Straub, D., & Karahanna, E. (1998). Knowledge Worker Communications and Recipient Availability: Toward a Task Closure Explanation of Media Choice. *Organisation Science,* 9(2), 160-175.

Strauss, S., & McGrath, J. (1994). Does the Medium Matter: The Interaction of Task and Technology on Group Performance and Member Reactions. *Journal of Applied Psychology,* 79, 87-97.

Strom, D. (1997). Videoconferencing in Focus. *Internet World Magazine*, 8(9), Mecklermedia Corporation.

Strom, J.C., & Buck, R.W. (1979). Staring and Participants' Sex: Physiological and Subjective Reactive Responses. *Personality and Social Psychology Bulletin*, 5, 114-117.

Suchman, L. & Wynn, E. (1984). Procedures and Problems in the Office. *Office Technology and People*, 2, 133-154.

Sullivan, A.M., Andrews, E.A., Hollinghurst, F., Maddigan, R., & Noseworthy, C.M. (1976). The Relative Effectiveness of Instructional Television. *Interchange*, 7(1), 46-51.

Summerfield, Q. (1992). Lipreading and Audiovisual Speech Perception. *Philosophical Transactions of the Royal Society of London*, B335, 71-78.

Sutherland, S. (1995). *The Macmillan Dictionary of Psychology Second Edition*. Antony Rowe Ltd., Chippenham, Wiltshire.

Tang, J.C. (1991). Findings From Observational Studies of Collaborative Work. *International Journal of Man-Machine Studies*, 34, 143-160.

Tang, J.C., & Isaacs, E.A. (1992). *Why Do Users Like Video? Studies of Multimedia-Supported Collaboration* (SMLI Tech. Rep. No. 92-5). Mountain View, CA: SUN Microsystem Laboratories.

Tang, J., Isaacs, E., & Rua, M. (1994). Supporting Distributed Groups with a Montage of Lightweight Interactions. *Proceedings of Conference on Computer Supported Co-operative Work* (pp 23-34). New York: ACM.

Tankard, J.W. (1970). Effects of Eye Position on Person Perception. *Perceptual and Motor Skills*, 31, 883-893.

Taylor, I. (1976). *Introduction to Psycholinguistics*. Holt, Rineheart and Winston, New York.

Temple, B. (1999). *Entrepreneurship: Report on Multi-National, Multi-Disciplinary Team Projects.* In W. Alexander., C. Higgins & N. Mogey (Eds) *Videoconferencing in Teaching and Learning: Case Studies*. LTDI and Talisman, Institute for Computer Based Learning, Heriot-Watt University, Edinburgh (1999).

Thayer,S. (1969). The Effect of Interpersonal Looking Duration on Dominance Judgements. *Journal of Social Psychology*, 79, 285-286.

Thayer, S., & Schiff, W. (1974). Observer Judgement of Social Interaction: Eye Contact and Relationship Inferences. *Journal of Personality and Social Psychology*, 30, 110-114.

Tiffin, J.W. & Rajasingham, L. (1995). *In Search of the Virtual Class. Education in an Information Society.* Chapter 6: Telelearning in Cyberspace.

Trent, J.W., & Cohen, A.W. (1973). Research on teaching in higher education. In Travers, R.M.W (Ed), *Second hand book of research on teaching.* Chicago: Rand McNally.

Valacich, J.S., Mennecke, B.E., Wachter, R.M., & Wheeler, B.C. (1994). Extensions to Media Richness Theory: A Test of the Task-Media Fit Hypothesis, *Proc. Twenty-Seventh Hawaiian International Conf. On System Sciences, Maui, Hawaii,* 11-20.

Velichkovsky, B.M. (1995). Communicating Attention: Gaze Position Transfer in Cooperative Problem Solving. *Pragmatics and Cognition,* 3, 199-222.

Vertegaal, R. (1999). The GAZE Groupware System: Mediating Joint Attention in Multiparty Communication and Collaboration. **In** M.G. Williams, M.W. Altom, K. Ehrlich & W. Newman (Eds). *CHI '99,* Pittsburgh, P.A. (pp 294-301), New York: ACM Press.

Videotalk: Retrieved on February 17, 1999, from the World Wide Web: http://www.videotalk.co.uk/home.html. © 1998 Wincroft Inc.

Walker, M. (1993). *Information Redundancy in Dialogue.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Walker, M., & Whittaker, S. (1990). Mixed Initiative in Dialogue. **In** *proceedings of 28th Annual Meeting of the Conference on Computational Linguistics* (pp 70-78), Morristown, NJ: ACM.

Watson, A., & Sasse, M.A. (1996). Evaluating Audio and Video Quality in Low Cost Multimedia Conferencing Systems. *Interacting with Computers,* 8(3), 255-275.

Webster, J., & Hackley, P. (1996). *Learners' Reactions To Distance Telelearning.* Paper presented at the annual meeting of the *Academy of Management,* August, Cincinnati, Ohio, 1996.

Weiner, M., Shilkret, R., & Devoe, S. (1980). *Acquisition of Communication Competence: Is Language Enough?* **In** M.R Key (Ed). *The relationship between verbal and non-verbal communication* (pp 275-293). The Hague, The Netherlands: Mouton.

Whitney, L., & Cook, A. (1990). *The Use of Closed-Circuit Television in New Zealand Courts: The First Trials.* New Zealand Department of Justice.

Whittaker, S. (1995). Video as a Technology for Interpersonal Communication: A New Perspective. *IS&T SPIE Symposium on electronic imaging science and technology*, 2417, 294-304.

Whittaker, S. (1996). Rethinking Video as a Technology for Interpersonal Communication: Theory and Design Implications. *International Journal of Human Computer Studies*, 42, 501-529.

Whittaker, S., Brennan, S.E. & Clark, H.H. (1991). *Coordinating Activity: An Analysis of Interaction in Computer-Supported Cooperative Work.* **In** *Proceedings of ACM CHI '91 Conference on Human Factors in Computing Systems* (pp 361-367), New York: ACM Press.

Whittaker, S., Frolich, D., & Daly-Jones, O. (1994). Informal Workplace Communication: What is it Like and How Might We Support it? **In** *Proceedings of CHI '94 Human Factors in Computing Systems* (pp 130-137), New York: ACM.

Whittaker, S., Geelhoed, E., & Robinson, E. (1993). Shared Workspaces: How Do They Work and When Are They Useful? *International Journal of Man-Machine Studies,* 39, 813-842.

Whittaker, S., & O'Conaill, B. (1997). *The Role of Vision in Face to Face and Mediated Communication.* **In** K.E. Finn, A.J. Sellen & S.B. Wilbur (Eds). *Video Mediated Communication.* Lawrence Erlbaum Associates, New Jersey.

Williams, E. (1975). Coalition Formation over Telecommunications Media. *European Journal of Social Psychology*, 5, 503-507.

Williams, E. (1977). Experimental Comparisons of Face-to-Face and Mediated Communication: A Review, *Psychological Bulletin*, 84(5), 963-976.

Yerkes, R.M, & Dodson, J.D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.

Yngve, V.H. (1970). On Getting a Word in Edgewise. In *Papers from the sixth Regional Meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society.

Zuckerman, M., & Driver, R.E. (1985). *Telling Lies: Verbal and Nonverbal Correlates of Deception.* **In** A.W. Seigman & S. Fledstein (Eds). *Multichannel Integrations of Nonverbal Behaviours.* Hillsdale, NJ, Erlbaum.

Appendices

Appendix 1: Interview with Bill Howell and Donald West of 'Worldwide Videoconferencing.'

CF – "So do you think in some ways then erm, when people use videoconferencing erm, the fact that they are communicating via a screen erm, acts as some kind of barrier so that people don't you know don't treat the other person on the other side of the link as being real in some ways?"

BH – "Erm well I think that erm yes, yes in some ways. I think the problem is here that the first impression anybody gets of videoconferencing it does exactly that. Erm as, as you become an experienced videoconferencing user you, and this is something that I also have found in the years and in the experience that I have got, is that most of the content of a video conference happens to be audio, believe it or not, about 85% from our estimations: in a sense that you may convey a picture and as long as you are real and I sort of try and make a personal or have some sort of personal relationship with you to judge whether you are someone who is genuine maybe someone who I believe in erm is not dissimilar from the, from the initial erm personal contact you make by meeting people and providing you get to a level which is good enough erm in terms of the visual and the audio quality, you mentioned audio synch erm video-audio synching in terms of lip synch and audio, providing you get to a level such as that that is acceptable then suddenly the audio content or the or what I call the intelligent content takes over, because you know the picture is unintelligent, a picture is unintelligent, but the content of audio is actually erm the stuff that decisions are made on. What happens is, as you become more experienced as a videoconference user then the environment you are in becomes fundamentally more important and particularly the audio content because if you can have a conference and feel as though you know your emphasis perhaps in, in linguistic terms is accepted by the far end it really doesn't matter if you are laughing or crying. The, the content of the seriousness of the audio content is quite important and what I find is that the better the audio and the better the erm video quality to a certain plateau then people can become experienced and, and cheerful if you like or erm, consistent users of videoconferencing and I think therein lies one of the big problems you're going to get because, because the, in your research. Because there is a benchmark, T.V, where audio and video is brilliant you know in comparison, if you don't get to that benchmark, if you don't get within to, close to that benchmark with videoconferencing, what ends up happening is people don't use it and therefore the impression of it is that it isn't a useful business tool because they haven't reached a certain level of capability and that is almost exclusively in my opinion down to you know the use of the product as well as the quality of the product in the first place."

CF – "You mentioned that benchmark, erm that T.V being the benchmark. Do you think videoconferencing will ever reach that benchmark, as far as quality….."

BH – "…yeah I mean, yeah let me, I mean we are on 128 conference here, I'll actually dial in 384 perhaps a little later on and what I want to do give you an example of the difference between 128 and 384. It's not really detectable in a sense, it clearly is in terms of a long videoconference you'll get to find that the high quality is so close to what I would call broadcast video in a business environment; what I mean by that well in a business environment you don't see formula one cars going past at 200 miles an hour, so the impression of speed which is the one thing that really does ruin videoconferencing images is

much lower. We are just looking face to face here so the frame rates that we are exchanging in other words the perception of jerkiness if you like of the picture is very small, because there's not a lot of movement. So if you actually compare it to a T.V image the image that we are actually erm exchanging is good enough for a videoconference; good enough for a business meeting because it isn't particularly movement orientated. Now there again lies, if you like, a technique of videoconferencing. If I actually want to get rid of the fact that you are moving your hands around which could deteriorate the picture I could just zoom in on your face and therefore techniques of using videoconferencing can actually improve, if you like, the perception of the video quality simply by using a few tricks of the trade perhaps."

CF – "So in zooming in on the face obviously you get less movement so it would be less jerky erm but in doing that you cut out some of the body signals. Do you think these body signals are important in a video communication erm scenario?"

B H - "Again I think you need to put this into some perspective in that most videoconferencing is used for intra-company meetings; in other words it is unusual for me, for example, to meet you in this sort of environment and videoconference. Erm in general terms I know what *Donald* looks like; I sort of know what he's going to do if I ask him a difficult question and so in a way it's not a fact that the body language is there it's a fact that the body language is sort of expected and I guess in a sort of, in a, in a. Videoconferencing is very good. Let's go back one stage. Videoconferencing is very good in an iterative environment, so if I talk to *Donald* everyday as in London to Slough as normal, or maybe London to Edinburgh, erm then my expectation of what he is going to do is only confirmed by what he does..."

CF – "....that's because you have met him in the flesh?"

BH –"... Yes if you like. It's the negative aspects. If he doesn't do it I want to know why, so in a videoconference you actually get, and erm this another, and this another perhaps erm point, point to make, because I am looking at a very focused position you know, I am looking at you in a room on a clearly small flat T.V and therefore my focus is completely on what you are doing at that end and it's very difficult for me to be distracted by things at this end, but you know that potentially will take away, if you like, the focus of the meeting. So you find videoconferencing actually increases the meeting focus because you are actually looking, to coin a phrase, at you know a sort of porthole, it's the porthole effect you know, you've only got one thing to look at you know and it's the sea and you are in a boat and so therefore you look at it a lot and the sort of impression is that you focus on that, on that image. So in effect I don't think it's the lack of body language that potentially causes a problem because you can zoom cameras in and out and move them around. What does happen I think is that because videoconferencing does focus, for example, on the person rather than the room it doesn't matter whether a door opens at your end, I wont notice it and therefore my focus tends to be much more on the person in the meeting irrespective you know of whether they are waving their hands about as I am sure they do in France as they are expressing with their hands but certainly in this country and international conferences elsewhere erm it's really the content as I go back to the point, it's really the content that starts to become quite important regarding the audio level and the audio quality rather than necessarily the body language, but I also think that it's the focus to the extent that body language if it's missing starts to become very obvious rather than if it's there."

CF – "If the audio quality, if the audio content is more important erm, then what advantages does V.M.C have over just using the telephone?"

BH – "Well I think this this, and this has often been asked as a question. I think the important thing is that if I wanted to bring *Donald* into this conversation, I know that he's not in the picture right now, I could simply say *Donald*, you know, what's your opinion of this and because he's been party to both the vid... you know if I'm angry he's been party to that image, well that's not necessarily obvious on an audio conference, erm but what it does do is allow many people to meet many other people all at the same time, with exactly the same content without having to have interpretation or perhaps of the of a telephone call. In other words perhaps if I was on a telephone call I would call *Donald* up and say what was the guy like, was he a nice guy, friendly, was he unfriendly? In other words I couldn't judge, if you like, the whole of the content and the whole of the participants in the meeting if it wasn't on video and that doesn't necessarily mean that you need video to do that, but I think what it is, is it is very clearly a group and a team, if you like, meeting technology or meeting business application which probably isn't reproduced anywhere else. The classic example is one of the marketing director I met many years ago, very first years in videoconferencing. He said to me the good thing about videoconferencing is you know I go to the states once every three months and I know everything they are going to do, it was an American company with a U.K subsidiary, he said I know everything they're going to go and do but by the time I get back to the U.K I have forgotten you know fifty percent of the things that were important and when I debrief my team and I talk to the marketing people you know I don't convey the same enthusiasm or the same interest that I would when I was there. You know people sold to me and I knew what was going on or what was coming out of you, but add the new product XYZ in three months time is not necessarily you know that interesting to people but when I tell them to, but what is interesting is what that product does but I don't have the time or the erm if you like the opportunity to be able disperse that information to lots of people in that environment. So he said videoconferencing was great because what I actually got was twenty people together in one go and the guy that I, but I still went to the states interestingly enough, but what I did was get this guy to talk to these people but in the same way that I had found it you know that we discussed it and then everybody got the sort of buzz and the information from a videoconference and they also respected the fact that you know a quite important person had bothered to talk to them. In that sense it sort of 'de-layers' an organisation in that it doesn't matter if I am the M.D of the business and you know you are coming to look at this you know we are all equal in 'videoconference land'. I think, I think that perhaps tends to define in some instance, because it is a technology that everybody can use, there isn't any sort of badges if you know what I mean you know we are not in the executive board room where you couldn't meet you know the director of a company and so in some ways it does "de-layer" the organisation by allowing people to meet who perhaps wouldn't have met before I mean not sort of deliberately but just because of time demands."

CF – "Perhaps you could tell me a little bit about the erm technology that you use here erm if you've got the time.

BH - "Yeah sure. Erm, the erm the businesses really started in the early eighties. The fact many years ago many years before that BT invented compression technology to be able to send pictures on cable basically. Erm but really the development of videoconferencing is clearly defined as being erm available, available to all about the mid eighties when ICN, the digital telephonist service really gave sufficient coverage for people to be able to specify it, install it and actually use it. Erm and because in the early days most people's attention was

on the lowest common denominator which is 2x64k channels or ISDN2 erm then everybody manufactured their equipment to compress as best they could, you know put lots and lots of effort into designing their product that would compress the best video quality they possibly could into that lowest common denominator and the good news in my opinion is this, is that because it is the lowest common denominator around the world then if you go to you know very many of the worlds countries then they still maintain or still have the ability to produce that lowest common denominator and for all the technologies that are out there now that might be brilliant and good like yes they deliver video but do they deliver it to erm Moscow or can they actually deliver it to the erm outback's of Australia? And the key is here the answer is generally no because the basic common denominator which is allowing this to happen in the first place is built around the telephone system and because the worlds companies and telecoms companies infrastructure is built around the digital backbone you know digital ISDN backbone then the provisions will then point just about anywhere in the world, providing it is reasonably developed or has the capability to let them in, is going to be a telephone connection and the fact that videoconferencing can use that telephone connection, all be it perhaps a slightly different 'flavour' than just a voice connection means that videoconferencing if it is going to be deployed, is going to be deployed on that lowest common denominator. So as technology increases I still don't think necessarily that the method of transmission, wide area, is gonna change, but maybe there is gonna be some changes internally to companies. Typically IP or LAN video as it's been dubbed erm is a strong candidate for that. I think you've gotta go back and ask the question 'What is videoconferencing used for primarily?' There is still a very large proportion of companies that do a, you know, a London or Slough to Edinburgh connection with reasonable audio and video quality. Potentially you know a very large proportion of it is internal and therefore having established connections such as ISDN as we've got here erm means we can dial up and connect and really the transport mechanism is fairly irrelevant to, to erm most users of videoconferencing since they are actually discussing business problems rather than necessarily technology problems. So I think there are a number of emerging technologies, all of which have a price tag but I think the key issue is here is that the, that the method of transport may change but the fundamental 'connectivity' is gonna be some sort of telephony. Erm you know ATM has been mentioned erm a lot of frame relay type solutions have been mentioned but these are really, these are really not available in more than a, you know, small proportion of locations. So if you wanted one tomorrow in your home for example you have no choice; it will have to be some sort of ISDN BT highway type connection and then you could then videoconference with anyone else that you wanted to anywhere effectively around the world... (Market shares, business talk, I.T industry vs. telecommunications) I think there is a difference in if you like mentality between wide area telecommunications people whose sort of motto is if you like is always up-time, you know up-time all the time and compared to compared to the IT oriented people where perhaps where some of the emerging video technologies are well if it works it works but if it crashes reboot your PC, you know? That seems to be perhaps erm perhaps erm an example of where the two erm, convergence which is often used in this industry of the IT and the telecommunications industry becoming erm you know becoming one entity doesn't actually matter because the mentality of the two sorts of 'camps' if you like are very different. Telecoms tend to be very long-term in five year cycles but very you know very reliable, very simple in general terms whereas the IT industry changes every five minutes you know new technology emerges...(technology listed e.g. fibre optic lanterns etc.). The IT industry changes so often that to get a stable platform for communications is actually quite difficult and very expensive in actual fact whereas the telecommunications world is sort grumbling on at much lower change level speeds. So I do think clearly they're gonna migrate towards each other but I don't think we're clearly there

yet and I do think there is a very big concern perhaps is a better word because the LAN technology, local area network technology just does not have the just has the fundamental design philosophy of LANs but is not suitable to synchronous to synchronous video...(LAN more technical talk)."

CF- "As far as application is concerned err you mentioned earlier on that videoconferencing at the moment is mainly used in a business environment. Erm obviously it's also used in academic environments as well, for distance learning and things like that but as far as the future's concerned what applications do you think videoconferencing will be used for, or what will they best be applied to?"

BH - "I think that the education market and particularly medical erm remote expertise is a generic term but I think that's very important here. I think the price tag has always been definitive in terms of where it's used most of all, you know academically other than a few locations it's actually you know it's quite an expenditure to make or actually acquire what apparently cannot be necessarily used in, in a whole bundle of applications but distance learning is in my opinion is very important and very strong in places like Australia you know you can attend university courses on video simply because of the distances involved erm MIT in the US has the same sort of thing, and yes I think those technologies are actually extremely useful in those environments and will only in my opinion continue to grow. I think the only problem is all the end points right now other than desktop stuff which perhaps we can talk a little bit about separately, the end points are expensive and it doesn't bode well in my opinion for erm you know for five thousand pounds worth erm the system that you've got there thirty grand's worth you know that you're not gonna get that deployed erm universally to actually generate any benefits from perhaps erm erm a technology such as erm erm or in the application in perhaps distance learning unless the prices come down I think that makes a big difference. Tele-medicine I think is an interesting one because of the deregulation, if you like, and the, and the use of private medicine particularly in this country where frankly again the charge for erm you know medical help or operations on the, in the private sector erm means that people spend an awful lot of money on it and I can see that becoming very important particularly as the video quality improve but there are some legislative problems that occur there of course particularly from America which is litigation mad erm in the fact that the video quality or video resolution isn't sufficient erm to do for example things like the analysis of an X-ray and if there is erm there are diagnoses or misdiagnoses made because of the use of the technology then people are very concerned that they can be sued, particularly in America because someone has made it got it wrong and because they can't see the quality of the X-ray so there are some I think barriers for those sort of applications. Erm I see applications in kiosks, Global as a company and the previous company before we've done funerals, we've done weddings, erm we've done all sorts of applications in regards to videoconferencing. We sort of point to the leisure, I wouldn't call a funeral a leisure industry, but point towards the sort leisure industry...(golf, David Ledbetter)...those are the applications which are very much undervalued right now erm which could be useful...(Sky sports match reports)...The leisure industry is an interesting one and I do see a lot more money being spent in that industry you know as disposable income starts to mean, as we grow richer, as we are told we will grow richer in this country I think that could be a very interesting form of visual communications. I'm not convinced in terms of technology used in banking at this stage although there has been a lot of trials, we've been involved in a lot of trials in erm kiosk work, people actually being able to order mortgages and stuff. Erm it seems to me that the risk in getting that wrong seems quite a lot higher, you'd much rather sort of pin someone down face to face and say 'how much is this going to cost me' rather than

tap in perhaps to erm to a computer you know talking to someone on video and tap in and they get the numbers wrong. Erm it seems to me it seems to me it's a little more risky in terms of their in terms of their application and I certainly know from experience that the take-up of kiosks in particular a trial was done by a financial institution which shall remain nameless, they put a kiosk in each of their, in each of their branch offices in the *Diss* area and just about everybody went to a person at a desk. The predominant reason for that is that a very large proportion of the people that were going into them were old people with savings you know people taking a few pounds out erm from savings accounts and predominately the people who were interested in doing it were young. It would seem there's an obvious connection. But, the young are not always that interested in saving money, I know it seems a general sort of statement but in effect what they were saying was that people didn't trust this new technology and would much rather talk to a real person even though they had to queue up because it was perceived as being quite difficult to interface to. I don't think necessarily that it was the video component that was difficult it's actually the buttons you had to press to make it happen. So I think there's a lot of perception in that sort of area where erm it may be, it may be that in terms of a self starting application difficult to sell video as a component."

CF- "Okay erm can I just ask you a quick question about desktop videoconferencing?

PS- "Yeah".

CF- "Erm as far as desktop videoconferencing is concerned it's obviously not, the quality isn't as good as specialist conferencing centres like yourselves erm do you know, as far as the technology is concerned though do you know whether that is going to improve much in the future at all?"

BH - "Well I sort of erm I'd also take a step back in this. I mean my experience with desktop videoconferencing is that it can be as good as if not better in some instances than conference room video. I think the problem we've got here is that desktop video fundamentally doesn't work in my opinion because I have a desk out there with a laptop on it. It would be as easy for me to talk to you on that laptop technologically speaking and with the same sort of quality providing I throw enough bandwidth at it and a few other bits and pieces. Err although it might be a more expensive call it would still give us the sort of capability we've got, the problem is I can't sit out there and conference while there are six or seven other people in an open office, so we would be interrupted. And if you think about that it would give me a very, and I have had this opinion for a long time. It makes it very difficult for videoconferencing erm to become a fundamental component of an office-based environment and therefore I don't think desktops fundamentally would work. Desktop videoconferencing fundamentally doesn't work because the last thing you want if you've got a meeting perhaps is for, is for the phone to ring on the desk next door whilst you are trying to talk to a colleague or maybe even a team of colleagues talking about critical electronic design or test results or something when they actually say 'sorry we can't actually hear you' so it doesn't matter whether it's a video you know video connection in there or not. Just the environment of desktop videoconferencing makes it quite hard to use. Now that doesn't mean to say you can't use a PC in videoconferencing but I still think you know that predominately and if you ask erm *Donald* on this, he's had some experience as well, people use rooms or small offices almost exclusively for videoconferencing because, one that they want more people to be involved in the meeting than just one and that sort of obviates the indenture of phone call, requires you to have some sort of videoconference at least a videoconference. And two you're actually talking about something that requires a little bit of concentration and a little bit of decision

making because there's no point in having a meeting in my opinion erm on videoconferencing unless there is a decision making process involved in, in erm the conference. So in other words a desktop thing is probably erm you can E-mail somebody, or you can send a spreadsheet or discuss data conference and that's quite easy on desktop because it generally doesn't require audio input. Where an audio input is required you can't do it in a open office and that seems to be to me a very sort of simplistic thing that everybody seems to miss...(bought video systems)...The point being that I think it would appear to be a great technology but in fact in practice it adds no more than you'll probably be able to do on the telephone and I think that is also a result of the fact that you tend to be one to one on PCs. You know that it's rare that you sit around a PC all day isn't it I mean its rare that you don't have a PC to yourself so all of your PC interaction tends to be at least one from your end to somebody else and typically one to one. You e-mail somebody although it goes through a LAN it's just a one to one connection and PCs are sort of designed to be personal you know that's what they are personal computers but it doesn't bode well if you've got a video system there and two people sitting around it though it can be done because you can't you know, although it's a cheap video system, you can't actually use most of the P.C facilities or both of you can't use those facilities. So I think desktop video is one of those things that seems to be a bit like the *laser* you know a brilliant solution to something that doesn't have a problem and I think that will probably continue to be the case until such time as perhaps you know headsets maybe have been tried before but the people don't like to sit around all day with a headset on, on the basis that they are going to have a conference for an hour. And it's not a critical component of their everyday lives. So if it's a problem for me to get up and walk five feet to come into a meeting room I'm gonna do that every time rather erm than do it on a PC."

DW- "There's another angle to this"

BH- "Sorry?"

DW- "There's another angle to the desktop systems where people keep saying people are going to have them at home because you are sitting in your study at home communicating with people at work and you've got no distractions at home you're sitting in your study with your PC system and you're talking to colleagues at work. That's where I think it'll work. The problem with that is ISDN at the moment for the average man at home is too expensive to install and so that negates you know we're talking about home workers and telly workers so that negates that, that thing. But that's where I see desktop conferencing coming into its own if it does where people have their own privacy and they don't want a big system like this sitting in their study."

CF- "So when the prices come down at home"

DW- "Well the prices aren't that expensive anyway it's the ISDN connections that put people off like B.T are charging £200 to install it and £75 a quarter. You know the call charges locally aren't that much but when the cable companies start piping ISDN2 into your homes then I think people will start, teleworkers might start looking at that, but you know we've got one there, that's a desktop system and if you can imagine us being in an office with lots of other noise you've got to concentrate that much harder to hear someone that you are looking at and it negates the whole reason that you are doing it, you know I think that's what's Phil trying to put across. I mean that you know it's it's a bit of a novel thing I think and if people do use it with companies it's between two departments, you know it's an internal, we don't

never use, we never charge people to use them, to talk to someone round the world we use the bigger systems."

BH -"Yeah I think that's a pretty good summary of Desktop, although you know to be honest if you're talking about home users, if you are talking about the price you know I don't know anybody that buys anything for five thousand pounds in the home. I don't know of many people that actually want to go buy a thousand P.C believe it or not, even though you, I think it's about 27% of the U.K populations got a P.C now and that's been driven completely by the Internet not by you know home use or you know home accounts or anything like that. But I think from that perspective you've got to perhaps consider that the addition of a thousand-pound video system you know to a thousand pound P.C is actually a substantial component of anybody's erm home, home expenditure even if it's paid for by the company...(M.D's have video's at home)..."

CF- "I'm just going to ask you a few questions about the technology again. It's concerning, erm concerning the sort of options that you have when you're videoconferencing. Erm do you have like for example voice activated switching modes, erm can you show graphics that kind of thing?"

BH – "Yeah all of those. I mean in a sense video, videoconferencing technology is about a acquisition of a video source, so if I can convert a P.C to video I can show you it, if I can convert erm a document with a document camera, Dave will probably show you one around, erm it can convert paper to a video source, yeah I can do it. There's sort of nothing that you can't do, I mean we've run television programmes over video, tapes you can produce, so in a sense any visual source, by definition a visual source is something that is recordable with some sort of camera, erm can be replayed. So, switching erm, there is a very large proportion of features and facilities in video systems that people never use frankly, and they don't use that very often because the technology erm or the environment that they are in doesn't make it usable in a very simple sense. I think people sort of would data conference a lot more if it was very easy to data conference whilst you're in a video conference, but not many people have reengineered their business practices to do that. You know a good example I know, we introduced the customers to budgeting on a multi-point video conference using an excel spreadsheet on a big screen and it was very easy to do, providing you know the client did some simple things, it was very easy for everyone to participate in the budget of this company all over the world, and as soon as they saw that it was easy to do, suddenly it became a thing that they did all the time, but without actually sort of, without trying to be too complicated about what was going on then doing one thing at a time well actually made, had a great reel of benefits and rather than if I was saying, what I'm trying to get to you is, rather than trying to perhaps look at a word document doing a press release, doing a spreadsheet on budgets at the same time and they're actually gonna go out on presentation, you know something that became complicated erm didn't work particularly well on video conference because the focus of attention, I think I'll make this point, perhaps the focus of attention of the point in question, the meeting that we are having about, you know maybe I'm telling some things that I've learnt in videoconferencing experience, I don't need to show you a chart in excel, perhaps we do that later on, when maybe the next iteration of this meeting, I'm not suggesting we have one, as the next iteration of this meeting is looking at some results in these surveys and where the graphical content might then be more interesting than just actually meeting you for the first time. I think as I say, the iteration process that most companies go through internally when they use videoconferencing becomes extremely valuable or you know when we perhaps meet the next time you say I only want to show you this chart, is this representative of the

videoconferencing industry? I say yes or no and it's that simple in a sense, but if you try to make things too complicated in videoconferencing particularly with a large audience is that when there a multiple people at each end what you tend to find is that people lose the plot. It's allright if you've got one spreadsheet, one number and someone actually saying that this is our raise margin this year and what can we do next year to **object?** it, but if you've got a you know a whole, a whole erm spreadsheet defining all sorts of things in it, not everybody's gonna be able to follow the plot in what you are trying to achieve, so you know there is a sort of limitation in a sense that the way that meetings are conducted and with that sort of information, and videoconferencing doesn't add or subtract form that it just means that more people participate and therefore you have the potential to confuse more people rather than educating them. I think those are sort of, erm in technological terms, they have yet be explored and I think that the manufacturers have yet to put things in place and I do see that changing now really that will allow simple integration erm if you have a look at the **p.k frame??** You've got millions of buttons on there but you actually only press two or three and once maybe those buttons instead of they come up with hieroglyphics or funny pictures on them, maybe one of them will say share spreadsheet or you know data conference and then pops up on this screen and it's the spread sheet you've loaded earlier and I think that those are the tools are now becoming quite, you know erm, they are much more available and much easier to use. The interface will start to look in a way which, will start the way that you as an individual does business, but I think that is another important factor in that you know we are talking camera and telephone technology here without really any integration, you know there's really very little integration in terms of a user interface that will allow me to you know graphically perhaps bring up a spreadsheet…(Unimportant)…"

CF- "So, as far as novice users are concerned erm this key pad that you have right in front of you erm all of the novice users I assume when they first start to use this technology they don't actually play around with the keypad it's something that's you know done for them. Is that correct?

BH – "You know I think it goes back to this T.V thing, I often watch home video's on the television you know sort of You've been framed or whatever, but people look really crap on video don't they?"

CF- "Yes"

BH - "I mean they do look really crap. And one of the problems is that people are really self conscious about their own image on the P.C I'm sorry on the T.V…(talks to *Donald*)…people are extremely self conscious about seeing themselves on T.V and the first thing you'll find is they do this (gesture) or they do this (gesture) you know or do the, because they actually wa.., because they are actually part of what we call erm industrial theatre. The people who really like videoconferencing are the right show offs who think they are really good on video telly and want to broadcast to everybody that they are real smart, that they look the you know erm the, they've sort of got a presence and in a way to talk about parallels because there are people who will not use videoconferencing because they don't like being on T.V… (repeat)… I think the good news is that videoconferencing tends to be one technology which once you've done it for fifteen minutes it's sort of much less intimidating than it was before, because again the purpose of what you are doing is actually on content and when people are articulate about their content they tend to forget that fact that they are on T.V. You know get someone to read a poem who is not artistic or has no interest in it and it's very difficult to do. Get someone to tell about the latest videoconferencing technology and it's very easy to do, because in their

minds there is already a picture of the information they want to convey and videoconferencing tends to bring that out very specifically so the thing to stop people videoconferncing is that hate that they are going to be on there and worse thing I see is that they do this (gesture) erm they sort of fiddle around with buttons and oh blimey I've lost your picture I mean erm people used to come up and say oh Christ I've gone a bit too far and that becomes, because then the meeting then gives the impression that you actually don't know what you are doing on this, people don't like erm touching the buttons."

CF – " Do you think then in, if people were given some kind of erm training experience to use the technology, use the keypad, you know to you know to fiddle around with it and in that way they can manipulate it when they are conferencing. Do you think that will increase satisfaction of use?"

BH - "Yeah absolutely. Erm and I think it's, and my answer today is if you do it when you walked in did Donald sort of show you how to do it when you walked in today?"

CF- "No"

BH - "But I think the question is o.k you've not touched any of the buttons or explored maybe some of the things that we could do on video. Not that I'm saying you should but I think you know you sort of leave it alone and not touch it because it looks as though it's gonna be, it's gonna bite you when you do these things. I think, I think my perception is, there are two things to get over here, one which is the personal perception of what people look like on T.V and if you are self conscious that tends to be quite important. The second part is when you are in a conference you actually want to make more use of it and you actually want to do presets or something to show *Donald* as well, I mean I'm not just talking to you I might want to look at *Donald*. It's easy for me to actually to identify that I'm going to say you know I want to see both of you rather than just the one of you then suddenly the usa..usability of the system becomes much easier and I know that next time I go into that meeting if I want to see two people I need to set that preset up for, so when I press that button I can actually see the whole room, and from that perspective I think you don't need to do too much training erm to allow people to use it properly. What you do is you've got to build a little bit of erm confidence if you like, and I don't mean through training it's probably confidence of use and part of that clearly is training but part of it is that you are not going to look really really stupid you know to the person at the other end because everything's O.K."

CF – " So part of it's experience then yeah?"

BH - "Yeah it is, I think experience does actually achieve that but people do not and I do know that people do not use videoconferencing at all on one try if actually they don't think they would have got to that position."

DW – "But when we, when we sell someone a product and the engineer goes in an installs it they do give that person who's going to be operating it, whether it be the I.T manager or the conference, the room office manager, they do give them training, we don't just set it to them and say there you are and clear off. So yeah they are told what buttons to press."

CF – " and that does help?"

DW – "Oh yeah and if they do have a problem with it because someone's always going to go and press that grey button and shut the call down by mistake. Because when they come into my studio because I hire it out to business people I can hear it go "beep" and I know that they have pressed the button because they are trying to do something, which is great because they are trying to learn themselves and what I find is if a company comes back sort of two or three times I find a lot of the time that they've worked it out themselves what the buttons do; they haven't told me they've been fiddling with it unless they drop the call and come out *"Donald we've lost the call"*; I said what did you do, oh we pressed that grey button by mistake, well don't do it again, but if you want to do this you can move the camera up, down, left, right, zoom it in or out and you find after three or four visits you go in there and they say it's o.k we know how to do it now, we know how to mute it, we know how to move the camera and half the you know the training is done by people that come into the studio, and hire it by the hour and by the time they come to buy they are up and running anyway and they realise that it's not quite a frightening piece of kit, they you know choose the buttons which are relevant to them at the time. You know zoom in and out, left, right, volume up and down and mute, which is basically all most people use because they are quite happy for the camera to stay in…"

BH – "… I think the process is that actually pressing the buttons and understanding how to press the buttons is quite difficult, it's actually the confidence of doing it whilst you are sort of on air and it's you know the subject to come back to…(talk amongst *Donald* and *Bill* about the camera zooming in)…I think the point being as I said is that it's not just a matter of confidence, it's not just a matter of technology erm unlike most everything else. As I said I'll draw a parallel again with P.C. If your P.C, if you shut down an application erm accidentally, no one's actually going to give you a bollocking about it, no one's going to know. If you actually shut down a videoconference erm accidentally everybody's going to know if the call's wrong. So again you know experience you know experience is an example. Pressing the buttons is not a problem, actually making productive use of the video communication actually is and people are a little bit weary whether on a live call or on air if you like of making a mistake and you draw a parallel again of the newsreaders when they get it wrong it's so bloody obvious that it appears to be so embarrassing that you can't possibly believe how this bloke could end up showing his face again, but erm I know it's somewhat of an extreme example but that is exactly what people do if they feel they've you know made a mistake on videoconferencing erm and I think just to sort of round it out most videoconferencing tends to be used higher up in an organisation at the more senior level and the risk of getting it wrong or looking stupid is actually quite high and, and the more that, as I say, the more that it de-layers the sort of erm the organisation, very often the better that the videoconference take-up of business actually is. Now people just walk in and use it rather than feel as though they have got to be you know on their guard all the time and be and look their best and you know say something that is intelligible rather than not."

CF – "Just one point erm to make here. You some to be gesturing quite naturally erm when you are talking to me here. Most people or a lot of people that I have spoken to or I have sent questionnaires to especially people who have just started to use the equipment find it difficult to gesture to a screen. Do you think that is something that you have learned through experience or…"

BH – "Yes, no question. Erm again just from my past experience, I've been doing this since 92 so erm I'm not an expert, I'm not an expert at all, but just experience of doing it a lot. Erm, it, it is, again if you feel you are talking with some authority or intelligence on a

particular subject then I think you can become quite, you, you actually become very close to a normal person to person meeting. I think the problem is that videoconferencing can be somewhat intimidating to start with which inhibits your you know your sort of natural erm your natural way of doing things but appearing on television having this sort of appearing on television issue, because you don't often do videoconferencing all day everyday, it might be a once a month thing you sort of lose if you like, you lose a little bit of confidence, so the more you use it, it's not necessarily experience of using it, but the more you use it the less intimidating the whole thing becomes and the more you know confident you become in pressing the buttons, but just being in a video conference is actually still intimidating to people if they are only doing it once a month so the more you use it the easier it becomes, the easier it becomes the more you use it. It's once of those things as I said earlier, if you get to a plateau level of quality erm then people just walk in and use it, and if you allow them to walk in and use it then it becomes almost second nature."

CF – " As far as confidence and erm communicative ability, do you think erm video feedback is some way to deal with that, some method. Do you actually use it yourself?"

BH - "Yes, yes. I mean I actually did that in the early days because erm what we tried to assume, what we assumed in the early days was is that because not very many people videoconferenced they had to be taught to do it. Erm, in a way it's the same thing as television training for senior executives of companies, you know if you are going to be in front of the media then you've got to portray this company well. It's a completely different kettle of fish from just talking to people in a general meeting for example. So erm you know really that's where erm the phrase is coined industrial theatre in the sense that you need to be a bit like a performer in a way to carry it off. What I subsequently found is it's not like that at all, I mean that does occur no question but what you do need to do is really assume as so though the meeting is going to be productive, it's ,going to save you stress, time and aggravation. If I had to go up to Edinburgh for this meeting for example you know sort of six hours each way for the same sort of content time wise, you know the stress and aggravation would have been huge for me erm meeting you would have been a huge stress that's for sure erm (laugh) the point being I guess is that because it is a productive erm productive tool and I do want to use it all the time that I don't feel inhibited by just doing and it's no just personality I mean it's just an experience, it's just an experience thing and I don't feel as though I need to perform again internally, it gets a little stressful but I have to say that's actually external and I think maybe you know your experience today of doing this is an example and you may be a bit weary about it to start with but once you get into it and start you know start having a conference then it's actually O.K and internally it never becomes a problem."

CF – "Can I sort of direct a question to the both of you? Erm I'm just going to backtrack here a little bit but these are you know kind of important things that I need to find out."

BH - "yeah sure"

CF – "Erm, what kind of people do you supply video conferencing technologies to or you know your actual videoconferencing suites. To whom do you supply them to. You mentioned a few…"

BH – "…erm in general terms it's more the level of senior management individuals, erm very few, very few that I would say that are erm outside that sort of range and they tend to be people that call meetings they tend to be people that would erm organise other staff."

DW – "Yeah, it depends upon the industry that are using it. We have a lot of advertising agencies that use us, Sachi, Grey, these sort of people and you get a lot of the account directors come in who are talking to proctor and gamble about the new Persil soap ad, and you get the designers and you get the post production people. You get the creative people come in with advertising agencies. If you get people like Inchcape or Phisor or these sort of people or like the banks you tend to get the top guys like Eddie George has been in from the Bank of England and you tend to get the C.E.O, the chairmen, the F.D or the M.D. Erm we had a big one the other night when G-AC and JRB merged, you've heard of them, well we had the two top guys in the other night to talk to bankers in Canada because they only announced it that day, so we get, then we get the other end of the scale, we get people coming in for job interviews, that is big business for us, like we get a guy who wants a job as a, I don't know, a financial analyst for state street bank in New York, so a head-hunter in America would interview him in London. We get people in the I.T industry getting interviewed, not top, top jobs, but a guy who wants to become a network analyst for telecom New Zealand. So we get that, that scale, but when it's business to business meetings we do tend to get the top echelon in erm but the we get film stars in, we Judy Dench, we get Phil Collins we get Seal, we get Nick Ross and we get Freddy Truman, we get those sort of people"

CF- "What do they use it for?"

DW –"…(Promote films and albums)…So they are many, I mean I've seen all sorts of applications, many, many, many applications, but they do tend to be senior management upwards, you'll get the P.A come in and you know you'll get the erm someone coming in and making notes for instance although with videoconferencing they don't have to any more because we'll just videotape the meeting for them, so they've got a visual and a verbal record of the whole meeting and that happens a lot where we've been involved in company take-overs where we have to be very erm confidential to the client, we mustn't open our mouth and say anything. There was one particular company that took over a file-o-fax type company and I had them in for eight months in a row while the negotiations went on and all that was videotaped because, and minuted, so that you know that everything that was said and discussed was actually there as if they had been in the room"

CF- "So it seems to be quite, the applications seem to be interview, business meetings, negotiation…"

DW – "…Yeah, yeah…"

CF – "…do you ever get any people designing things or collaborating on tasks?"

DW – " Erm you get people showing costumes for adverts, PowerPoint, we had a guy who was selling to shareholders in America a new financial plan, one of the big management consultants in London, he did a laptop presentation across the, the link to investors in America, he hit six American towns in six days without actually getting on a plane and he was able to put across his case why they should go for his advice, erm…"

CF – "…do you think he would have been more successful face to face?"

DW – "No, because they were appreciative of the fact that he was a forward thinking erm consultant and he was saving himself time and money by meeting them over a videoconference link and it worked in his favour, so they were quite happy…"

BH – " …That's the same thing, that's the same thing also I see for people that are coming from sort of a broadcast mode that actually do conference and can actually be a bit of a character, actually do better on video than they do live erm in a, in a sense the more charismatic you are the better you are on T.V and from that perspective you can actually cover half, let's also consider that you know it's much more productive to do what he did and he's actually, he can do that very much more quickly than he can ever do live erm and I think in some instances I've seen people say well I'm using videoconferencing because I am going to demonstrate how I can be more productive and how he should be more productive and in a way the argument is actually very simple and very logical"

DW - " There is another factor, is where a lot of major companies, let's take Marks and Spencers, I'm not saying this is what they do, but let's take Marks and Spencers who have lots of vendors all over the country and in Europe who make their products. A lot of the big corporate companies are now saying to their suppliers if you want to supply us with 2 million socks a month you've got the install videoconferencing because we are not going to keep coming to you and we don't want you to keep coming to us we want to talk to you about a problem now and we want to be able to talk about it and get it resolved quickly so you have got to buy videoconferencing if you want to do business with us and that is happening quite a lot now especially in the states…(Harley Davidson as an example – particular makers of certain components)…Advertising agencies have saved weeks on T.V ads by showing them across the system to their client in another country rather than flying out there playing the tape and flying back again. So it is, it is, it becomes part of company culture with a lot of companies and, and, basically it is saving them time, management time, money because of air-fare and accommodation etc., and it improves their quality of life because people are spending more time at home and they are not spending eight hours sitting on a jet twiddling their thumbs."

BH – "I think the other part of that *Donald* is that is to consider what the contact that is and mostly as I said earlier is that is actually iterative, it's actually tweaked things erm from a basic contact, I mean if you are talking about T.V ads, about maybe machine and parts erm quality control I think is an iterative process which says maybe the first meeting is going to be face to face to establish it, the second meeting might also be face to face to develop it, the third meeting is a tweak and therefore is video and as you establish that sort of process erm it becomes much more productive to continue that iteration more often than little tweaks than go to client and say absolutely hate that colour it's terrible, need to change it, you know before you are off the blocks erm you know you end up with a situation that you have to go away after a two minute meeting and re-evaluate, where you can do that iteration process in a fifteen minute meeting on video and you can say well actually I don't like the colour we'll go back and we'll modify it now, we even might do it on screen while we are there, while we're doing it, what do you think of this? And that might be an extreme example but it's typically the sort of the thing where erm maybe the project work's established and it does work very well in project management type situations, the project work is established but what you are doing is fine tuning it and actually reducing significantly the time to market by arranging lots of short modifications, because you actually need to do that, Ford is very good example, typically they say they shave eighteen months off the development of the Mondeo because

they actually talked visually about the engine mounting, who, where they were making erm Bridgend in Wales I believe or the engine to the mountings that they were being developed in Germany and because they could actually share you know practically here's the engine bits and this is a problem we've got, you know because they could actually share that, modify that bit like this, you go away, we'll go away come back tomorrow and see what we've done, could not have been done in any other way and computer design works you know you need to make it at some stage and you need to understand why that engine wont go in it's engine bag because when you drop at the end of this crane you know the engine's sitting at this angle, so in a way erm there is a common thing here where videoconferencing becomes very much more productive and that's why it's inter –company, and that's why very rarely, sorry intra-company, that's why it's very rare to try and supply a relationship in this sense, because you sort of don't dr.., often enough in general terms to make videoconferencing any more productive than a face to face meeting and I think that if you look at one that carries that thing through you end up with erm a lot of very productive sort of situations which, which a fifteen minute videoconference would fix real good which you would never do otherwise, and perhaps then you could improve the quality or the time to market in small perceptible steps rather than in one giant re-design, which then may not actually work"

CF- "As I mentioned earlier, sorry I'll just check on the tape to make sure it's still recording, as I mentioned earlier one of the major erm elements of my work is to look at the kind of problems that exist in videoconferencing both technological, human you know whatever. Could any of you two answer this question? What kinds of difficulties do your customers report or in essence what kind of problems do you think exist in the use of V.M.C technologies?"

DW- "Some of the, well we get a lot of phone calls in London from customers, in fact they usually ring Slough, but the technical centre's where *Bill* is, but if they can't get through to Slough they'll come through to us and I have to be blunt about this. A lot of the errors that we get that we deal with have got nothing to do with the equipment at all. It's either users erm inexperience or they don't understand the keypad because the supplier, not normally us, but a, or it could be someone who has come into the room who has not been trained wants to know which button to press, why has the picture disappeared off the second screen. The biggest problem tends to be with I.S.D.N lines whether, you know, the I.T guy has connected, if they have got cat 5 cable in then they haven't organised the server or the switch to feed the lines into the V.C system. So what we can do from Slough is actually test that line remotely from Slough and let them know A if it's live B if it's engaged or C if it exists at all and that is a lot of the problems where we can't physically help them because it's erm a B.T or a line problem within their own facility and that's frustrating for us because there ain't a lot we can do about that apart from go round there and plug them in"

BH – "In the old erm in the old adage of erm an estate agent, what's the most important thing for a house, position, position, position. In the videoconferencing world it's network getting fixed, network getting fixed, network getting fixed. Erm, a usual problem 65% of the time they make a successful call first time, it's something like 99.999% in terms of a telephone conversation. So, so not getting through is a big problem. You've experienced a little bit of delay when we didn't get audio into the room for some reason. Now there are, there are still erm a lag of technology in confidence of connection, which is the number one issue without a question. There is, there is, that problem tends to be compounded by the fact that people don't use video conferencing, or the people that don't use videoconferencing very often forget what they have to do to actually make videoconferencing work. Dial 9 for an outside line, or

dial three zero's to get to America, because if you want to go across the Atlantic you have to add three zero's instead of two zeros. So there are simple practical things which is, which are generally as a result of the network or of the interface that are used in videoconferencing. Erm one of the things that stopped people in practical terms using it very well you know is the time difference we are talking about. Sometimes video systems are turned off, actually making the appointment to meet someone else in Australia and find a common time to do it is actually quite hard. So there's, there's sort of windows of opportunity in videoconferencing for example erm and I think in, in general terms these are not the nicest and most friendly erm interfaces that you'll ever find. When you actually did the dial to us, you saw *Donald* type in a phone erm hieroglyphics. It's all right for him who does it quite often, but if you get a P.A who comes in, I'm not trying to be disparaging here, but a P.A who comes in once a month and knows bugger all about the technology those erm hieroglyphics, you can actually have a star instead of a hash and end up without a connection. So, so, the, the I guess the practical things here are erm the network isn't reliable and continues not to be reliable erm the video systems tend to be a little unfriendly in places and I can see that's getting better and the confidence that, that erm the connection is going to be made first for either, for any of the above reasons tends to be relatively low so guys like *Donald* who have got some experience in all the problems that have ever occurred probably in a video conference tend to be invaluable to clients here who rely on that expertise just to make it work in the first place."

CF – "Now there's a lot of people who have also mentioned that familiarity is a factor which effects communicative success and what I mean by that is erm if the people have met before hand, so if they've met in a face to face situation before hand and if the know person, they're, they're familiar with their communicative style, the way they gesture, who they are erm then when they meet them over the video for the second time it makes it a lot easier than say the people who just meet...."

DW – " ...Yeah absolutely, absolutely yeah"

CF – "Do you both agree with that?"

DW – "Yes"

BH - "Yeah erm I think it's interesting, I'll actually just sort of counter that in a way. I often, I've met, I've met so many people in so many countries that I would have never have met before and I think you have to, you have to maybe erm think a little bit about the opportunity erm option here which is I've only met those people because I have given, they have given me their video numbers and I've actually dialled them up and talked to them and we used to have a policy if we wanted to get someone to do a demonstration in the U.S we would never, we would never phone them up but we would always videoconference with them, we would find their video numbers and then videoconference with them because it is not difficult for a guy to say when they have actually met a couple of Brits over the other side in an office and you know and have managed to create the sort of rapport with them erm and they would never say no you know they would always say yes of course we'd actually do that. Now and I am not saying you couldn't do that on the telephone but I think it is very easy on the telephone to say no because you are actually talking, you are talking to an abstract voice. It's very difficult for me to say no I'm not going to answer any more of your questions because we have sort of got a little bit of a personal relationship, however short it is it's a lot better than sitting on a telephone doing it and because you don't see me you don't know me erm on the telephone you could walk past me in the street and not give me any aggravation (laugh) whereas as now

if I actually walk past you and you don't say hello you know in the Scottish office perhaps when I'm there, you know we have actually established here so I think it's also an issue that the opportunity of meeting people grows enormously erm where that first meeting will actually permit erm a second or third perhaps personal meeting. So It actually works the other way round and I think in establishing business I've found much more success in the people that I've met erm on video that I would never have met personally and that's from a whole bundle of countries I've never been to: Australia, Saudi Arabia, I've met people there on video erm well I think I could probably call them up and, and erm call in a favour if you like and make it good as I have actually videoconfernced with them. I think that opportunity is significantly higher than it would ever be else where than in any other medium...(too quite)..."

DW –"It, it is fairly rare for people to meet on video for the first time except in a job interview situation. Most companies or people that meet have either spoken over the phone, they know who the other people are, unless the marketing director's come into the meeting that the guy has never met before, oh this is John our marketing director, but most people know each other erm and what it does, what videoconferencing tends to do it does, seem to encourage people to meet more often. What it also does, what we are not trying to do is stop people flying or going and meeting face to face, we are not trying to do that. What videoconferencing does is allow people to meet more often, so in other words the guy who went to New York six times a month now only needs to go three times a month because the other three meetings they can meet over the video."

CF- "One application in which erm, I mean I agree with you that most of the time people would have some background with each other if they were videoconferencing with each other, but one application in which that wouldn't happen would be in a distance learning situation where the lecture wouldn't..."

DW – "No that's true, that's true no..."

CF – "have had any previous contact at all with his students. Do you think if those students would have met the lecturer beforehand say in a half hour or twenty minute meeting where they could familiarise with each other that would increase erm the conversational benefits"

DW –"I don't know because it's not really a one to one is it? They wouldn't have met that teacher anyway, if he had come to their college as a consultant from another town they wouldn't have met him before either. It depends on how..."

BH – "I think actually, I would say it would but as I said before, I think the most important about videoconferencing is it's an unusual way of meeting people, but it actually does mimic pretty closely a personal meeting, other than smelling you which is perhaps the only other, perhaps the only other component which that we need to consider here, you need to perhaps a little sniff, how many you give for stuff like that, but I think the point is if you actually have the ability to erm converse, providing it wasn't and this is the other thing, providing it wasn't the first time that you did it, I think it would be extremely valuable because in the first few minutes of any meeting initially ...(Boy's room) you see whether they interested in erm you know their communications, whether they are enthusiastic and whether they actually want to do this or not and that is actually very difficult to replace...(unintelligible)...and I think video, because it does focus the attention on you know one person to the other person you

can't get sort of lost in a crowd and it's very easy in a distance learning environment to be one of a number you know there's the broadcaster, there's the broadcastees. I think the, the simple analogy is if it, if I was to erm meet a celebrity for example I'd feel I'd have a huge advantage to get him to focus and talk to me about something that he and I were maybe more interested in. If I went to a erm, erm a show or something where the celebrity was in a crowd but I didn't meet him, it would do me no good. I wouldn't know any more about the guy, it would be, it would be, might be a story I could tell my friends about, but I wouldn't be particularly proud of it. So I think there's this, there's this perhaps erm misconception in a way that video does something different from a personal you know from a personal meeting. I think it is extremely important to create personal relationships in this way and video is a very good method of doing it, probably more focused actually than meeting people face to face. Again you don't tend to find that people's minds wonder or you talk about things or you are distracted elsewhere and in an educational environment I think you need absolutely to believe in the fact that the guy is actually going to tell you the truth and most hopefully make it interesting for you to want to learn, because you know, there's no, there's no payback. You don't get a wage for learning something and you know the motivation and the interest that has been created is actually quite important and I think that is something that is missed and would be extremely useful in a distance learning environment. Perhaps sort of hype it up a little bit, you know it seems a strange think to say, but make education and, erm meeting quite an important professor form the states for example you know you'd never meet that guy any other way so lets have a conversation with him see what it's like you know what's it like in the states teaching for example and that's a very effective way of erm I think of creating an interesting (cough *Donald*) ..in a job or in a videoconference?"

CF- "Another thing, I mentioned briefly eye contact. Say we were having a one on one meeting and we both had access to each other's eye's and face erm, it's difficult to see your eyes because you are quite far away from the screen, the screen isn't focused on your face and your head. Erm do you think eye contact would cause a problem, because we couldn't actually get eye contact with this erm in a one and one meeting over videoconference because you know the camera is positioned above the monitor. Do you find yourself for example looking towards the camera when you are emphasising points and things like that, because if you looked at the monitor as I was looking at the camera it would look like I was looking into you eyes"

BH – " Yes what I am going to do erm I mean answer that question and I do erm and I do erm, you were going to add something here? Is that I can actually get your head twice normal size, it's sounds crazy to say but I can actually focus in on you erm and actually get your eye contact at twice the size I would normally have it, yeah? So in fact that if I really wanted to be a bastard and that's the way that some of the people that we sold, and \I have actually heard this in a statement. The chairman of the meeting grabs hold of this (Hand control) I mean in control of this meeting, so he zooms into your far end and he gets as close as possibly can and he says and what's the sales figures this month? So in effect the technology allows me to change the emphasis or the perspective of the meeting quite significantly by actually focusing on you and saying O/K are you going to get a first in your studies? Now tell me are you going to work really hard, you know is that what you are really going to do because if you are I want to know about it, so in effect it is actually the opposite in that you can effect the dynamics of the meeting by using this technology. Though we choose to don't I'm being rather aggressive now by zooming I on your face, I really only want to see the whites of your eyes type issue but I'm actually in control so you are not going to see mine. I am not going to allow you to do that so I'm in a, in a, in a meeting sense I'm in control of this big time,

because I can see you big time and you can't see what I'm thinking or saying because you can't actually see, you can't actually see what I'm doing but in that, I think that's slightly erm perhaps example of the technology could actually have effected this (meeting) quite significantly, the fact that, that we choose not to in this environment is really down to the dynamics of the meeting rather than necessarily, than, than erm fixed by choice or fixed by design within the environment we are in. It can happen but the question is whether we want it to or not."

CF- "To clarify a bit further erm some studies in psychology have shown that eye contact is important in conversation as in regulating conversational turns. People tend to look more at another person, in their eyes when they want them to take over and this has also been studies in videoconferencing in that there are more interruptions. People don't know when to come in, don't know when it's there turn and erm you know the conversation will usually take longer aswell. Erm if you an I both could see each other's full head and face, If I looked up at the camera you could then see my eyes and it would look like I was looking at you or looking into your eyes. Do you think that would help at all in that situation?"

BH – " Yeah I think, again the layout of the meeting room is important because the further you are away from the camera the more obvious it looks like me looking at beings looking at the T.V. (talk amongst *Donald* and *Bill)*. To answer your question yes I absolutely agree, but the thing is if I keep, psychologically I agree yeah I would agree with you, I'm actually not going to interrupt you because I am going to keep encouraging you to talk and I am going to keep nodding and I am agreeing and it is one of those examples of erm you know of psychology of meetings in fact and eye contact absolutely is true in video. As I said before because you tend to be focused towards the T.V screen that's your realm of erm that's your realm of interest then they become even more accentuated in my opinion, but you do need to actually use that technology to make it happen, and the presets buttons ...(talk amongst selves about zooming etc.)..."

CF – "So I mean what would, in essence as I mentioned is like, I mean obviously I mean just from using this here and talking to yourself you know it is very satisfying. It, it is you know, I am quite impressed with the technology but to increase satisfaction to increase you know erm not just satisfaction but the use of it erm do you think something like training people to be aware of their own bodily movements, you know what they do and also to be more aware of what other people do and signals that people use to infer that they want someone else to tale over or to infer you know that they mean some other point. Intonation for example."

BH – "yeah absolutely and that is what I am saying. Because the focus of attention, because you can use technology to switch pictures between, I can move the camera around I can switch it to *Donald* now erm then the technology can be used to control meetings in that way, but it isn't natural erm in a sense that the focus is always on the speaker, so *Donald* is irrelevant to me at the moment, although he is in the room with you and can hear the conversation I can't communicate with him in any other way. Erm we could do that by setting up a preset, we could flip between those two pictures, we could focus on what I, I could control the meeting to the extent that I could focus on those `individuals by pressing buttons..."

CF- "...As far as..."

BH - " And there are some technologies that would, *Donald* mentioned a product called "" which is a voice tracking camera which we could employ to do the same thing, and in, in a sense that is a much more natural meeting environment there's loads of settings to, to reduce the control of the meeting and therefore reduce the impact of being able to do what I'm doing with you which is zooming in close..."

CF - "...How about sort of the hydra set up..."

BH - "Sorry"

CF- " Have you heard of the hydra set up, where it is like one monitor for each person?"

BH "Yes"

CF – "Do you think that would be advantageous over a voice activated mode?"

BH – "Yeah I mean that, I've seen hundreds of implementations of video switching, video personal screens, screens behind people or the screen up on the wall, other people on the sides of the screens here, voice tracking cameras, or movement detection cameras. What they tend to do is actually change the way you would normally have a meeting and I think what happens is that becomes less natural in a meeting environment than using the sort of processes we are using now in a very simplistic sense, but they do work no question bur you become rather erm uncommunicative with your, with your far end if you like, you become you become less focused on the content and more focused on the fact that there is a T.V screen sitting right in front of you and because other people maybe participating in the meeting and exchange as *Donald* is at your end then maintaining the focus of the one way, you know the one way porthole effect means that he knows what I am saying, you know what I am saying and we are not seeing anything different, so focuses his attention perhaps on that one output device and rather more than necessarily than flicking cameras around or video images around, what you tend to do is focus on the content of the meeting, and erm I think it's, I think it's not exercised here again but erm with that sort of focus you tend to be able to know when someone is going to go in and make a point because you can actually argue over it, you can actually talk you know and it's full duplex audio capability does allow you to interrupt and say shut up you know talking too much and I want to talk to someone else..."

DW – "It's quite an interesting point actually because my experience, I had 275 different companies come and use our room in London and generally if there is 3, 4 , 5,6 people they want the view like we've got here and I've had the limelight camera on and everytime they speak it zoomed in on them they say oh no we don't want that I don't want a head and shoulders I would rather we kept it with the room view"

CF – "what that be you think because they'd be you know self conscious..."

DW – " I don't know because I don't say oh why what's the matter with you, don't you want him to see the white's, I don't you know, it's none of my, it's what they want, I give them what they want. The only time you ever do a close up is for job interviews, where the guy is, obviously they need to, it's a bit more of a critical meeting because they have got to see how the guy reacts and everything, but for meetings generally for 4, 5,or 6 people they tend to want the room view and if someone wants to speak, say we are sitting here now talking to Phil and you are talking then what normally happens is the next guy who wants to speak will

go like that (hand up) which you might do in normal meeting, if someone is going on and on and on, you might want to, before you interrupt you might just go can I say something, you automatically do that don't you?"

CF –"yes"

DW-"Hang on John can I come in here, you don't just say, most people either make a signal that they want to interrupt a floe and that happens in videoconferencing you'd be surprised. You know even if people in London want to interrupt a guy in Chicago, the guy in London will go like that (Hand up) hang on guys and even though, he's saying to them I'm going to interrupt here and he'll put his hand up and say hang on John let me just come in there you know, so it's amazing how people adapt and, and just get used to it."

End Interview

Appendix 2: Example of completed map used in 'collaborative map reading task.'
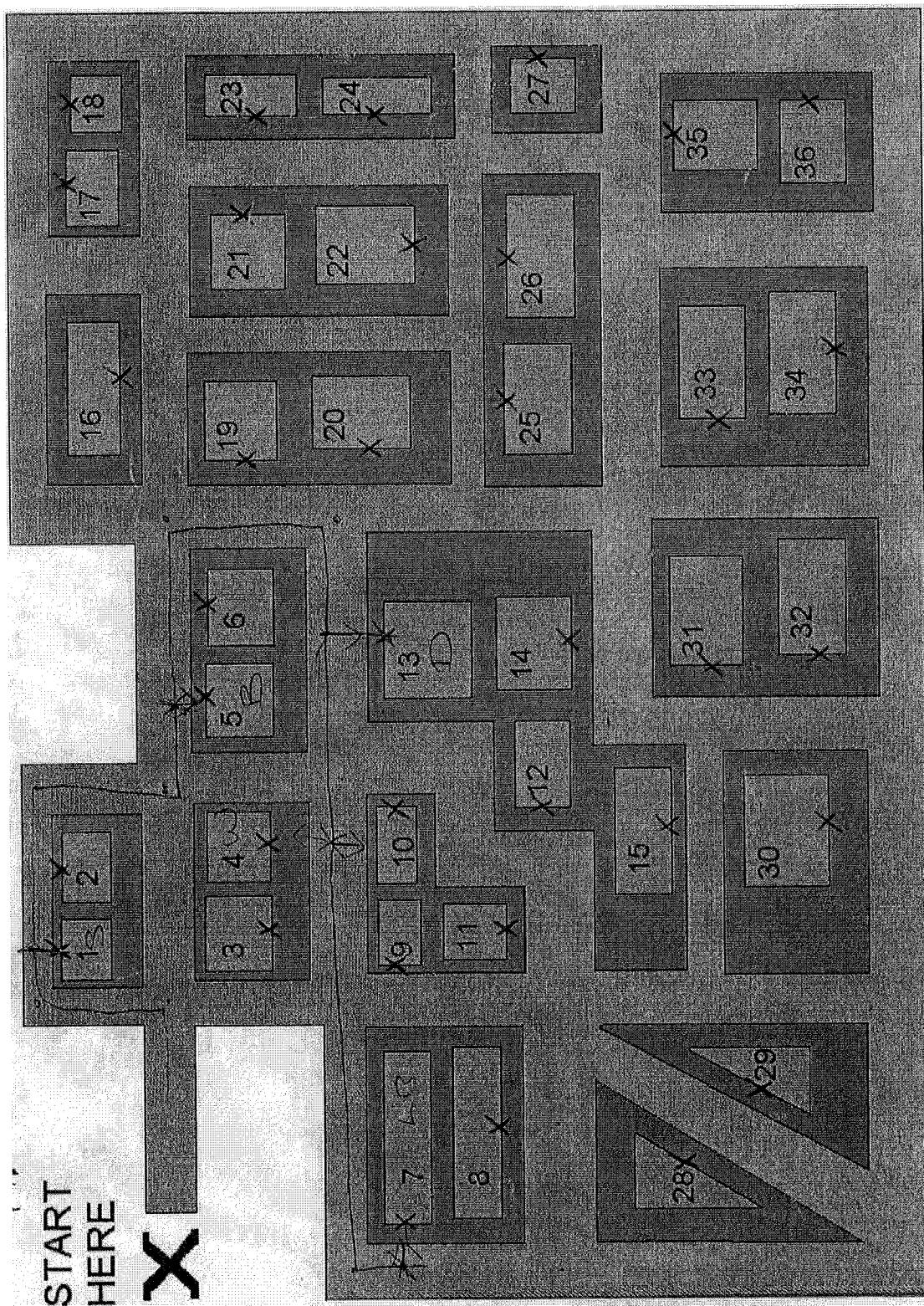
Appendix 3: Accompanying list used in 'collaborative map reading task.'

1. Scriven's (Steak - £2. 50)
2. Post Office
3. Alan's Bakery (Bread - 30p)
4. Shopmart (Bread – 55p, Wine - £5. 60)
5. Molly's (Bread - 25p)
6. Pub
7. Electric City (Light Bulbs - £2. 05)
8. Martin Smith's (Bread - 50p, Steak - £3. 05)
9. Animal Heaven (Dog Food - £1. 15)
10. Shoe Shop
11. The Wine Cooler (Wine - £5. 85)
12. Sava Penny (Light Bulbs - £2. 75, Dog Food - £1. 15)
13. Pet Palace (Dog Food - 95p)
14. Record Shop
15. Thornton and Sons (Steak - £2. 35)
16. Pets-R-Us (Dog Food - £1. 00)
17. Green Grocers
18. Basement 10 (Wine - £5. 25)
19. Travel Agents
20. Samson's (Light Bulbs - £1. 95)
21. Price Cutters (Light Bulbs - £2. 10, Dog Food - 80p)
22. Café.
23. Eastside Corner Shop (Bread - 40p, Light Bulbs - £2. 00)
24. The Grape Vine (Wine - £5. 00)
25. Newsagent
26. Mr. Fixit (Light Bulbs - £1. 80)
27. Granny's Bakery (Bread - 35p)
28. Bank
29. Spendthrift (Dog Food - £1. 10, Wine - £6. 50)
30. Parkinson and Bowyer (Steak - £2. 20)
31. Animal House (Dog Food - 35p)
32. Clothes Shop
33. Flower Shop
34. Speedymart (Wine - £5. 60, Steak - £2. 20)
35. Mulling's (Steak - £2. 00)
36. Charity Shop

Appendix 4: Instructions given to participants in 'collaborative map reading task.'

COLLABORATIVE MAP READING TASK.

Instruction Set One.

You are a sales assistant working in a town centre office. You are on your lunch break and have promised to drive a friend round the shops in order that he may buy five items from his shopping list. These items consist of bread, steak, wine, light bulbs and dog food. You have been given a map that is a representation of the town centre. On the map the grey area represents the road, and the numbered orange blocks represent the shops. The 'x' marked on these blocks represents the entrance to the shop. The block numbers correspond to a list of 36 shop names that you have also been given. On this list, shops that sell any of the items on the shopping list are highlighted with the item(s) printed next to the shop name followed by the price of that item(s). There are also shops on the map that do not sell any of the items on the shopping list. Considering the fact that you are on your lunch break, your objective is to take the shortest route possible so that you can return to the office and enjoy the remainder of your break. You must however, collaborate with the purchaser so that both of you finish with the same route. In order to purchase an item from a shop, you must drive past the entrance of that shop. No back-tracking is permitted; you must decide upon a continuous route from start to finish that does not double back on any of the roads you have traversed. You may plan your route before hand or make changes to your route through erasing and drawing again. Your route will terminate at the last shop you purchase an item from. Remember, your aim is to draw the shortest route possible

COLLABORATIVE MAP READING TASK.

Instruction Set Two.

You are a student and you need to buy some items from the town centre. These items consist of bread, steak, wine, light bulbs and dog food. Your friend is on his lunch break and has offered to drive you around the shops. You have been given a map that is a representation of the town centre. On the map the grey area represents the road, and the numbered orange blocks represent the shops. The 'x' marked on these blocks represents the entrance to the shop. The block numbers correspond to a list of 36 shop names that you have also been given. On this list, shops that sell any of the items on the shopping list are highlighted with the item(s) printed next to the shop name followed by the price of that item(s). There are also shops on the map that do not sell any of the items on the shopping list. As you are a student you do not have much money and so your objective is to purchase the cheapest items possible from the various shops. You must however, collaborate with the driver so that both of you finish with the same route. In order to purchase an item from a shop, you must drive past the entrance of that shop. No back-tracking is permitted; you must decide upon a continuous route from start to finish that does not double back on any of the roads you have traversed. You may plan your route before hand or make changes to your route through erasing and drawing again. Your route will terminate at the last shop you purchase an item from. **Remember**, your aim is to buy the cheapest items possible.

Appendix 5: Story on 'Vernon the very old wolf' with list of important points from story.

Vernon – a very old wolf

Once upon a time, long, long ago lived a very old wolf called Vernon. Although Vernon lived in a beautiful green forest in which there was plenty of food, he was getting too old to catch anything to eat. This was a real shame as his favourite meal was deer, which were too quick for him to catch.

One day Vernon was walking through the forest feeling very hungry when he came upon a black castle on the top of a hill which had blue flags at the top of it's turrets. On these blue flags were symbols of a golden lion, with a red cross behind it. This could only mean one thing – this castle belonged to the Queen.

Vernon noticed a pleasant smell in his nostrils. There was no mistaking it....he could smell the delightful smell of roast chicken. He couldn't resist the smell so he snuck inside the castle. Then he snuck inside the kitchen and as quick as a flash grabbed the chicken off of a large, brown oak table and gobbled it all up.

Back in the kitchen things were not so happy. When the Queen heard news that her dinner had been stolen she got very angry and stomped off to bed with a very rumbly tummy. Nothing could console her, not even the sound of the harp, which was her favourite instrument.

Vernon had enjoyed his dinner so much that he decided to come back for some more the next day. Again he sniffed the air, but this time he could smell pot roast. He couldn't resist the smell so he snuck inside the castle. Then he snuck inside the kitchen and as quick as a flash grabbed the pot roast and gobbled it all up.

When the Queen heard the news she was very angry and very hungry. This time, however, she decided to call upon her best huntsman,

who's name was John. "Tomorrow night" she said, "I want you to lie in wait for the wolf so I don't lose another dinner."

The next night Vernon came back to the castle and sniffed the air. This time he could smell fish. He snuck into the castle and snuck up to the kitchen. When he peered around the door, however, he saw a big black boot sticking out from behind the table. Vernon was a very clever wolf so he knew this was a trap, so he went into the dining room and hid behind a large purple velvet curtain.

The Queen thought that the wolf was not coming this night, but when her food was about to be served up, Vernon spotted his chance and grabbed the fish and gobbled it all up.

This time the Queen was not angry, however, as she realised what it was like to be hungry. So from that day she decreed that food would be placed outside of the castle for the wolf to eat so that the wolf could be happy and so that she wouldn't lose any more dinner.

Old wolf called Vernon.

-Lived in a beautiful green forest.

- Too old to catch anything to eat.

- Favourite meal was deer.

- Came to a black castle.

- Blue flags on turrets.

- Golden lion symbol with red cross (Queen's flag).

- Took chicken off a brown oak table.

- Queen went to bed hungry.

- Harp couldn't console her.

- Vernon took pot roast.

- John the huntsman called upon.

- Vernon smells fish.

- Big black boot sticking out from behind the table.

- Vernon hid behind a large purple velvet curtain.

- Took fish.

- Food outside castle for Vernon.

- Everybody's happy.

Appendix 6: List of advice given to participants in study 5.

Please read the following information carefully:

Videoconferencing systems employ the use of a camera (to pick up and send a signal) and a television screen (to show the signal). Videoconferencing has been described as the best alternative to face-to-face communication as, unlike using the telephone, visual as well as audio signals can be transmitted. For a typical videoconferencing set-up, please see the figure below.
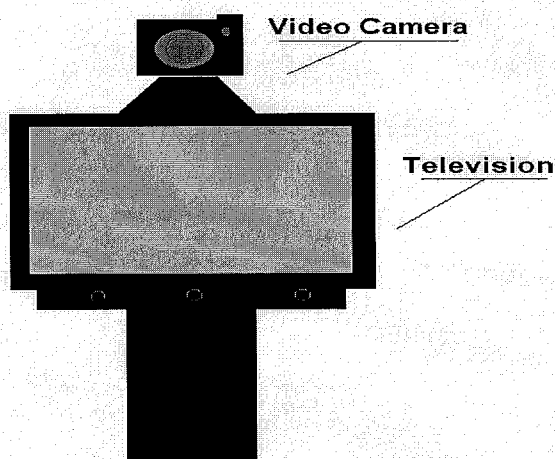
Videoconferencing, however, does have its problems, and is not as effective as face to face interaction. Many reasons have been attributed to this, including the fact that audio quality can often be quite poor, and also there is often a delay between the transmission of visual and audio signals. Even in the use of high quality equipment (with no delays and clear audio), people still perform less efficiently on a number of tasks over a videoconference compared to face to face interaction.

Another major problem associated with the use of videoconferencing concerns the fact that the camera is mostly placed above, and not inside the monitor (television screen). This particular set-up results in problems with eye contact.

Most users of such technology "look at the image of the person they are talking to on a monitor and not at the camera. The result is they appear to be looking off camera and they do not have the eye contact which gives conviction to what they are saying" (Tiffin and Rajasingham, pg.112). Therefore, a normal videoconferencing set-up can be quite off-putting as the person you are talking to will appear as if they are looking away from you, almost as if they are disinterested. Because eye contact is also very important in human interaction, the lack of eye contact can also make communication less productive. Eye contact, for example, has been shown to play important roles in the turn-taking process and evaluations of confidence, honesty and attraction.

Devices called "Video tunnels" have been developed which allow people to have eye contact through the use of mirrors being placed inside the monitors. Unfortunately this type of set-up is unfeasible for long distance communication. Studies have shown, however that people communicate as successfully in video tunnel conditions compared to face to face, which seems to imply that the major barrier effecting videoconferencing being as effective as face to face interaction is the lack of eye contact.

Measures, however, can be taken by users to improve this predicament. Simply by looking up at the camera the impression will be given that you are looking your partner in the eyes. Even though this cannot be considered natural eye contact, it is, however, better than none at all. As far as when the user looks up at the camera is a matter of choice. For example, the user may want to look at the camera to emphasise a particular point in the dialogue, thus giving conviction to what they are saying. The user may want to look up at the camera when they want to interrupt. In essence the most important thing is that some form a eye contact can be achieved which makes the videoconference less intimidating, more friendly and closer to face to face interaction. Users, however, will also need to strike a fine balance, as too much eye contact can also be considered intimidating.