The Role of Perceived Proximity in Video-Mediated Communication

David Michael Grayson

Thesis submitted for Doctorate of Philosophy at the University of Glasgow.

Department of Psychology
Faculty of Social Sciences
University of Glasgow
Glasgow
Scotland

September 2000

Abstract

As technology for remote communication continues to advance and become more widespread, there is a need for research to attempt to understand the manner in which such technology may most suitably support human communication. This thesis describes a series of experiments which investigated the role of proximity within video-mediated communication.

Proximity is one of the most fundamental forms of non-verbal communication used in a face-to-face interaction. Even subtle changes in interpersonal positioning are rich in information which people use to attempt to regulate the behaviour of themselves and others. At present it is unknown whether this type of non-verbal communication is preserved in video-mediated interactions. The aim of the present research project was to investigate whether impressions of proximity could be conveyed across a video link. In addition the research attempts to illuminate the physical parameters which may underpin the perception of proximity and to explore the impact upon users that any changes in perceived proximity may cause.

The research uses a wide range of approaches to study the potential impact of proximity including analyses of the structure and content of dialogue, objective and subjective task outcome measures. The research demonstrates that perceptions of proximity can exist in a video-mediated environment and when they do, they can lead to differences in the communication behaviour of individuals communicating across a video link. It is found that when participants interact with a remote interlocutor who appears to be close, they tend to be more interactive. The research goes on to investigate the perceptual basis behind this effect and also considers how this relates to other variables which are known to affect communication, most notably familiarity.

Acknowledgements

I would like to extend my grateful thanks to all the individuals who have helped me during the course of this thesis. In particular I was fortunate enough to have two supervisors who were always generous with their time when their help was required.

My sincere thanks go to my academic supervisor at the University of Glasgow, Prof. Anne Anderson, whose support and guidance was invaluable throughout the course of this thesis. Thanks also to Jim Mullin and all at the Multimedia Communications Group in the University of Glasgow for their expertise and assistance with the technical aspects of this research.

Thanks also to NCR Scotland and the ESRC who co-funded this research. In particular I would like to thank Dr. Lynne Coventry, my industrial supervisor at NCR, who always ensured that I got the most out of my time there.

I would also like to thank my present employer, Dr Andrew Monk at the University of York for his understanding during the latter stages of writing this thesis.

Finally, my thanks go out to all the friends and family whose faith, encouragement, good humour and cooking was invaluable throughout.

Chapter 1. Executive Summary

As technology advances, opportunities for remote communication using high bandwidth technology are increasing. This has led to the emergence of a new field of research in human communication in which researchers explore the impact of such technologies on their users. Video-mediated communication, which relays visual signals between remote interlocutors, is the medium that appears to offer the closest simulation of face-to-face communication. So far however, studies that have compared communication between users of different media have shown that video supported interactions share more similarities with those by audio communication technology than face-to-face conversations.

It is hypothesised that the reason for this is that video conferencing technologies do not support all of the communication channels present in face-to-face interaction. A video-mediated communication system is not a single entity however. There are many variables within a video-mediated communication system (image size, image view, frame rate, screen positioning, task purpose, camera position, eye contact etc.) which may affect the manner in which individuals communicate when using a video link.

One source of communication information present in a face-to-face interaction is proximity. Even subtle changes in interpersonal positioning are rich in information which people use to attempt to regulate the behaviour of themselves and others. At present it is unknown whether this type of non-verbal communication is preserved in video-mediated situations. The aim of the present research project is to investigate whether impressions of proximity can be conveyed across a video link. In addition the research attempts to illuminate the physical parameters which underpin the perception of proximity and to explore the impact of perceived changes in proximity.

1.1 Can Impressions of Proximity Exist in Video-mediated Communication?

The first aim of the thesis was to investigate whether perceptions of proximity could be supported by a video link. To do this, it was necessary to demonstrate that it was possible to systematically manipulate the impression of distance of an individual represented by a video image.

There are several points to consider when thinking of the perceived distance of an individual across a video link. The projected image of the remote collaborator is on a screen that is fixed. No matter what the image looks like, the distance between individual and screen is static and so the distance between the individual and the image of the individual can be considered static. It is possible to manipulate visual cues such as the size of the image, the amount of visible background and the amount of the individual that is visible. It is probable that these cues may have some effect on how far away an individual appears to be.

Chapter 3 describes a study that addressed the fundamental question of whether or not manipulating a camera image by changing the zoom on a camera could affect the perception of distance of a remote individual. Participants sat between an individual and a monitor showing a video image of a remote collaborator. They were then asked to alter the camera zoom until the person in the video image appeared to be as close as the co-present individual. Each participant completed the experiment 5 times with the co-present individual at distances ranging from 2.5 feet to 12 feet away.

The results of this experiment revealed that changing the zoom on the camera could change the impression of distance of a remote collaborator in a systematic fashion. In the experiment measures of participants' feelings at each distance were recorded. These results showed that way participants regarded the remote individual largely mirrored their feelings towards the co-present individual.

1.2 Could Perceived Proximity Affect a Remote Interaction?

The results of Study 1 showed that perceptions of proximity could be preserved across a video link. Since it is known that a zoomed image can cause feelings of different distances, it is important to know if these changes in perceived proximity could affect the manner in which individuals interact when communicating using video technology. In face-to-face communication researchers have shown that changes in proximity can affect the behaviour of participants in several ways. Do differences in virtual proximity lead to similar effects? Study 2 explored the impact on interactions of differences in virtual proximity.

Study 2 sought answers the following question. Could changes in the level of camera zoom used to display an image of a remote interlocutor affect the way individuals communicate in a video-mediated interaction? If so are the effects of altering the level of camera zoom similar to the effects of proximity in a face-to-face situation?

To explore the impact of perceived proximity on interactions between users of the video communications link, another experiment was conducted, which is reported in Chapter 4. For Study 2 an experimental task was devised which simulated a customer seeking financial advice from a remote advisor. Participants played the role of the customer and interacted with an advisor who appeared at different distances. The distances were altered by manipulating the view of the camera by changing the zoom function. In the close condition, only the advisor's head was visible, whilst in the far condition, the head and upper torso were visible, along with a lot of background.

The experiment was a between subjects design with 30 participants taking part, split evenly between conditions. Data collection consisted of transcribing and analysing the experimental dialogues for structural dialogue measures (turns, interruptions, words, share of the dialogue etc.) and a questionnaire which participants completed answering questions about their subjective impressions of the task, the advisor, the advice and the technology itself. These measures allowed the analysis of effects on both behavioural and attitudinal levels.

Results of the analysis of the questionnaire indicated no significant differences between conditions. No questions indicated that participants' subjective impressions of any part of the simulation were affected in any way by the experimental condition. Analysis of the dialogues, on the other hand, revealed strong differences in the structure of the dialogues between conditions. The results indicated that when the financial advisor appeared close, participants talked more, more often and made more interruptions than participants who interacted with the advisor whose image appeared further away. These results lead to the conclusion that individuals were generally more interactive when communicating with a remote collaborator who appears closer to them.

This study indicated that differences in perceived distance could affect the manner in which individuals interact in a video-mediated interaction. This study had shown that differences in perceived distance could lead to structural differences in a dialogue with broad similarities to the effects found in face-to-face studies. The next aim was to determine whether these dialogues differed in terms of their communicative content. To address this issue a detailed system of dialogue analysis was applied to all the recorded interactions. This is described in Chapter 5 .

The analysis of the content of the dialogues sought to answer the following question: – What differences if any are there between the dialogues in each of the conditions of Study 2 in terms of communicative content? The analysis showed that there were indeed systematic differences in the dialogue content between the experimental conditions. When a participant was interacting with a remote interlocutor who appeared close, the participants tended to provide more conversational feedback than participants who were interacting with a financial advisor who appeared further away. The greater feedback was characterised by dialogues containing many more acknowledgements – typically short utterances such as "yes", "I understand" or "okay".

These results suggest that when the advisor appears close the customer has a different conversational style than when the advisor is far away. This style is defined primarily by a tendency to be make more verbal acknowledgements and be very interactive. When the advisor appears far away dialogues are much more "lecture-style" - where the customer tends to listen quietly without saying anything and only speaking when spoken to.

1.3 Which Factors Control Perceived Distance?

To this point, the thesis has demonstrated that changes in camera zoom can lead to differences in individuals' impressions of distance of a remote individual. It has also shown that these differences can lead to behavioural effects that affect interactions supported by a video link. The next question that was investigated is which factors are the most dominant in controlling the impressions of perceived distance?

Changing the zoom on the camera changes more than one perceptual factor. The physical size of the object image is changed, as is the amount of background, and the amount of the object image that is visible. Chapter 6 outlines an experiment that aimed to understand which factors were the most important in manipulating impressions of distance.

Study 3 involved participants studying a photograph and saying how far from the camera they thought the subject of the photograph was sitting. Each participant saw 96 photographs which differed in terms of background (object filled or blank), size (small, medium or large) and distance (subject sitting at 2.5, 4, 6, or 9 feet from the camera). The results showed that the presence of objects in the background made an individual appear closer than in reality, as did an increase in the size of the photograph. Small photographs made people appear farther away than they were.

These results show that the larger an image is, the closer it appears – regardless of the object to background ratio. Since people consider larger images closer and smaller images farther away, when the size of the video window was larger (and consequently the size of all images within the window) less zoom was required on the camera to give the impression of a particular distance.

It appears from these results that the most important factor in causing impressions of distance in the earlier studies had been the physical image size - as there was no change in the presence of objects in the background in the previous studies.

1.4 Does Image Size Regulate Perceived Distance in Interactive Situations?

Study 3 showed that image size was the key factor in determining impressions of distance in photographs. The next experiment sought to investigate whether or not this was also true in an interactive situation.

Study 4 (reported in Chapter 7) repeated the financial advice task used in Study 2 but extended it to include two more conditions than the original study. A projector screen was used to compare the effect of life-size images against those on a monitor screen. This allowed the study of the effects of image size and camera zoom independently.

Somewhat disappointingly, the results failed to replicate all the findings of the original financial advice task. One finding that was replicated was that participants who interacted with an advisor who appeared close again interrupted their collaborator more often than those for whom the advisor appeared far away. The results failed to replicate the finding that individuals speaking with an advisor who appears close make more speaking turns and words than those who speak to an advisor who appears further away. The failure to replicate effects in this study was most likely to be due to a different confederate, who used a different speaking style to the original financial advisor.

Despite these problems there were several patterns that suggested that, image size is a significant factor in a perceived distance in an interactive situation. The study also examined several other aspects of communication, which had not been analysed in the original experiment. These included analysing participants' body position, understanding, and a more detailed analysis of some of the social factors involved in the interaction such as persuasion.

It is known from the literature that taking part in a conversation leads to greater understanding than simply listening in to a conversation. This account was extended to investigate whether the level of interactivity within a conversation - how many times each individual spoke – would affect participants' level of understanding as measured by a recall task. The results of this analysis showed that those participants who spoke often during the interaction scored significantly higher on the recall task than those participants who spoke less. Although in Study 4 there was no significant effect of proximity on the number of speaking turns made in a interaction, the results tentatively suggest that in Study 2, participants who interacted with a financial advisor who appeared close, may have had a greater understanding of the concepts discussed in the task than those who interacted with the financial advisor appearing far away.

1.5 Why Does Perceived Proximity Cause These Effects?

It has been shown that changing the zoom on the camera can affect the "virtual distance" of an individual in a video-mediated interaction, and that this consequently can affect the behaviour of individuals in an interaction. Next the research explored possible causes of the proximity effect on communication.

From the literature there appeared to be a similarity between the effects of close proximity and familiarity. This is apparent both in studies of the face-to-face literature and video-mediated interactions. Familiarity has been shown to lead to longer dialogues, which contain more turns, shorter turns and fewer interruptions. In this thesis close proximity (whether face-to-face or video-mediated) has been shown to lead to longer dialogues, which contain more turns, shorter turns and more interruptions.

Does proximity lead to feelings of familiarity? Do interactions where a remote interlocutor appears close show similar patterns of dialogue content as those between familiar interlocutors? To address this question, Study 5 (reported in Chapter 8) compared the content of the dialogues coded in the proximity study (Study 2) with the content of another set of dialogues of which familiarity was the independent variable.

The results demonstrated a similarity in the pattern of particular types of conversational move between conditions of close proximity and of familiarity. Specifically, the results suggested that both familiarity and impressions of close proximity have been shown to lead to an increase in the richness of the feedback present in participants' communication.

Familiarity and proximity are known to be linked psychologically, as it is known that when familiar individuals interact, they generally do so at a closer distance than when they are unfamiliar. Also, when people come closer than usually appropriate, one of the behaviours seen is that individuals tend to change their body position to re-establish a comfortable space. If that space cannot be re-established then they tend to interact in a similar fashion to the way they would if familiar. Study 5 has provided some data at least suggesting a possible link between perceived distance and familiarity. It is possible that the reason that an individual appearing close affects video-mediated interactions is due to an impression of perceived familiarity.

1.6 The Main Contributions of this Thesis

This thesis contains a number of studies, applying a wide range of methodologies to explore the role of perceived distance in interactions supported by video conferencing technologies. The thesis has shown that impressions of proximity can exist in a video-mediated interaction, and that when they do, they can have systematic effects on the way in which individuals interact with one another. In light of this knowledge, the thesis goes on to explore the perceptual factors underpinning impressions of perceived distance, and finally attempts to gain an understanding of why such effects may occur.

It is hoped that the work in this thesis will be of relevance on both an empirical and a practical level. Potential implications of the work and possible future directions of this type of research are discussed in the Chapter 9.

2 Chapter 2: Literature Review

2.1 Introduction

As telecommunications technology continues to advance, so do the opportunities for interpersonal communication. Organisations are becoming more global, co-workers are less frequently co-located, and multimedia technology such as audio and video conferencing is being used to facilitate remote communication in both a social and a work context. One of the challenges this presents to researchers is to understand the impact of such technology on the communication of those who use it. Does technology provide a forum for successful communication? How do people adapt their communication when using technology? Does technology affect the way in which people interact?

To answer these questions requires an understanding of two diverse areas of interest. Technology for remote communication comes in many forms, encompassing technologies as diverse as email and virtual reality. Interpersonal communication is also an extremely wide ranging field, which includes many different actions such as speech, looking, body position, gestures and so on. Because of this diversity, it is necessary to take a bottom-up approach to the study of remote communication. Such an approach will focus on the attempt to isolate how particular types of communication can be supported remotely and conversely, how particular differences within remote technology may affect communication.

The aim of this thesis is to investigate one aspect of communication and its position with relation to video-mediated communication. Proximity – the physical distance between interlocutors - is a form of non-verbal communication information which is present in all interactions between individuals who are co-located. Changes in proximity have been shown to affect both the behaviour and attitudes of individuals involved in co-present interactions (Hall, 1966; Capella, 1994). Video-mediated communication refers to interpersonal communication in which individuals can both see and hear each other whilst interacting remotely in real time. The most common example of this is video conferencing. While this is one of the forms of remote communication richest in communication information, it is unknown whether impressions of proximity can be preserved using such a medium.

The present research seeks to find out whether the visual representations of remote individuals in video-mediated interactions may contain cues which guide perceptions of proximity. If perceptions of proximity do exist in such a situation, then the research aims to understand how they are governed, and what impact if any they are likely to have on the way individuals communicate in a video-mediated interaction.

This chapter will review literature of relevance to the study of proximity in video-mediated communication. In doing so, the review will be divided into five main areas. This chapter will begin with a review of communication in general, defining what will be meant by communication during the rest of this thesis. The review will continue to define and discuss remote communication and the technologies which have been developed to support it. This will be followed by a discussion of video-mediated communication and studies which have compared the use of video with other communication media, most notably audio-only and face-to-face communication. This chapter shall then continue with a discussion of the variables which combine to create the video-mediated communication experience. Finally, this chapter shall introduce the concept of proximity as a form of communication information and explain why it may be of interest to study with reference to video-mediated communication.

Before going into detail as to why the study of proximity in video-mediated communication may be important, it is worth examining what is meant by the terms "communication" and "remote communication" and why they are of interest in psychology.

2.2 Studying Communication

The study of human communication is a large and diverse field. Researchers from many disciplines are interested in human communication, from many different perspectives. Consequently communication itself is a term which has come to mean many things to many people. For this reason it is worth spending some time defining what will be meant by the term "communication" in the present thesis.

2.2.1 Defining Communication

Communication itself is a rather generic term, primarily due to the diversity of contexts in which it occurs. Communication can exist between computers, biological cells, animals, individual humans and entire cultures as well as many combinations of the above. What constitutes communication in each situation depends upon the subject of interest. This thesis is concerned with communication solely in terms of how individuals communicate with other human beings in given situations – interpersonal communication.

Even when considering only interpersonal communication, many definitions exist, all of which differ according to the context for which they are employed (Krauss & Fussell, 1996). Definitions tend to fall into two main categories however communication as an outcome, and communication as a process.

2.2.1.1 Communication as an Outcome

The dictionary definition of communication is:

> "conveying one's meaning to others" (MacDonald, 1972)

This and many other definitions of communication include an insistence upon the successful transfer of information. The definition also requires that the receiver gains information that is in some sense correct. In this view, communication can be considered to be an object.

This definition is unsatisfactory for the present type of research, due to its narrowness. Considering communication in this way fails to take account of how meaning was conveyed. Differences in communication are defined simply as differences in the meaning being conveyed. When studying human communication on a practical level however, researchers tend to be interested in more than just what meaning is conveyed, but how, why, and whether or not it was intentional.

So a rather broader definition of communication is sought. A definition which includes not simply the conveyance of meaning, but also the manner of communication and the intention. An alternative view of communication is to consider not simply the outcome, but the process by which meaning is transferred.

2.2.1.2 *Communication as a Process*

Shannon & Weaver (1949) developed a model that was originally used by electrical engineers studying the carrying capacity of telephone lines. Their model consisted of a source, transmitter, channel, receiver, and destination. Communication is depicted as the flow of information between a source and a destination, focusing on the success with which information is imparted -- the physical passage of information from point-to-point or person-to-person.

In this approach communication is viewed as a process, as opposed to an outcome. The process itself focuses on the success with which a source or sender exercises control over a receiver or audience. Consequently, this definition shifts the emphasis away from being content based and towards a greater consideration of communication as an activity.

More recent researchers have continued to view communication as a process but have defined it somewhat less rigidly. Hovland, Janis & Kelley (1959) proposed the following definition:

> "[Communication is] The process by which an individual (the communicator), transmits stimuli (usually verbal), to modify the behaviour of other individuals." (Hovland et al., 1959).

This definition clearly states that communication should be considered as a process rather than an outcome. It also however, explicitly states a need for behaviour change to have taken place for communication to have been successful. The definition also retains a very definite emphasis on the flow of communication - as in the Shannon & Weaver model, the analyst is encouraged to think in terms of a transmitter and a receiver.

A more recent version of a process model of communication is provided by Krauss & Fussell (1990). They offer the following definition of communication:

> "[Communication is] the process by which knowledge that resides in one or more people comes to be represented in one or more others" (Krauss & Fussell, 1990)

At first glance this definition appears to differ only marginally with that of Hovland et al. (1959). Crucially however, Krauss & Fussell remove the need for behavioural change. In this definition, communication is now considered to have occurred when there is a change in the knowledge state of others. More subtle differences also change the emphasis away from a consideration of intention, by any party involved. In this definition, communication can clearly be either intentional or unintentional.

Process models of communication stand as the dominant viewpoint in communication research. These models are highly analytical -- logical, linear, methodical, process oriented, and descriptive in a third-party or "observational" sense. They lend themselves well to observational analysis of interactions, allowing researchers to identify behaviours representing different forms of communication and quantify their use under different experimental settings.

2.2.2 Criticisms of the process model of communication

> "The whole is something over and above its parts, and not just the sum of them all." -- *Aristotle* (Cited in Corning, 1997)

This process driven model of communication is not without its critics. The above quote by Aristotle captures the essence of the criticisms. Eighmey (1993) warns against viewing communication as merely the sum of its linguistic parts, lest the researcher ignore the inherent character of the communication. Eighmey believes that thoughts, feelings and ideas are important in successful communication and yet are liable to be largely ignored by considering communication simply as a process.

Consequently, Eighmey comes up with his own definition of communication:

> "communication is the conceptualisation, expression, and reception of ideas" (Eighmey, 1993)

The aim of this definition is to emphasise the context in which communication occurs. Eighmey goes on to say:

> "ideas are the very medium in which communication takes place. It is ideas that give form to the information. It is ideas that enable a person to express a

perception, a thought, or a feeling. And, it is sensitivity to ideas that enables members of an audience to find meaning in an expression" (Eighmey, 1993)

Eighmey's argument is a valid one, of course communication does not exist without context. A related point is that interpersonal communication is inextricably linked to culture. Condon and Yousef (1975) state that "we cannot separate culture from communication, for as soon as we start to talk about one we are almost inevitably talking about the other, too" (p.34). This is true because consideration of communication requires an interpretation of meaning, which is tied to culture. Indeed, a definition of culture is extremely similar to a definition of communication:

> "Culture is a technical term used by anthropologists to refer to a system for creating, sending, storing, and processing information developed by human beings, which differentiates them from other life forms" (Hall & Hall, 1990, p.183).

It should be noted that Eighmey was interested in the use of communication as a tool for advertising. His concern was about the application of the analysis of communication to the construction of new communication messages – in the form of advertising. There are some crucial differences between this and research in the present field. The present research will be interested only in the analysis of communication, not the construction of communication. Consideration will be given to how certain situations can affect interpersonal communication. In the field of advertising the aim is to use communication itself to change behaviour. In advertising the aim is to investigate what is effective in a message and use that information to create new messages that it is hoped will be equally effective. The present research will investigate how a given situation will affect the way in which individuals create and interpret messages.

The points raised by Eighmey are important. The assertion that ideas are an integral part of communication is correct. However, it is of little practical application to researchers comparing interpersonal communication within the same context. From an analytical perspective, an observer can only assume the speaker's intention, as opposed to record it. Similarly, measures of understanding may also only be inferred. In an experimental situation however, firstly the contextual information available to the researcher is abundant and can usually be inferred with a fair degree of confidence, and secondly it is usually the same for all individuals studied.

The goals of this thesis are to understand how communication may vary in different situations. To achieve these goals, experimental methods will be used and quantitative data will be sought. The definition of Krauss & Fussell (1990) is the most appropriate for achieving these goals, because of its emphasis on the directly observable. In this thesis the term "communication" shall be used to apply to all of the behaviours which can be interpreted as having meaning. Communication does not have to be intentional, can be either conscious or unconscious and does not require to be "correct" in any way.

The points raised by Eighmey shall not be discarded, however. To truly understand communication requires an appreciation of why individuals derive particular meaning from a particular situation. Consequently, attempts will be made to gain such an understanding.

2.2.3 Types of Communication

In a face-to-face conversation between two or more individuals, there are many ways in which successful communication can be achieved. Broadly, communication can be divided into two forms, verbal and non-verbal communication. The following section shall discuss the main aspects of each of these in turn.

2.2.3.1 Verbal communication

Essentially, verbal communication is anything that an individual says and how they say it. Analyses of verbal communication may be concerned either with how the interaction was structured or the content of what was said.

In a face-to-face environment, verbal communication is a delicate and complex process. Usually only one person is speaking at a time and each time the floor is taken by another person constitutes a speaking turn. The transitions between speakers are generally smooth and effortless (Sacks, Shegloff & Jefferson, 1974). Indeed the delay between one individual and another speaking in normal conversation is often less than the normal reaction time of 150-200 milliseconds

(Walker, 1982). Such short delays indicate the presence of cues to aid the process of knowing when one person will stop speaking and another should start.

Structural analyses of dialogue measure each individual's number of speaking turns, the number of words that they say, how often speakers interrupt each other and so on. The aim of this type of analysis is to compare objectively how the act of verbal communication differs between situations.

An example of the use of studying the structure of dialogue is a study by Boyle, Anderson & Newlands (1994) who investigated, amongst other things, the effect of familiarity on dialogues in a problem solving task. Boyle et al. found that familiar pairs spoke more often, used more words and interrupted each other less often than unfamiliar pairs. This result shows how an analysis of the structure of dialogue can indicate a systematic difference in the way people communicate with each other depending upon their relationship.

Analysis of the content of dialogues is concerned with the purpose of an individual's utterances. Researchers have used several methods to analyse the content of dialogues, something that will be expanded upon in Chapter 6 of this thesis. Analyses of content all seek to measure the amount or the types of things that individuals say. It may be of interest to know for instance the number of questions asked, amount of social talk, how often the same question is clarified and so on. These can be measured to give an impression of the pattern of a dialogue, as well as a guideline to how much is being understood, what the main focus of attention was and for how long etc. As with structural measures, it is possible to compare how different situations may affect the relative composition of dialogues.

An example of the use of examining the content of dialogues is a study by Olson, Olson & Meader (1997). They had participants work in a group to design an automated post office. Groups either worked face-to-face or remotely using either a audio-only or an audio-visual link. As part of their results, Olson et al. recorded how long each group spent completing a particular activity. The results showed that groups in both remote conditions (audio and video) spent significantly longer clarifying what they meant to each other than face-to-face groups. They also showed that audio-only groups spent longer stating and clarifying the issues involved than the video groups. This study showed that although there was no significant difference in the length of time of interactions, or the number of speaking turns between conditions, there were significant differences in the content of the interactions.

2.2.3.2 Non-Verbal communication

Communication can also occur without any speech at all. Communication using non-language signals is known as non-verbal communication. This type of communication may involve virtually any of the senses other than aural.

Gestural information, for instance, can be used either to emphasise things that are being said (for example to communicate the message that something is important) or instead of speaking such as head nods to indicate agreement (iconic gestures). Other non-verbal communication behaviour can be used to indicate information about emotions (facial expressions), attention (gaze, body position), relationships (body position, touching), understanding (gestures), see Argyle (1988) or Kendon (1990).

It is worth noting that much of non-verbal communication may not be consciously transmitted or interpreted. There is a proverb, "The eyes are more eloquent than the lips" (Okada et al., 1994) which hints at how facial expression can betray information without the need for a conscious decision. As far as the analysis of communication is concerned however, any behaviour which provides a communicative function is of relevance to the communicative process, whether conscious or not.

In a face-to-face situation, it is the combination of verbal and non-verbal communication that leads to a smooth overall communication process. Efficient verbal turn-taking for instance is the result of a combination of several communicative cues including gaze, speech inflection, gestures and grammatical structures (Argyle, 1988, p115-116). When communicating using another medium, individuals are required to adapt their communication strategies to suit whatever communication channels are available. The success of their communication may depend largely upon their ability to adapt to suit the environment.

2.2.4 Measuring Communicative Success

To measure the impact of a given situation on communication, as well as comparing the communication itself, it is possible to study the results of the communication in the form of task outcomes. While the aim of studying communication directly is to identify differences in patterns, styles or relative frequencies of specific methods of communication, the aim of outcome measures is to examine the impact of a given situation on the relative success of communication.

Measures of task performance will inevitably be dependent on the aim of the task but could include such measures as time taken, quality of result, happiness of participants, etc. Some researchers have used problem solving tasks with strongly defined outcome measures in order to provide a quantifiable measure of communicative success (Anderson et al., 1997).

An example of a task with a well defined outcome measure which has been used extensively is the Map Task (Brown, Anderson, Yule & Shilcock, 1984). The Map Task is a collaborative problem solving task which is designed as a dyadic interaction. Two participants aim to complete a map route as accurately as possible. Both participant have schematic maps which refer to the same location. One participant is given the role of Information Giver and has a route marked on their map. Their task is to convey the route to the other participant without physically showing them. There is a further difference between the maps in that they do not have all the same landmarks marked on their maps. The aim of the task is for the Instruction Follower to replicate the route marked on the map of the Instruction Giver without seeing it. This leads to a well defined outcome measure as the route drawn by the Instruction Follower can be compared to the one described by the Instruction Giver, and the deviation measured. This gives an objective measure of the success of the communication throughout the task.

The Map Task has been used by several researchers to compare the success of communication in a variety of different contexts. Groups of participants investigated include children (Doherty-Sneddon, 1995), aphasic adults (Anderson et al., 1997) and non-native speakers of English (Veinott, Olson, Olson, & Fu, 1999).

Communication is a delicate process, and to study it in detail, it is best to adopt a multidimensional approach (Monk, McCarthy, Watts & Daly-Jones, 1996). To do this, experiments will be designed which will include measures of the structure and content of verbal communication, non-verbal communication, and objective measures of the success of communication. It is only with a multidimensional approach that it will be possible to gain a full understanding of the impact upon communication of any technological variables studied.

2.3 Remote Communication

So far the discussion of communication has been limited to face-to-face interactions. A result of recent advances in technology however, has been an increase in the availability of methods of communicating with individuals who are not located face-to-face. The definition of remote communication is the same as that for communication in general, with the exception that the individuals involved are not located in the same place at the same time. Similarly, research into remote communication shares the same goals as for co-present communication in general. Research aims to enlighten how technology may impact upon the human communication process. With this understanding, communication research can aid the development process of technologies as well as understanding when remote communication may be appropriate.

2.3.1 Technology for Remote Communication

Remote communication in the form of the telephone has already made an enormous impact upon our society; new technologies such as e-mail, the World Wide Web, audio and video conferencing may have the potential to do the same. Remote communication technology provides people with the ability to see, hear and exchange written information with others, all over the globe.

Many technologies presently exist to facilitate remote communication. They can be broadly categorised according to the channels of communication which they support. Channels of communication refer to the major sources of information which can be used to communicate (visual, audio, textual etc.). The major differences which define different technologies include whether the communication they support is synchronous or asynchronous, if they support group communication or merely person to person communication, and whether or not there is an audio channel, a visual channel, and a data channel present. Table 2.1 gives a list of examples of types of remote communication technologies that exist, along with the channels of communication that they generally support.

Technology	Communication Channels				
	Synchronous?	Data	Audio	Visual	Group?
Bulletin Boards / Email	No	Yes	No	No	Yes
Instant messaging	Yes	Yes	No	No	No
Telephone	Yes	No	Yes	No	No
Audio Conferencing	Yes	Possibly	Yes	No	Yes
Video Conferencing	Yes	Possibly	Yes	Yes	Yes

Table 2.1 Examples of Remote Communication technologies.

Almost any face-to-face situation that could be conducted remotely offers an opportunity for savings in time, and in many situations cost. Benefits to using remote communication technology can come in the form of efficiency savings for the user, such as making the completion of some task faster or cheaper, or subjective benefits such as user enjoyment. Without any appreciable benefit however, a technology is unlikely to become successful (see for example Hewson, Allison & Clarke, 1995; Kraut, 1996). In a work context, organisations are becoming more global, colleagues are often no longer to be found down the hall. Travelling to meetings is both expensive and time consuming, so companies are turning to remote communication technologies as a potential solution (Veinott, Olson, Olson & Fu, 1999). Similar benefits can be imagined in a social context, with worldwide travel on the increase, so staying in touch is considered important. Anything that may enrich the experience of communicating with someone with whom you cannot be with would be seen as a benefit.

2.3.2 Why Study Remote Communication?

These applications of remote communciation provide the context in which research in this field exists. Just as the potential applications of remote commnication are vast, so the potential alternatives are diverse. Research aims to identify the factors within technology which may be of relative importance for potential applications, by illuminating the probable impact each may have on the communication processes of users in a given situation.

Suppose for example, that a remote communication system is designed with the aim of providing a more efficient method of conducting brainstorming meetings. This hypothetical system would connect three rooms by an audio-visual link, thus allowing participants to take part in a meeting without all having to travel to the same destination. It would be useful in such a circumstance to know whether the technology led to an increase or a decrease in the amount of time that a typical meeting may take; the quality of the results of the brainstorming; how happy people are with the results; how easy people found it to make their contributions; how much co-present individuals talked in comparison to remote individuals; how well each individual understood the ideas presented by the others; and so on. When evaluating the effectiveness of the new technology all of these factors may be important and would need to be taken into account and compared with any alternative method (in this case face-to-face meetings).

The impact of a technology may manifest itself in physical differences such as the time taken to complete a task, or it may exist in more subjective effects, such as the extent to which people enjoy using the system, trust others they speak to, or how they behave whilst using one technology as opposed to an alternative method. If the potential impacts on communication of a particular technology are known then this information can be used to aid in the design process, to ensure that new products successfully achieve their goals.

Due to the diversity of remote communication technologies, it is unrealistic to study "remote communication" per se. Rather, communication must be studied within a particular context. Consequently, the present thesis is concerned only with video-mediated communication. This was chosen for two reasons, firstly it is a technology which has been around for a few years but yet has failed to meet early expectations of uptake (Tang & Isaacs, 1993) and secondly because it is the form of remote communication technology which should most closely resemble face-to-face interactions.

2.4 Video-mediated communication

The term video-mediated communication refers to synchronous communication conducted remotely, using both the visual and the audio channels. There may or may not be other shared information available to the participants. It is the presence of both the visual and audio channels that makes video-mediated communication potentially the richest form of remote communication available. It is worth noting that video-mediated communication differs from the term "multimedia". "Multimedia" requires the use of a data channel as well as audio and visual (Fussel & Benimoff, 1995), thus video-mediated communication is a subset of multimedia communication.

The most common form of video-mediated communication is videoconferencing. A distinction is often made between desktop systems which are based upon personal computing technology, and 'room-based' videoconferencing systems. This difference implies a difference between personal, office-based equipment which is PC-based and has the above sharable features, and facilities which have many users and have to be scheduled and booked in advance. The advantage of the latter is usually to have a more 'studio' like environment with higher quality equipment - higher specification, a dedicated facility, larger screens, peripherals, lighting adjustments and technical support (Gornall, Pengelly, & Shearn, 1999).

Cunningham (1996) points out that videoconferencing can be used in various modes.

> *One-to-one meetings* - often known as point-to-point; this provides full two-way (duplex) communication, with audio/data and video. Desktop videoconferencing can include file sharing, interactive whiteboard, use of a remote 'OHP' via document camera and so on
>
> *Many-to-many conferencing* - multi-point communication. Some systems show only one site on screen at a time : the camera switching is voice activated, so the current speaker is displayed. 'Continuous presence' software

allows all sites to be seen co-present on each screen up to a limit often of 4-5. Participants may also join in on audio only.

One-to-many communications - such as the use of videoconferencing to 'broadcast' a talk or presentation from a main site to other sites.

These different modes for using videoconferencing reflect the diversity of the system. It shows that a variety of different types of video communications technologies have been developed which aim to support different kinds of communication. It also illustrates that although video-mediated communication technology exists to facilitate communication, that communication will always be constrained by the configuration of the technology.

2.4.1 Applications of video-mediated communication systems

Despite the fact that video-mediated communication is still an emerging technology, it has been trialed in a diverse set of situations. There has been a steady increase in the number of users of multimedia applications in general over the last few years (Watson & Sasse, 1998). To emphasise the wide ranging extent of the context involved in remote communication it is worth mentioning some of the applications for which it has been implemented.

2.4.1.1 Telemedicine

One of the earlest applications of video mediated communication technology is in the field of public health, as an aspect of telemedicine. In an attempt to clarify what was meant by the term telemedicine, a committee of the Institute of Medicine in the USA defined telemedicine as "the use of electronic information and communications technologies to provide and support health care when distance separates the participants" (Field, 1996). Within this context, video conferencing is one solution for telemedicine, rather than being synonymous with the term (Grigsby & Sanders, 1998).

One area in which video has been used to provide telemedicine is in mental health. The problems of distance and the delivery of mental health services have been studied for over 50 years (Baer, et al., 1997; Shannon, Bashshur, Lovett, 1986). An early experiment with video links resulted in a live one-way and later two-way connection between a psychiatric unit in Nebraska and a state mental hospital. The connection allowed university psychiatrists to perform neurological assessments, provide consultation, and education (Wittson & Benschoter, 1972). Since this time telehealth has made significant strides into virtually all areas of health services (Fields, 1996; Baer et al., 1997). However, only in the current decade have health services via distance technologies become both prevalent and financially feasible (Baer, et al., 1997; Dakins & Jones, 1996). Recently the field has experienced a rapid explosion. It was estimated that by the end of 1996 almost 30% of rural hospitals in the USA would be using some form of telehealth (Nickelson, 1996).

There are two ways in which video is likely to be particularly useful. Firstly, the ability to connect remote hospitals to specialist centres, enabling the possibility to obtain specialist opinion on patients, scans, test results and so on without the need for travel (Furnace, Hamilton, Helms & Duguid, 1996). This is being piloted in hospitals in the North of Scotland but information as to its success is not presently available. Studies comparing psychiatric assessment by 2-way video as opposed to face-to-face have been largely positive (Ball, Scott, McLaren, & Watson, 1993; Schneider, 1999).

An alternative application for video in the medical sector uses the video channel to display images of an operation or diagnosis so that an entire medical team may be able to view the case. Nardi, Kuchinsky, Whittaker, Leichner, & Schwarz (1997) studied this type of use for neurosurical operations. The aim of the system being trialed was to allow the neurophysiologists to monitor operations remotely, thus allowing them to service a greater patient population, and spread their expertise over a larger number of operations. This application of video-mediated communication was shown to be successful but rather different to the more common interpersonal uses of video communication.

2.4.1.2 Education

Just as health care can be dispensed remotely, so too can education. Video-mediated communication systems are currently being used in a number of distance learning settings.

Gornall, Pengelly, & Shearn (1999) report a case study of a telematics project called 'Alps' (Adult Learner Partnerships) which featured the use of PC-based videoconferencing systems. It was located in South Wales, and involved the piloting of open learning, self-paced materials for basic maths study, to be used at community centres and combined with interactive 'distance' tutorials between the centres and the University.

They found the project to be a success, and reported that the University continued to use the system after the end of the project. They found that younger age groups found the medium more attractive to use than adults, but that given extensive training most people were happy to use the system.

A less interactive use of videoconferencing in the educational sector is reported by Knowles and Dillon (1996). They detail an example where satellite video-conferencing was used successfully to inform American architects about new legal requirements in their profession. Nineteen thousand people attended one of three videoconferences explaining the implications of the "Americans with Disabilities Act".

2.4.1.3 Speech Therapy

An application which blends health with education is the field of speech therapy. Rose, et al. (1999) present a descriptive overview of the TeachSpeech experiment which combined both tele-health and tele-education over a three year period. The project set out to assess whether videoconferencing technologies could be used to link a speech and language therapist (SLT) with children in mainstream schools who needed the services of the therapist. Supporting children locally in the schools was an Educational Support Assistant (ESA) who had the responsibility of managing the daily therapy program in consultation with the professional SLT. At the time the article was written, the program included eleven children in four different schools.

Because not enough resources have been available, SLTs typically visit local schools only once or twice per term. When modifications in therapy were needed as a child's program progressed, the needed changes often were not identified and implemented as quickly as desired due to the long time between SLT visits at the schools. The videoconferencing capability was seen as a possible way to deliver these valuable services in a more timely way to more schools.

Rose et al. report that the Study shows that professionals, who have had little or no prior experience in the use of multimedia technologies, quickly become accustomed to using these advanced multimedia systems on a day-to-day basis. They also suggest that the data retrieved from parents and the SLT and ESAs was overwhelmingly positive regarding the experience and the outcomes.

2.4.1.4 Banking

Another large potential implementation is in the banking sector. Banking providers are considering implementing greater accessibility to self-service banking, one possible feature would include a human expert available on-line to help customers and to deal with any difficulties. This video-mediated representation would be attempting to create a greater feeling of human presence than an audio-only line, as presently used with telephone banking.

2.4.2 Studies Comparing Video-Mediated Communication With Other Media

Although video-mediated communication is still an emerging technology, several studies have investigated its use in a number of situations. Up until now the implicit aim of video-mediated communication systems has been to create remote interactions which are as close as possible to a face-to-face style interaction (Sellen, 1995). Because of this, most of these studies have compared video with other modes of communication, primarily audio only and face-to-face. Due to the fact that communication is so diverse, these studies have investigated different aspects of communicative style. These studies can be categorised according to the aim of the research. These studies will be discussed based on whether they investigated how video-mediated communication technology affected participants' dialogue, task performance or subjective impressions of using the technology.

2.4.2.1 The Effects of Video-Mediated Communication on Dialogue

Research has shown that the way people talk to each other on the telephone is different to the equivalent in a face-to-face situation. Studies have compared the structure of dialogues in experimental situations, and have found that when participants were communicating by audio alone, turn taking tended to be much more formal, with longer turns, longer pauses between turns and more explicit handovers (Rutter & Stephenson, 1977; Cohen, 1982; O'Connail et al., 1993).

Anderson, Bard, Sotillo, Doherty-Sneddon & Newlands (1994) compared the intelligibility of communication when participants were completing the Map Task in either an audio-only or a face-to-face condition. They reported that individuals speak more intelligibly when there is no visual channel available. These results are hypothesised to be due to the lack of the visual channel, thus making gaze and gesture both ineffectual turn taking cues, forcing individuals to rely only on verbal signals. The increasing pressure on the verbal cues to turn-taking may lead to an increase in the clarity of those signals.

It had been expected that adding high bandwidth video to existing audio would produce more social cues and consequently increase social presence (e.g. Short et al. 1976; Rutter, 1987). This in turn was expected to lead to video-mediated interactions that were more similar to face-to-face than audio-only conversations. In actual fact video-mediated communication has been shown to bear a greater similarity to audio-only interactions than face-to-face. Researchers have found that video-mediated interactions lasted as long as audio-only, longer than face-to-face and contained more interruptions than either face-to-face or audio-only (Jameson, Hobsley, O'Hanlon & Buckton, 1996; Anderson, Newlands, Mullin, Fleming, Doherty-Sneddon, & Van der Velden, 1996). O'Connail (1994) concluded that desktop conferencing appears to be more like the telephone than face-to-face, with conversations being briefer and more task focused. Turn taking does not appear to be greatly improved in video-mediated communication, compared with audio only, studies have shown that the addition of a video channel to mediated communication does not lead to any greater fluency in turn taking (Anderson et al. 1996; Doherty-Sneddon et al. 1997).

When there is an option to speak or not, participants may be less likely to speak when they are in a remote situation than in a face-to-face situation. Furnace, Hamilton, Helms & Duguid (1996) assessed the potential of videoconferencing as a medium for teaching students remotely. The study compared tutorials with medical students conducted either face-to-face in a teaching room at a hospital in Aberdeen or via videoconferencing between Aberdeen and a paediatric unit in Inverness (171 km away). They found that when tutorials were conducted remotely, students were less likely to speak and get involved than if they were face-to-face.

It is not just the structure of dialogue that has been studied with respect to communicating across different media. Olson, Olson & Meader (1994) compared the content of dialogue during face-to-face, video-mediated and audio only interactions where participants completed a problem-solving task. They found that there was a difference in the amount of verbal checking in the dialogues with more occurring in the remote interactions, particularly in the audio only situation. Audio only conversations were also rated by participants as having a lower discussion quality. These results suggest that although the video channel led to a greater quality of discussion, the fluency of the video-mediated interaction was more like the audio-only interaction than a face-to-face interaction.

2.4.2.2 The Effects of Video-Mediated Communication on Task Performance

Many of the studies comparing communication between different media have used tasks containing a measure of performance. Such studies have been concerned with measuring how successful communication is in different situations.

The most commonly used outcome measure is performance on problem solving tasks, such as the Map Task (Brown et al., 1984) which was described earlier. Several researchers have studied the differences between performance on problem solving tasks between face-to-face, audio-only and video-mediated communication systems. Virtually all of these studies have shown that there is no significant difference between participants performance using any of these conditions. Tasks have included the Map Task (e.g. Doherty-Sneddon *et al.* 1997); designing an automated post office (Olson et al., 1997); locating the nearest physician to a given address (Chapanis *et al.* 1972) and scheduling on a hypothetical college timetable (Weeks & Chapanis, 1976).

There is however, one study which did show a performance advantage on a problem solving task for video-mediated communication over audio-only communication. A study by Veinott, Olson, Olson & Fu (1997) showed that when speakers who were not native speakers of English completed the Map Task in English, performance was significantly better in the video-mediated condition than in an audio-only condition. When native speakers were interacting, there was no advantage for video over audio.

A related finding is that by Reisberg, McLean & Goldberg (1987) who found that English speaking subjects performed a shadowing task in a foreign language better if they could see the speaker. This suggests the ability to lip-read may be important for non-native speakers and video may be useful to non-native speakers as a result.

Other types of task which have been used to compare performance, have attempted to create a more social element to the task. Morley & Stephenson (1969) asked pairs of subjects to take part in a simulated industrial dispute using either the telephone or in a face-to-face situation. Each participant was given arguments with which to argue their case, of which one was stronger than the other. Their results showed that if bargaining was done without vision, the individual with the stronger argument would tend to win. In a face-to-face situation however, interpersonal processes such as liking and wanting to be liked interfered and the case was less likely to be resolved on the merits of the case.

Following this, Short, Williams & Christie (1976) conducted a similar task but comparing video-mediated communication with audio-only and face-to-face. This task involved one participant playing the role of a manager and one of a union representative. The two participants had to negotiate a deal, and each option had different payoffs for each individual. The results showed that if the participant was supporting a view at conflict to their own view then they performed better using the telephone than face-to-face. If they were arguing in accordance with their own views then they did better in the face-to-face condition. In opposition to the majority of studies comparing video and face-to-face communication however, results in the video condition more closely resembled those conducted face-to-face than the audio-only condition. Results also showed video to facilitate conflict resolution, by being more personalised, less argumentative and more polite (for a review see Whittaker, 1995).

The research outlined above shows that certain types of task are sensitive to differences in communication medium. Although in problem solving tasks, there appears little difference between the success of participants using audio only or video-mediated communication, visual contact may be more important when the task involves some conflict or negotiation, or if communication is stressed, as in the case of speakers not sharing a common first language.

2.4.2.3 The Subjective Effects of Video-Mediated Communication

Evaluations of technology usually concentrate on efficiency and effectiveness, as is the case in studies of performance. For some applications however, particularly those aiming at an informal, home environment, social presence and satisfaction may be more appropriate criteria for evaluation (de Greef & Ijsselsteijn, 2000). To this end there have been some studies which have investigated the subjective impressions of participants communicating using different media.

It is known that users tend to rate an audio-only interaction as having less social presence, and being more socially distant than a face-to-face interaction (Argyle, 1988, p119). Studies comparing how video-mediated communication compares to that of face-to-face and audio-only suggest that video is somewhere in between the two (Short et al., 1976). Gale (1990) conducted an experiment in which participants had to collaborate to complete three separate tasks over a time period of three weeks. Results showed that the participants ratings of the 'social presence' of the others in the group increased with the communication bandwidth. Thus video was considered to lead to a greater sense of social presence than audio only communication. Gale (1990) also discovered that participants impressions of their productivity increased as communication bandwidth increased, despite no significant difference in the quality of the output or the time taken to complete the tasks between video and audio conditions. This shows that although users may not be performing any better they may think they are.

Just as people have been shown to prefer those people they meet face-to-face as opposed to those they only spoke to on the telephone, so groups conversing with each other by video were also found to like each other more than those negotiating by audio alone (Short et al. 1976). These results have been attributed to the transmission of affective cues e.g. facial expressions, posture and gesture (Whittaker, 1995; Whittaker & O'Connail, 1997). Sellen (1995), using a debate situation reported a similar finding. She found that although dialogue analysis showed no objective differences between audio only and video-mediated conversations in terms of the way participants spoke when completing the task, there were subjective differences. Most individuals tended to prefer the video-mediated communication rather than audio. The benefits of video reported by participants in this study included providing non-verbal feedback; identifying who was talking; increasing the social presence of the group and increasing the interactivity of the conversation.

Although users may feel more positively about video-mediated rather than audio-only communication, it is still not equivalent to face-to-face communication. Storck & Sproull (1995) found that people who interacted remotely by video form less positive impressions of each other than those who interacted face-to-face. In forming these impressions, individuals were also shown to be taking account of different kinds of information. Liking has been shown to affect how individuals rate the performance of others on a task. Weisband & Atwater (1999) found that individuals when asked to rate the contribution of other members of the group on a problem solving task, liking biased the reports of those in face-to-face groups but not in electronically mediated groups.

In general, research has suggested that users are more positive about video-mediated communication than audio-only communication (e.g. O'Conaill *et al.*, 1993; Tang & Isaacs, 1993; Isaacs & Tang 1994). Reasons given for this have included feeling able to engage in interaction more freely in VMC (Anderson *et al.*, 1996), feeling less dislocated (Sellen, 1995), and feeling that users can receive some non-verbal feedback from the other user (Tang & Isaacs, 1993). There are also, of course, some negative subjective beliefs, Sellen (1995) found that some users feel more self-conscious when using video communication systems.

Taken together, these results suggest that interacting by video is generally preferred to interacting by audio but is less favoured than face-to-face interaction.

2.4.2.4 The Novelty of New Technology

Some researchers claim that the lack of significant results showing video conferencing as leading to improved performance than audio only situations may be due to the novelty effect and users not being used to the technology (e.g. Anderson *et* *al.*, 1996). Using the telephone requires a different communication skill to face-to-face communication, which is learned (Argyle, 1988, p 118). It should therefore not be unexpected if individuals require to take some time before learning to use a technology most effectively. Research has shown that users can adapt to a technology. Watson & Sasse (1996) found that some subjects' ratings of audio quality increased over time, despite the quality not improving. Rudman & Dykstra-Erikson (1994) showed that students using a desktop video conferencing system became more sensitive to information in the visual channel. Facial expressions were used more effectively and non-verbal signals were substituted for verbal ones.

It is worth noting however, that it is also possible that the positive subjective reports often found in studies (e.g. Gale, 1990; Tang & Isaacs, 1993) may also simply be due to the novelty of the technology. This is particularly relevant in those situations where there may be no performance improvement but there appears to be a subjective advantage for video (e.g. Tang & Isacs, 1993).

2.4.2.5 Summary of Communication Conducted Using Different Media

Although individuals' communicative style differs between face-to-face and audio-only interactions, and even though the way individuals perceive others they talk to on the phone differs from face-to-face, there appears to be no difference in the success of individuals on completing problem solving tasks. This result has been replicated using a number of tasks (e.g. Boyle, Anderson & Newlands, 1994) and shows that although communicative style is different in an audio-only interaction it does not necessarily equate to being less effective.

Taking the results together suggest that when social presence is unimportant for completion of the task then there is no need for video to be added to audio in terms of successful completion of the task (Sellen, 1994). When it is important, video can provide benefits.

A point raised by Monk, McCarthy, Watts & Daly-Jones (1996) was that while subjects may have been able to perform equally well without video, it may have taken greater effort. This in turn may lead to participants preferring video and consequently the subjective improvements shown for video. A study by Watson &

Sasse (1996) has suggested that individuals may indeed find poorer quality video conferencing harder to use than video with a high frame rate and high quality audio.

2.5 Variables That Differ Between Studies Investigating Video-Mediated Communication Technology

The results of these studies give a flavour of how video-mediated communication compares with face-to-face and audio-only communication. As well as comparing how video-mediated communication compares with other media, studies have investigated how different video communication systems may compare with each other.

2.5.1 Physical Variables Within Video Technology

So far video-mediated communication technology has been described as a singular entity. It is worth noting however, that whereas face-to-face and audio-only interactions tend to have little scope for diversity, video-mediated communication technology systems may differ in many ways. These differences may consequently lead to variations in the way in which individuals communicate using the technology. For this reason, much of the more recent research into video-mediated communication takes an intramodal approach as opposed to the intermodal studies discussed so far. In this section, some of the variables of video-mediated communication which have been studied will be outlined.

As with intermodal experiments, the effects of an individual variable on video-mediated communication can be measured in terms of both behavioural effects and subjective effects. Behavioural effects refer to any objectively measurable entity, for instance level of success, errors made, time taken and so on. Subjective appraisals refer to how individuals feel about some element of their interaction using the technology, it could refer to their performance, the equipment they were using or any other element of the task.

2.5.1.1 Quality of Video

The variable that has been studied most extensively with reference to video-mediated communication systems is that of quality. Quality is an issue which concerns both the audio and the video channels. When considering quality of a video-mediated communication system, there are several variables to consider, including packet loss, audio delay, audio-visual synchrony and frame rate. All of these factors are connected to bandwidth such that if bandwidth is limited then one of these will be negatively affected.

2.5.1.1.1 Packet Loss

In multimedia communication, audio information is often sent in packets down a network. In low bandwidth situations, or when a lot of information is being sent concurrently with the audio information, high levels of packet loss can render communication almost impossible (Watson & Sasse, 1997) or at lower levels, conversants may require to check that they and their partners have heard correctly (Zhang, Wolf, Darjaval & Touna, 1998). This can lead to individuals taking longer to achieve grounding and consequently increase the task completion time (Jameson et al., 1996).

2.5.1.1.2 Audio delay

Audio delay refers to the time taken between an individual speaking and the audio being heard at the other end.

A study by Anderson et al. (1997) compared participants' performance on the Map Task when communicating using different media. There were two independent variables - delay versus no delay, and audio-only versus audio + video. The delay was approximately 500 msec. The results showed that participants who experienced a delay performed less well, with performance scores on average 36% poorer than those who experienced no delay. The addition of the video channel did not lead to any improvement in performance compared to those who only used the audio channel. Audio delay also led to a significant increase in the percentage of turns that were interrupted, with over 50% of a speakers turns being interrupted by the listener in the condition with delay, compared to 15% without.

Other studies have not used audio lag as an independent variable but also indicate audio lag to lead to negative effects. A study by Cohen (1982) compared the communication processes of face-to-face communication with low quality video conferencing with an audio delay of 705msec on a series of lab tasks. Participants found it hard to switch speakers and to ask clarifying questions. Participants made twice as many turns in the mediated condition as in the face-to-face situation. This is a greater differential in the number of turns than has been found comparing high quality audio only interactions with face-to-face (e.g. Rutter & Stephenson, 1977).

Other problems which have been encountered experimentally by participants using systems with audio lag include difficulty in turn taking (Tang & Isaacs, 1993; O'Connail, Whittaker & Wilbur, 1993) and asking clarifying questions (Whittaker, 1995). Audio lag is also subjectively unpopular with participants who become frustrated (Isaacs & Tang, 1997) and may negatively influence interpersonal opinions (Angiolillo et al. 1997). Egido (1990) for instance, reported anecdotal evidence of users taking a dislike to others they had only met through using the Picturephone – a system afflicted with a large delay (700ms).

So what length of lag causes disruption to the communicative processes? Tang & Isaacs (1993) indicate that a lag of 220-420ms is acceptable but this figure is disputed by other research (Whittaker, 1995). A delay of 500msec or more has consistently been found to be severely disruptive, impairing subjects ability to complete a task successfully (e.g. O'Conaill, Whittaker & Wilbur, 1993; Tang & Isaacs, 1993; Anderson *et al.,* 1996). For fast turn taking Rudman et al. (1997) report that a delay of 60msec or less is necessary.

2.5.1.1.3 Audio-visual synchrony

The reason that an audio lag may be introduced is to allow the video to be compressed – necessary even at low frame rates if it is to be sent across a network. The lag then keeps the synchrony between the audio and the video. The alternative to introducing this audio lag is to lose the synchrony between the audio and the visual channels. Some researchers have reported that some individuals preferred a system which loses the synchrony of the video and audio but has a low level of lag (Tang and Isaacs, 1993; Kurita et al. 1993).

The only advantage that has been shown for keeping audio-visual synchrony at the expense of a delay is that at high frame rates it may be possible to lip-read – which increases speech intelligibility (Reisberg, McLean & Goldberg, 1987; Summerfield, 1992). If the task is anything other than lip-reading however, this advantage is more than wiped out by the problems of audio lag discussed earlier. Reviewing the studies on audio lag, Whittaker (1995) concluded that introducing video at the expense of audio quality has a negative effect on the interaction processes.

2.5.1.1.4 Frame Rate

An issue of quality which refers solely to the video channel is frame rate. Broadcast quality frame rate is generally considered to be between 25-30 frames per second (fps), however when video signals are relayed over IT networks, there is often a danger that the frame rate be vastly reduced. In video conferencing systems, a frame rate lower than 5 frames per second is not uncommon due to limitations of bandwidth. Several researchers have investigated the effects of low frame rates on various aspects of non-verbal communication.

One of the uses of the visual channel in face-to-face communication is the ability to lipread. Vitkovich & Barber (1994) attempted to investigate at what frame rate video-mediated communication systems could support lipreading adequately. They asked participants to shadow verbal passages, which were presented against an irrelevant background message. They found that the presence of a video image always led to improved performance over a baseline audio-only condition. The results showed that when the video image was 25 Hz performance was significantly improved than when the frame rate was 8.3Hz. Some subjects also improved at 25Hz compared with 12.5Hz, but there was no difference between the performance at 16.7 Hz and 25 Hz. Their conclusion was that if users were to lipread adequately, video mediated communication systems needed to deliver video signals at 16.7fps or more (see also Barber & Laws, 1994).

Frowein, Smooresburg, Pyltes & Schinkel (1991) studied a group of subjects with hearing impairments. Hearing-impaired people should be particularly skilled at lipreading because they use this skill to communicate more than people without such an impairment. In their experiment they found that participants could lipread better if the frame rate was 15 fps as opposed to 5 fps. Above 15 fps there was no extra advantage.

Taken together, these studies both suggest that a frame rate of 15 fps or greater is preferable to aid lipreading. Other researchers have investigated the effects on other aspects of visual information. Barber & Laws (1994) showed that participants were successful at reading emotions at 5 fps. It is worth noting however, that at low frame rates expressions may be missed as the facial expression may happen too quickly to be caught on camera. A study by Bruce (1996) showed that face recognition does not require a high frame rate to be successful. She found that neither the recognition of faces nor the understanding of expression requires a high frame rate, presumably because of the lack of motion involved in either of them. It is, after all, possible to judge expression and face recognition from still photographs (Ekman, 1982; Bruce & Green, 1990), although it is unlikely to be possible to lipread from photographs.

All of the studies so far discussed have investigated the effects of low frame rate on the listener. Blokland & Anderson (1998) instead, investigated the effects of low frame rate on the speaker. They showed that a frame rate of under 5 frames per second causes users to alter their speech production as though they were experiencing communicative difficulty. This was despite having a high quality audio connection. The most likely explanation for this result is that participants in a conversation are experiencing difficulty when seeing their partner speak and consequently alter their own speech to make it more intelligible.

From a subjective point of view, little information is available outlining the lowest frame rate at which users consider video communication acceptable. This is likely to be highly dependent upon both the task and the equipment being used. Rimmel, Hollier & Voelcker (1998) suggest that users may be more tolerant of degraded video when viewing objects other than faces. Tang & Isaacs (1994) reported that people found 5 frames per second acceptable for desktop video conferencing, it seems unlikely that users of a large video conferencing suite would consider a similar rate acceptable. The metaphor which is closest to the experience of users in general is that of television. Most people are used to seeing people on television and thus when they video conference they may implicitly expect a similar quality.

2.5.1.1.5 Summary of Video Quality

The research cited above shows that depending upon the nature of the task a low frame rate may detract from the usefulness of the visual channel. Performance on problem solving tasks would appear to be hampered most by a delay in the audio channel. Subjectively, participants also appear to prefer the loss of synchrony between the audio and visual channels to experiencing an audio lag. Audio visual synchrony may be necessary for lipreading, as is a high frame rate, preferably of 15 fps or higher. When not requiring lipreading, low frame rate has not been shown to have any negative effect on performance in any tasks.

Although the audio and visual channels can be studied separately, research has shown that subjective perceptions of the quality of the audio and video channels are linked. Hollier & Voelcker (1997) showed that participants rated the quality of the audio differently depending on the quality of the accompanying video and vice versa. This was most profound when the image was of an individual speaking. The researchers claim that high quality audio should not be sacrificed in favour of adding video as the negative effects of poorer audio are too great.

2.5.1.2 Image Size

Of the physical variables which exist between different video-mediated systems, one of the most obvious is the size of the video window. In most desktop videoconferencing applications, a video window is used which is typically smaller than a quarter of a screen in size. Some researchers have studied the effects of image size in mediated communication.

In their study investigating lipreading, Barber & Laws (1994) found that a large image size was only necessary for lip reading when frame rates are very slow. They also found that emotion recognition was unaffected by image size. A later study by Anderson *et al.* (1996) compared participants using a large window (6.5 x 8 in.) and a small window (3.5 x 4.5 in.) on a problem-solving task with a social element - the travel game (developed by Anderson *et al.* 1994). Their results showed no significant differences between groups on either objective or subjective measures.

Monk and Watts (1995) found that enlarging the size of the video image led to an increase in the amount of simultaneous speech. Simultaneous speech can be caused by one of two things. Either it is caused by individuals mistiming the end of another's speech (a problem in turn taking), or it is due to individuals taking a turn of speaking before their conversational partner has finished speaking as they are eager to make a point or clarify something. The latter of these definitions is generally considered to be a sign that participants feel more at ease. It is worth noting that other research has shown an increase in simultaneous speech by users of video compared with users of audio (Jameson et al. 1996; Anderson et al. 1996). It could be suggested that increasing the size of the video image from small to large has similar conversational effects to the addition of a visual channel to audio-only communication.

It has been pointed out that reducing the video size from life-size is likely to detract from the usefulness of the visual channel (O'Donnell & Draper, 1997). Although some video-mediated communication systems being developed do use life-size video, none have so far been used for a structured experiment. Anecdotal reports however, do suggest that it may be popular with users. Okada et al. (1994) used the MAJIC system which uses life size video projection for interaction between two people. They reported that participants who used the system at a trade fair were more friendly and outgoing than is usual in video conferencing demonstrations. Participants also reported that sometimes they felt pressure when the image was zoomed in very close, but when students who knew the demonstrators used the system, they reported that they preferred the image close, or even larger than life size. Okada et al. suggest that this is because they wanted to talk at personal distances.

Of the studies cited, there appears to be only hints of a psychological effect of image size. It is possible that when larger images are used, people may feel more comfortable and consequently interrupt more than when the image is small. It seems reasonable to consider image size on a continuum from zero (audio-only) up to life-size.

2.5.1.3 *Camera View*

Another variable which may affect the communication process is the view that the camera should show. There is an inevitable trade-off between the amount one can see and the detail in which it is seen. Researchers have highlighted the need to choose between showing just the head or the head & shoulders vs. upper body (Mantei *et al.* 1991; Heath & Luff, 1991). If the whole upper body is visible then there may be more chance of seeing gestures but less chance to see expressions and other facial information. Research examining the effect of this difference is sparse, but in unpublished studies cited in Doherty-Sneddon *et al.* (1997) no effect of displays showing the users' head versus head and upper body was found.

2.5.1.4 *Eye Contact*

One of the major forms of non-verbal communication is gaze. Short *et al.* (1976) suggested eye gaze had three communicative functions: monitoring to provide feedback when required; regulation of turn taking, and for the expression of affect. Kleinke (1986) studied participants in a face-to-face situation and found that speakers gaze more at a listener's face when they are trying to be more persuasive, deceptive, ingratiating or assertive. One of the major challenges which have been faced by developers of video-mediated communication is how to create eye contact.

In the majority of systems the camera is placed on top or to the side of the monitor. This provides a full-face view but does not permit realistic eye contact. This is because the individual is not looking directly at the camera when they are looking at the image of the other participants. Several researchers have reported that users of these types of video-mediated communication systems have difficulty in effectively using gaze to regulate turn taking (Heath & Luff, 1991; Heath, Luff & Sellen, 1994; O'Malley & Langton, 1994).

There are presently two alternative systems for creating an impression of eye contact. The first involves a neat use of mirrors such that the individual looks at the monitor directly but through a semi-silvered mirror which reflects the image of the face on to a fully silvered mirror which the camera is trained on. Examples of such systems include the Gazecam set-up in Acker & Levitt (1987), and videotunnels in Anderson *et al.* (1997). The second option uses a one-way projection screen and the camera is placed directly behind it – an example of this approach is the MAJIC system (Okada et al., 1997).

Both of these systems purport to being able to provide realistic eye contact, however experimental evidence is inconclusive as to whether either system leads to the use of gaze in communication similar to that in face-to-face. Doherty-Sneddon *et al.* (1997) compared VMC with and without eye contact. They studied participants completing the map task and found that video with eye contact led to significantly more speech being used to gain the same level of task success as in either an audio-only or a video-mediated condition without eye contact. They concluded that this was due to excessive gazing occurring and interfering with cognitive mechanisms. Also, they concluded that individuals gave more redundant information in the video condition and thus that the audio only condition was more efficient.

It should be noted however, that the participants were not instructed to complete the task as efficiently as possible or as fast as possible. Neither did Doherty-Sneddon *et al.* study any subjective effects, so it is unknown whether or not their participants had any particular preferences for either type of link. Indeed other studies have indicated a preference for eye contact (e.g. Acker & Levitt, 1987; Austin, 1996). Acker & Levitt (1987) found that the presence of mutual gaze improved the satisfaction with video conferencing as a medium for negotiation - participants found it to be easier to evaluate their counterparts and participate more confidently in the exchange of information. So far no studies have investigated whether persuasion differs between users of video-mediated systems with or without eye contact.

The issue of the effect of eye contact on communication remains inconclusive. Eye contact seems to be preferred by users but may not lead to any improvement in communicative performance. More research is required to investigate if systems which claim to provide eye contact do in fact support mutual gaze in manner similar to that used in face-to-face communication.

2.5.1.5 Summary of Research Comparing Physical Variables Within Video Technology

The research cited in the previous sections shows how differences in the configuration of a video-mediated communication system may lead to differences in the manner in which individuals communicate using such systems. For this reason it seems worthwhile to describe video-mediated communication technology, not as a technology in itself but rather more accurately as a family of communication systems.

When considering the impact of any variable, or indeed medium, upon the communication process, impact has been considered in terms of any effect on performance or on subjective impressions. It should be noted that there could be a certain conflict between these two psychological variables, a point outlined by O'Connail (1994). Video is regarded as beneficial in negotiating and bargaining situations (Fish *et al.*, 1993), as fewer discussions ended in deadlock (Short *et al.* 1976; Williams, 1977). However, without the video the problem would be resolved "on the merits of the case" (O'Connail, 1994). Performance on this task can be evaluated in one of two ways. The first viewpoint is that taken by O'Conaill (1994) which implies poorer performance in the video condition - highlighting the shift away from the "merits of the case". The alternative view focuses on the fact that discussions were resolved more often in the video channel and therefore performance is improved. Which is more appropriate may depend upon the nature of the task – if a dispute was to be resolved at all costs then the addition of video may be helpful, if the nature of the resolution was more important, then video might be less useful.

Another example of a situation in which there is an apparent incongruency between what is liked by users and what appears to be the most effective, is with eye contact. Eye contact being preserved in a video link was preferred by individuals (Acker & Levitt, 1987) but was in actual fact found to be less efficient, since dialogues were actually longer to get a problem solved than without eye contact (Doherty-Sneddon *et al.* 1997). Again the nature of the task will determine which of these situations is preferable, it is not for the researchers to make a value judgement.

2.5.2 Experimental Variables in Studies of Video-Mediated Communication

When comparing studies which investigate the impact of technology on communication, it is necessary to take into account not only the physical differences between technologies used in the studies, but also the experimental differences inherent between different studies. Two of the main factors in considering studies of video-mediated communication are the task, that individuals used the system for, and the nature of users themselves.

2.5.2.1 Examples of Tasks Used in Studies of Video-Mediated Communication

It is possible that the nature of the task may impact upon the likelihood of communication being affected by video-mediated technology. Some of the differences which have been found between video-mediated communication and audio-only communication for instance may only occur under certain circumstances, circumstances which may include the nature of the task as well as the configuration of the technology.

One of the main distinctions between tasks is the purpose for which video conferencing technology is used Olson & Olson (1997) distinguish between tasks where the main use of the audio and video channels is to support verbal interaction (abstract material) or to show another form of data such as views of shared work objects. This is similar to a distinction that Whittaker (1995) made when describing the difference between "Video-as-data" or "concrete" material.

If the nature of the task involves another object which requires description to the other participant the video channel can be used to show the object. Marshall & Novick (1995) had participants complete a geometric puzzle. The video condition avoided lengthy and often inadequate verbal descriptions of the pieces which characterised the audio-only condition. Other applications of technology that have used video to show information rather than another conversant include the monitoring of surgical operations (e.g. Nardi, Schwarz, Kuchinsky, Leichner, Whittaker & Sclabassi, 1993) and engineering, training applications (Whittaker, 1995). These uses of video are virtually entirely impossible in an audio-only situation.

Watson & Sasse (1997) found that when participants were engaged in conversation using a system it was rated of lower quality than if they simply listened. This suggests that users may only notice any problems with communication when they are involved in communicating interactively via the technology.

Rudman et al (1997) speculated that difficulty of the task may lead to differences between the impacts on the user of video and audio quality. Olson, Olson & Meader (1994) found slight improvements in users task performance for video when using an automated post office task. Veinott et al. (1997) stressed communication by using non-native speakers and found an improvement in performance in the video condition. It seems fair to expect that increasing the complexity of the communication required would lead to a need to use all communication channels possible rather than simply resort to one or two.

One of the reasons that video-mediated communication has often been shown to appear similar to audio-only communication may be that individuals already have adapted to using the telephone and have this strategy within their personal communication abilities. If they are using video and they find that their face-to-face communication method does not work then they may revert to their telephone strategy. If a task was difficult enough for audio communication to be insufficient then perhaps users may adapt more to the video channel and succeed where they would fail in a task using audio-only communication. At present there appears to be no research which has aimed to systematically manipulate the difficulty of a task, although this may potentially be an area of future research interest.

2.5.2.2 Comparison of the Users of Video Technology

The final variable to be considered is the users of the equipment. Different groups of individuals may have their communication affected in different ways depending upon their expertise, culture, relationship, age, and so on. Research into the different ways in which video-mediated communication affects different users is sparse but is an area beginning to be looked at in detail.

As mentioned earlier, Veinott, Olson, Olson & Fu (1999) showed that individuals communicating in a language other than their native language showed improved performance on a problem solving task when they had used a video communication system compared with audio-only. The key to the video channel having been helpful may have been that communication between the users was more difficult.

Other situations where communication is stressed may be related to the amount of common ground between participants (Clark, 1996). Most studies have used people with a high level of common ground, a situation where an expert attempts to communicate to a naïve person would perhaps be more stressed. To converse successfully, people adjust their language to speak at the highest level of common ground (Isaacs & Clark, 1987). In real-life situations however, distributed groups are likely to have greater occupational specialisation (Fussell & Benimoff, 1995). Fussell & Benimoff suggest that one function of the video channel may be to provide information about the likely knowledge of the other people. This would clearly make the visual channel more important between unfamiliar people.

2.5.3 Differences Between Face-to-Face and Video-Mediated Communication

All of the channels of human communication described earlier exist in face-to-face communication. Not all of them however, can presently be supported in a video-mediated environment. Despite claims that video conferencing can produce "a face-to-face experience", the fact is that there will always be differences between a face-to-face situation and a video-mediated interaction.

One fundamental difference between a co-present situation and a remote interaction is in the focus of one's attention. In a face-to-face environment, you are most aware of your physical surroundings (e.g. what is at your feet, beside you etc. as opposed to what is further away). In a video-mediated interaction on the other hand, you are most aware of what is directly in front of you – often at a distance and have little or no awareness of the periphery in telepresence (Yamaashi et al., 1996). Related to this point is that in a video-mediated interaction, people may often feel constrained by the focus of the camera. Hiding tends to be much easier in a remote interaction – anything not seen by the camera is guaranteed not to be seen by one's conversational partner.

In the majority of video-mediated communication systems, the size of the image which users view is a lot smaller than life-size. Studies using desktop video conferencing have often used a video window as small as 4.5 inches (Anderson et al., 1996). It has been pointed out that reducing the video size from life-size is likely to detract from the usefulness of the visual channel (O'Donnell & Draper, 1997).

At present video-mediated communication is also still a novelty. Just as people who are now used to using the telephone have evolved strategies for communicating on the phone by their different conversational style, so will users of video. This is a point that must presently be remembered when attempting to draw conclusions about how people will communicate using multimedia technology. When considering how people are likely to communicate when using video-mediated communication and how it may be best designed to encourage successful communication, it is useful to consider which channels of communication may be available.

2.5.3.1 Which communication channels are present in video-mediated communication?

Of all the communication channels discussed earlier in this chapter, some are always available in video-mediated communication, some are never available and some may or may not be available depending upon the system architecture. The starting point for this discussion shall be simply an audio-visual link.

Of those communication channels that are always available in a video-mediated communication system, the most obvious is the visual channel. Even at the lowest quality, the visual channel allows for communication information about the person or people one is interacting with. This allows information to be conveyed which may give clues to their gender, age, culture, surroundings, social class and so on. Audio is also always available in an audio-visual system. Although the intelligibility of speech may be degraded, in general verbal communication information can be successfully transferred and is complete. The only audio information missing is arguably the cue of volume as a cue to proximity.

Equally, there are some communication channels which are never available in a video-mediated encounter. Olfactory information for instance, which can be a cue to close proximity in a face-to-face environment, is presently not possible. Similarly the ability to touch the remote individual, is currently impossible - although remote haptic devices are being developed which in the future may give an impression of touch (e.g. Brave, Ishii, & Daly, 1998).

Of most interest are those communication channels that may or may not be present in video-mediated communication. Some of these are purely a physical matter, such as the ability to transfer written information, by seeing someone writing. If a system has a shared whiteboard then this may be possible – without then one would have to rely on holding up paper to the camera and the chances of the writing being visible would be small. Other channels depend upon how various aspects of a video-mediated system are configured.

An example of such a communication channel is gestural information. Heath & Luff (1991) reported in a naturalistic study that gestures with interactional significance are relatively ineffectual across video. A similar result was reported by Tang, Isaacs & Rua (1994) who found that hand gestures were commonly missed by individuals using the Montage video communication system. There may be two reasons for this failure to successfully transmit gestural information. On the first hand, poor frame rate or a narrow camera angle may lead to a physical gesture actually failing to appear on the screen. Frame rates of around 5 frames per second are not uncommon for networked multimedia systems, and lead to a refresh only once every 0.20 seconds. A gesture may commonly take less than this length of time, particularly if it is something such as a wave or a nod in which movement is important. Alternatively, a gesture may be missed by a user because they may not have been looking at the screen. When using desktop video conferencing in particular, there are often many other things to look at, the camera, any other information on the screen etc. When coupled with the fact that the size of the image is often smaller when using video conferencing than in a face-to-face situation, it is clear that missing a gesture is a much greater possibility in general than in a face-to-face setting. If the frame rate was faster, and users were watching for gestures then the gestural channel would be available, if not then it is unavailable.

Another example of a communication channel that may or may not be available in video-mediated communication is gaze. In the majority of systems, eye contact is not possible, due to the camera being placed to the side of the monitor. Recently, some systems have developed with the camera either behind the screen or filming a mirror image to effectively create an impression of eye contact (Acker & Levitt, 1987; Okada et al., 1994). These systems may allow gaze to be used as a communication channel, although as described earlier, preliminary studies suggest it may still not be used in the same way as face-to-face (Anderson et al. 1996).

2.6 Proximity

Another example of non-verbal communication which presently remains unresearched with respect to video-mediated communication is that of proximity - the physical distance separating two individuals. For many years it has been known within social psychology that proximity is an important source of non-verbal communication information. Social rules appear to exist for how closely people interact, depending upon their relationship (Allegier & Byrne, 1973), culture (Vaksman & Ellysman, 1979), personality (Leary, 1983) and situation (Alexander & Rudd, 1981).

It is known that there are culturally distinct rules for how far apart people choose to be when they interact (Hall 1966). If people interact at a distance closer or further to that at which they would normally, behaviour may be affected in a number of ways (Capella 1994). Proximity can influence dialogue, gaze, body position, persuasion and subjective opinions about the individual with whom they are interacting.

Hall categorised interaction distances for seating as intimate space (6-18 inches), personal space (18 inches - 4 ft), and social space (4-12 feet). For each particular type of interaction there is an appropriate distance which corresponds to one of these distances. For example, talking to a close friend may be done within personal space, whereas talking to a stranger will usually be done within the social space. These distances are culture specific, although the general pattern of the areas of space appears to be fairly general across cultures.

If one person stands closer than would normally be expected then individuals tend to adjust their body position in order to preserve their personal space (Becker & Mayo, 1971; Efran & Cheyne, 1973; Sawitsky & Watson, 1975); decrease the amount of eye gaze they use (Patterson, 1977; Argyle & Dean, 1965); decrease the average duration of their responses (Johnson & Dabbs, 1976) and decrease the number of affiliative responses in a conversation (Capella, 1986). As well as all of this the time taken for an individual to leave the situation increases (Poyatos, 1975).

As well as these behavioural responses to the breaking of the normal rules of proximity in an interaction, there are also interesting subjective effects caused by individuals standing closer than normal to converse. Those who stand closer than is usual may be perceived as warmer and friendlier than those who stand farther away (Paterson, 1968).

The main purpose of this thesis is to investigate the role of proximity in video-mediated communication. The effects of proximity are only briefly introduced here to give a flavour of why proximity is important as a type of communication, and shall be discussed in more detail in Chapter three.

2.6.1 The Potential for Proximity in a Video-mediated Interaction

It is unknown whether video-mediated communication technology may be able to convey impressions of proximity. There are however, reasons for hypothesising either way whether it is likely to exist or not. When using videoconferencing, there are a number of factors which affect the actual proximity between individuals. How far an individual sits from the camera at one end and the distance from the screen at the other end are both fixed and are directly observable, physical distances. Then there is the actual physical distance of separation in location between the two rooms. Although measurable, this may not be directly observable to those individuals involved in the interaction. A view of an individual in a videoconferencing room in London probably differs little to a view of a videoconferencing room in New York, Adelaide or Cairo, for instance.

So in a sense, proximity is unlikely to differ much between video-mediated interactions. The distance people sit from the cameras and screens tend to be fairly standard, affected more by the relative comfort of seating positions than any attempt to transfer particular communication information. Alternatively, it is conceivable that changes in the presentation of an image may influence the appearance of distance of a video image. When considering a video camera, people talk of "zooming in" or "close-up views". Is it possible that such a difference may lead to impressions of closeness? or does the fact of being remote override any such possibility? This question shall hopefully be answered over the course of the next 7 chapters.

2.6.2 Why the Study of Proximity in a Video-mediated Communication is Interesting

There are several reasons why the study of proximity in this context is of particular interest. Empirically, it is interesting to investigate whether or not impressions of proximity can be conveyed in a video-mediated interaction. Some studies have shown that impressions of presence can differ between technological configurations (e.g. Gale 1990) but none have so actually investigated how presence may equate to proximity.

In face-to-face communication, proximity has been shown to affect both individual's communication behaviour and their subjective attitudes. If proximity were to have similar effects with video-mediated interactions then that would illustrate how a single factor within video-mediated technology could impact upon an interaction.

Another point of interest is that in a face-to-face interaction, interlocutors know exactly how close they appear to be to the other parties present. Apart from in exceptional circumstances they also have virtually total control over their appearance. In a video-mediated interaction however, individuals are often no longer in control. The image that people see is often determined by the receiver, dependent upon the configuration of their screen as much as the configuration of the cameras at the end of the transmitter. Should perceptions of proximity in video-mediated interactions differ according to the configuration of the equipment, then it is possible that individuals will no longer be in control of how far away they appear to be. This may lead individuals to disregard any information they receive that they know not to be under the other individuals control or they may be susceptible to this ultimately erroneous data.

A related issue is that video-mediated communication technology could lead to novel types of interaction. In a face-to-face environment, an individual is always equally far away from the individual with whom they are conversing as that individual is to them. In a video-mediated interaction, this natural equilibrium may no longer exist. It is conceivable to consider a situation where one individual appears close and the other appears far away in a dyadic interaction. How this may affect the way people interact is unknown, as these situations are entirely novel.

As well as being empirically interesting, this research should also be of worth in an extremely practical sense. If perceptions of proximity lead to systematic differences in the way people interact when using video-mediated communication technology then that knowledge could lead to certain usability benefits. People would be able to configure their equipment for what they consider to be an optimal interaction.

2.6.3 Approaching the Problem

The aim of this thesis was to examine the potential impact on communication of a presently unstudied area of video-mediated communication systems. There are two main ways in which this aim could have been approached. Most of the studies so far described have employed an experimental approach to the study of communication, where a variable is identified for investigation and experiments conducted to examine the effects of manipulating this variable (e.g. Veinott et al., 1997; Anderson et al., 1996).

An alternative to this methodology is to take a more top-down approach to the problem. A more ethnographic approach would seek to understand the impact of particular types of system by studying the way that individuals use the system in a real-world situation. An example of this ethnographic approach to the way in which communication systems are used was a study by Hughes, J., Kristofferson, S., O'Brien, J., & Rouncefield, M. (1995) who studied the use of a video mediated communication system to support meetings. This approach can be particularly beneficial for informing the design of systems for it ensures that the context of a situation is taken into account (see also e.g. Randall, D., J.A. Hughes and D. Shapiro, 1994).

In the present thesis, it was decided to use an experimental approach to initially investigate the issue of proximity in video-mediated communication. The reasons for this choice were both theoretical and practical. Theoretically, there were strong reasons for predicting that proximity may potentially affect remote communication, based on the work of Hall and others in face-to-face communication. To understand perceived proximity in a remote however, requires stringent controls on issues such as the configuration of the camera and seating position. On a practical level, the low level of implementation of video-mediated communication technology makes access to users of these systems very difficult to get. It was felt that while it would have been desirable to use field studies to infer whether or not proximity was likely to be an issue worthy of extensive investigation prior to conducting experimental research, it was both practical and theoretically valid to adopt an experimental approach from the outset.

The first goal of this thesis is to investigate whether users experience perceptions of proximity in a video-mediated interaction. If successful at isolating a method for systematically varying impressions of proximity than the next challenge will be to understand what effects perceived proximity may have on an interaction. If effects are found to be caused by changes in perceived proximity then the research will seek to discover the extent of these effects and to understand why they occur.

In investigating these effects, some consideration to context must be given in order to keep the problem to a manageable level. Firstly, the research is attempting to compare only within video-mediated interactions, not across media. Secondly, only dyadic interactions will be studied, in an attempt to reduce the complexity of the conversations. Finally, the research shall use video equipment which is of higher quality than much of the commercially available equipment. This is an attempt to keep up with the advance in technology – as standards are constantly improving, it makes no sense to research the poorer present day standard of video mediated communication on an empirical level.

To optimise the potential of new communication opportunities, it is important to understand how individuals communicate when using them. Understanding the impact of technology on communication will enable technology to be used and designed with greater consideration to the manner in which individuals will use it. New situations however, may lead to new ways of communicating. The study of proximity in video-mediated communication may potentially enrich our understanding of the way people communicate remotely. Firstly our aim is to discover whether or not proximity can be supported in a video-mediated interaction.

3 Chapter 3. Study 1: The Effects of Camera Zoom on Perceived Distance

As noted in Chapter 2, it is unknown whether proximity, as a form of non-verbal communication information, may be preserved in a video-mediated interaction. This chapter reports a study which attempted to explore the possibility that changes in the configuration of a video image may lead to a viewer experiencing differing perceptions of proximity. The question that the Study sought to answer was: Can differences in a video representation of a remote individual lead to differences in another's perception of their proximity?

Chapter 2 outlined the existence of proximity as a form of non-verbal communication in a face-to-face setting. The aim of this chapter is to consider how this information may be manifested in a video-mediated interaction. Before considering how to investigate such an issue, it is necessary to understand in more detail the role of proximity in a face-to-face interaction.

3.1 Proximity in Face-to-Face Interactions

In a face-to-face interaction, proximity is the physical distance of separation between interlocutors. Well defined cultural norms exist regarding the appropriate distance at which certain types of interaction take place.

Hall (1959) categorised 4 different areas of communication space which exist for people from a Western culture. These areas of space were intimate space (0-18 inches), personal space (18 inches to 4 feet), social space (4-12 feet) and public space (beyond 12 feet). Hall, an anthropologist, also outlined the differences between these spaces and the types of interaction that were appropriate in each situation.

Intimate Space. Intimate space is reserved only for lovers, family, small children and very close friends. At this distance vision is distorted with large objects appearing blurry. Olfaction, touch and thermal sensation are also available at this distance while verbal communication is kept at a whisper.

Personal Space. Within personal space, individuals can touch each other, but contact is not necessary, visual distortion does not occur, olfaction is minimal and thermal perception is non-existent. This space is reserved for friends.

Social Space. Social space is the distance at which casual interactions between acquaintances and strangers take place. Physical barriers such as desks, tables and counters can be used to make people keep this distance.

Public Space. Public distance allows the amount of space desired among strangers. At farther distances speech must be projected or amplified to be heard.

Hall believed that the existence of these types of space were consistent across cultures, although the actual distances corresponding to each space may differ between different cultures. Research investigating proximity in other cultures has shown that North Americans have one of the largest areas of personal space, and that in certain "contact cultures" such as Southern Europeans, Arabs and Latin Americans this distance is less than half the size (Watson, 1970; Nydell, 1987).

A lot of research has focussed on what happens if these social rules are violated in an interaction. Should a stranger stand closer than is appropriate, then an individual may feel that their personal space is being invaded. This generally leads to feelings of unease and several compensatory behaviours have been observed by which people attempt to regain their personal space (Argyle, 1988; Capella, 1994).

These areas of space are attached to the individual but are also dynamic, and can be affected by other situational factors. In crowded situations, for example a train or an elevator, strangers often stand within a distance that would usually be considered personal space but may not usually be considered to be an invasion of territory (Vine, 1973).

3.2 Could Proximity Exist in a Video Representation of an Individual?

The relevance of the face-to-face research on proximity as far as remote interactions are concerned is unknown. If video-mediated communication technology can transmit proxemic information then there is the possibility that the areas of psychological space may retain some influence on individuals' feelings. Alternatively, should the images of remote interlocutors fail to elicit any sense of proximity then the psychological effects of space are unlikely to have any impact on a video-mediated interaction.

In any video-mediated encounter, there are some definite differences to co-present interaction. Firstly, it is always clear in a video-mediated interaction that one's partner is remote. Although systems may purport to give the impression of being in the same room as someone else, this is certainly not the case. Secondly, not all communication channels are available. There is no opportunity for touching, or for olfactory information to be transmitted, both of which were highlighted as factors which help define areas of space. Both of these factors are also considered to be important in the invasion of personal space (Argyle, 1988).

In a video-mediated interaction there are three physical distances which may be of importance in determining the concept of distance. These distances are, the distance from the remote individual to the camera, the distance of physical separation between locations, and the distance from the screen to the viewer. In terms of the visual representation, the most relevant distance of relevance may be that between the viewer and the screen. This is also the only distance over which the viewer has any control. In a sense, no matter the physical separation that may exist between individuals, the actual appearance of one's partner is only as far away as the screen on which they can be viewed.

Studies of film and television have suggested that images presented on screen may indeed lead to different impressions of distance. Persson (1998) suggests that images viewed in closeup can lead to feelings of both threat and intimacy. In describing this, Persson alludes among others, to the effects of close-up images of spiders leading to viewers feeling uncomfortable. Persson claims that close up images lead to viewers feeling spatially close to the characters being viewed, however no quantifiable evidence of this is offered.

Reeves, Lombard & Melwani (1992) conducted an experimental study which led them to conclude that faces on the screen can invade an individuals' personal space. They concluded that faces on a big screen, faces viewed close to the screen, and close up faces gained greater attention, received stronger evaluations of character and resulted in better recognition memory. This, they claimed showed that characters on the screen could appear to be close.

These studies have not been done in an interactive setting, nor was the focus of the study the proximity of individuals. Most of this work has been done under the study of "telepresence". Defined by Steuer (1992) as "the experience of presence in an environment by means of a communication medium" (p.76), telepresence is a generic term referring to the involvement people feel when looking at mediated images. It is not clear how generalisable studies which have investigated individuals' perceptions of screen characters and television images are, to perceptions involving people with whom one can interact. Films In particular involve a suspension of disbelief (Bleich, 1984; Laurel, 1991) and the viewer's intention is also to a large extent to imagine being somewhere they are not.

A review by Lombard & Ditton (1997) on the concept of "presence" outlined six alternative conceptualisations of presence within the study of different media. Of these, the ones which are of most relevance to interpersonal communication are "social richness" and "realism". Social richness is the extent to which a medium is perceived as sociable, warm, sensitive, personal and intimate when used to interact with other people. This includes the effect of the medium on the way people feel about others they interact with remotely. Realism is the degree to which a medium can produce accurate representations of objects, events and people. In the present study, the aim is to investigate whether a video image can lead to the impression of a "real" distance or changes in distance. The primary aim of this study is to investigate how a video link may be able to provide cues towards a "real" appearance of distance.

3.3 Perceiving distance

Despite there being an abundance of work on depth cues in the perception literature (see any introductory perception text e.g. Gregory, 1988; Bruce & Young, 1986), so far no research has investigated how people judge absolute distance (also known as egocentric distance) from a live video image. This is unsurprising, due to the fairly abstract nature of the question. When considering how an individual might approach such a task, the answer may be informative as to how individuals conceive a video image.

If there are no background cues, such as objects which are a known size, to provide information on depth in a picture, then there are two theoretically possible strategies for determining distance in a video image. These differ according to how the relationship between the camera and the video window is perceived. Distance may either be estimated on the basis of the physical size of the object image (e.g. a face) on the screen or by the amount of the visual field of the picture an object image takes up (the image to background ratio). In a face-to-face situation these two strategies are inextricably linked – the size of an image on the retina is directly related to the amount of the visual field it occupies. Depending on how individuals consider a video window however, these cues may be separated.

If a video window is considered to be a representation of all the camera can see, then the window can be considered to show an entire visual field. In this case, distance is most accurately judged by considering the amount of the background occluded by the object image. Alternatively, if the window is considered to be equivalent to viewing part of a larger visual field, then the amount of background the object image takes up is less relevant to an interpretation of distance as the entire visual field cannot be judged. In this latter case, the actual size of the image per se is the relevant factor in determining distance.

The present study sought to determine which perceptual strategy people used to determine distance in a video image. This was done by having two conditions, which differed in the size of the video window. Participants either completed the task viewing a small or a large video window. In both conditions the view of the camera was identical - a large video window versus a small one only meant that the pixels were enlarged, thus the resolution was slightly lower in the large condition, but not noticeably. The amount of background visible in a large video window is identical to the amount of background in a small video window.

3.3.1 Hypothesis 1 – The viewer treats the video window as an entire visual field.

This theory is the equivalent of treating the camera as another pair of eyes. The video window is effectively considered to be representing the entire view of the camera. Using this theory, an estimation of distance and changes in distance should take into account the size of the object relative to the size of the background. In other words distance is directly related to the area of the visual field that the object fills. Thus changing the size of the video window should have no effect on the perception of distance, as the image inside the window is kept the same, with the same relative dimensions between objects within the image.

3.3.2 Hypothesis 2 – Viewer treats the video window as a partial vision into a larger visual space.

This theory differs from the previous theory on the basis that although the camera is still seen as a pair of eyes, the video window is seen as having a blinkering effect. In this theory, distance is seen only as a function of the size of the image, and the amount of background visible plays no part in the perception of distance. Increasing the size of the video window is seen as equivalent to widening the view of the camera, and therefore more of the visual field can be seen. Therefore, changing the size of the video window will have an effect on the perception of distance. When seen in a smaller video window objects will be likely to be judged as further away due to their smaller appearance.

3.4 Appearing Close vs. Feeling Close

When considering video-mediated proximity or virtual distance, it is important to note that there may be a crucial difference between appearing close and feeling close. Even if an individual appears close the knowledge that they are actually remote may prevent them really feeling close. This can be explained by considering the example of a camera image of a ball. Zooming the camera in on the ball may well make it appear closer but would it feel closer? No matter what the focal length of the camera, the size of the screen, or the proximity of an individual to the screen, the ball will always be remote and can never be touched, kicked, moved, bounced or interacted with in any way. In this sense it is conceivable that although the ball may appear closer it does not feel closer. The present study sought to differentiate between the appearance of distance and the feeling of distance by asking questions about how participants felt about images relayed in different ways.

3.5 Designing the Task

There were a number of aims to the design of the task. The first was to investigate whether or not changes in camera zoom could relate to differences in feelings of proximity of an individual presented in a video image. Secondly, if camera zoom does indeed lead to differences in impressions of perceived proximity, how do these relate to the areas of space proposed by Hall (1966) for face-to-face interactions? Can video-mediated images of individuals invade personal space?

The study also sought to discover the determinants of perceived distance across a video link. Do people perceive impressions of distance from image size alone or is the relationship to the background taken into account? Does the size of the video window affect distance perception?

3.5.1 Methods used in face-to-face research

There have been a number of ways of studying proximity in a face-to-face situation, not all of which would be suitable for studying the potential for proximity in a video-mediated situation. Argyle (1988) outlines 5 methods of studying proximity and human behaviour. These are as follows:

Move and stop: Subjects are approached by the experimenter or a confederate and are asked to stop the approach when he or she begins to feel uncomfortable.

Real-life observation: Hall (1966) used observation to study proximity in different cultures. He took photographs of people in real-life situations and compared them. This method produces high levels of validity but cannot be used to conduct controlled experiments.

Naturalistic laboratory methods: This method involves giving subjects a choice of where to position themselves to carry out any particular task. For example, there might be a number of chairs that the subject could choose to sit in when left in a waiting room. Recordings can be made of where they choose to position themselves, without the participants ever being aware that the interest of the experimenter was based on where they sat.

Projective methods. Hypothetical situations can be set up using dolls or other equipment, with participants being asked to imagine a situation and indicate how they would feel at a particular distance, or at what distance they would choose to interact in a given situation. Mehrabian (1972) asked subjects to interact with a hat rack, imagining it to be another person. McCroskey, Young & Richmond (1977) demonstrated that projective methods were highly valid. They demonstrated a close correlation between the distances subjects said they would choose to sit at in varous situations and the distances that individuals actually chose to sit at.

Invasion of territory. In this type of experiment, the experimenter records participants' reactions when their personal space is invaded. This may include types of stress symptoms, motor responses, attitudes etc.

None of these methods clearly suit the task of investigating whether distance can be represented by video. The core of the problem is to measure participants' impressions of the appearance of distance in an image. Participants are being asked to express a two-dimensional image in terms of a real three dimensional environment. This required a method of judging virtual differences by physical measures of distance.

Since comparing distances is easier and more reliable than judging absolute distances, the chosen method was to compare the image of an individual with a co-present individual placed at a particular distance. Participants were asked to alter the level of camera zoom until the individual on the screen appeared to be the same distance as the person in the room.

The question "which of these people is further away?" is usually much easier to answer than the question "how far away is that person?". Absolute judgement is harder because it requires a mental representation of a particular unit of measurement. Comparison, on the other hand, has all necessary information available.

The main advantage of this methodology is that all need for units of measurement is removed. A participant will not be penalised for being unable to judge six feet. Were individuals asked to zoom the camera until the individual in the image appeared to be four feet away, there would be a lot of variance based on participants' skill in judging four feet, before taking into account the variance caused by different peoples' impressions.

In order to keep a relationship with the areas of space proposed by Hall (1966) participants were asked to judge distances which related directly to those areas of space. Participants were also asked by questionnaire to rate how they felt about each distance, both in terms of face-to-face and video images.

3.6 Method

3.6.1 Participants

20 participants (10 male, 10 female, age range 23-59) agreed to take part in an experiment which investigated perceptual judgement. All participants were were employees within a large corporation, based at one site. Participants completed the task in pairs, all of whom knew each other.

3.6.2 Technology

The video link was provided by connecting a camera directly into a Pentium 200MHz PC running Windows 95. The software used was that running VDO phone software. One remote control camera (Sony, model PV PTZ Camera 300-NTSC; part No: 528-000421) was connected directly to the computer. This camera could be zoomed remotely by the individual on whom it was focused by using an infra-red handset. The monitor had a 17 inch screen.

3.6.3 Conditions

The experiment was a between subjects design with one independent variable (window size). There were 2 conditions with 10 participants in each condition. The conditions differed only in the size of the video window. The large window measured 200mm x 160mm (width x height). The window in the small condition measured 100mm x 80mm (width x height). In each condition the video window was placed in the centre of the screen, the rest of the desktop was plain. The image projected within each window was identical, simply enlarged or shrunk to the size of the window.

3.6.4 Procedure

Participants were informed that they would be taking part in a short psychology experiment into the nature of distance perception across a video link. All participants consented to still image captures of their faces being recorded for research purposes.

In each trial there were three roles. Participants were run in pairs, with one playing the role of the participant and one the role of the confederate. At the end of each session they swapped roles.

The Participant. The participant sat on a swivel chair by a computer monitor. On one side was the monitor - on which was an image of the confederate - and on the other side (180 degrees to the monitor) was the co-present person. The participant was instructed to ask the confederate to change the camera zoom until they felt that the image on the screen appeared the same distance away as the co-present person.

The Co-present Person. This role was played by the experimenter and involved sitting on a chair facing the participant at a predefined distance.

The Confederate. This role involved sitting facing a remote controlled camera. The confederate had the power to alter the zoom of the camera as directed to by the participant. They sat behind a screen on a prepositioned chair, six feet from the camera. At the shortest focal length, the image from the camera showed less than the entire face, at the longest, it showed all of the upper torso, face and a lot of background.

When the participant was satisfied with the image, the confederate placed a marked ruler on their forehead and a screen snapshot was taken which could be analysed later and used as an objective measurement. Each session consisted of 5 trials all at different distances (2.5 ft, 4ft, 6ft, 9ft, 12ft). These distances corresponded to Hall's (1966) theories of space. 2.5ft is well within personal space, 4 ft is at the edge of personal space, 6ft is in social space, 9ft and 12 ft are at the far end and outwith social space.

3.6.5 Questionnaires

At the end of the experiment participants completed a short questionnaire. This contained questions designed to gain an insight into how participants felt about the images and also how comfortable they felt about the co-present person. A copy of the questionnaire can be found in Appendix A and a full breakdown of the results of the questionnaire is available in Appendix B. Within the breakdown of the questionnaires, results were analysed both between conditions, comparing the results for those with the large window with those for the small window, and across conditions, looking at the opinions of participants overall.

3.7 Results

Two types of data were collected, one subjective and one objective. The questionnaire provided an insight into participants' subjective impressions of the images and the feelings caused by both the images and the co-present individual. Analysis of the snapshots taken during the task involved measuring the length of the ruler on screen, this was present in every snapshot and provided an objective comparison of the image size chosen to represent each particular distance.

3.7.1 Questionnaire Results

3.7.1.1 Most Appropriate Distance for Interaction?

The questionnaire contained 3 questions which asked which of the five images (representing each distance) would be most appropriate to use in a particular video-mediated situation. The remote individual was either a call centre agent dealing with a loan enquiry, a friend, or a stranger. The scenario of a call centre agent was closely related to the line of business that all participants were involved in and thus was a familiar situation to each of them. The results are shown in Table 3.1 and Figs. 3.1-3.3.

Distance	Scenario					
	Call centre agent		Friend		Stranger	
	Small	Large	Small	Large	Small	Large
2.5 ft	10%	10%	10%	20%	10%	10%
4ft	20%	40%	30%	40%	10%	20%
6ft	30%	50%	50%	30%	30%	40%
9ft	30%	0%	10%	10%	30%	30%
12ft	10%	0%	0%	0%	20%	0%

Table 3.1. Percentage of respondents who responded that the image in a particular distance was the most appropriate for remote individuals to use in a video-mediated interaction. Results split by condition (small window vs. large window).

Fig 3.1 Call centre agent

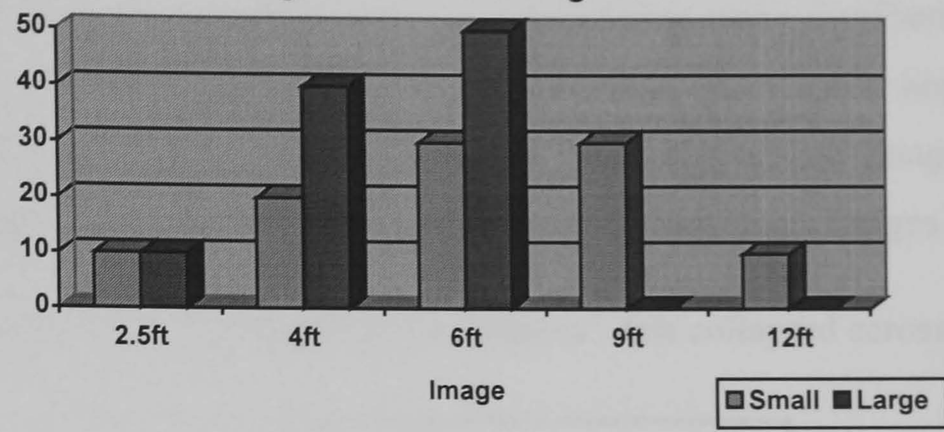


Fig 3.2 Stranger

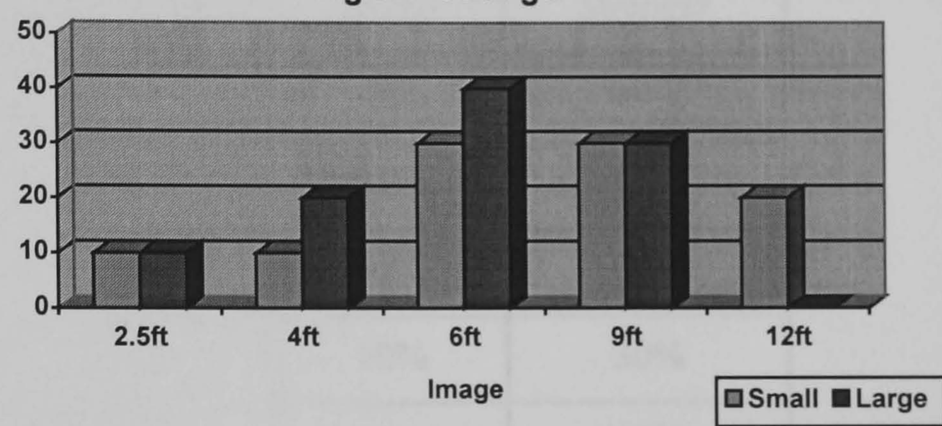
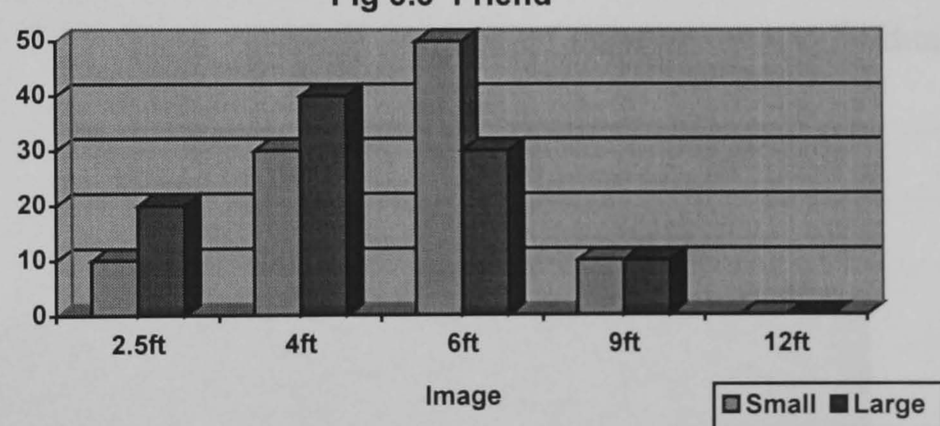


Fig 3.3 Friend



Chi square tests were conducted on the data for each question independently, comparing the answers of those in the close condition with those in the far condition. None of the Chi square tests showed a significant effect of window size on the most appropriate image to use in a particular situation.

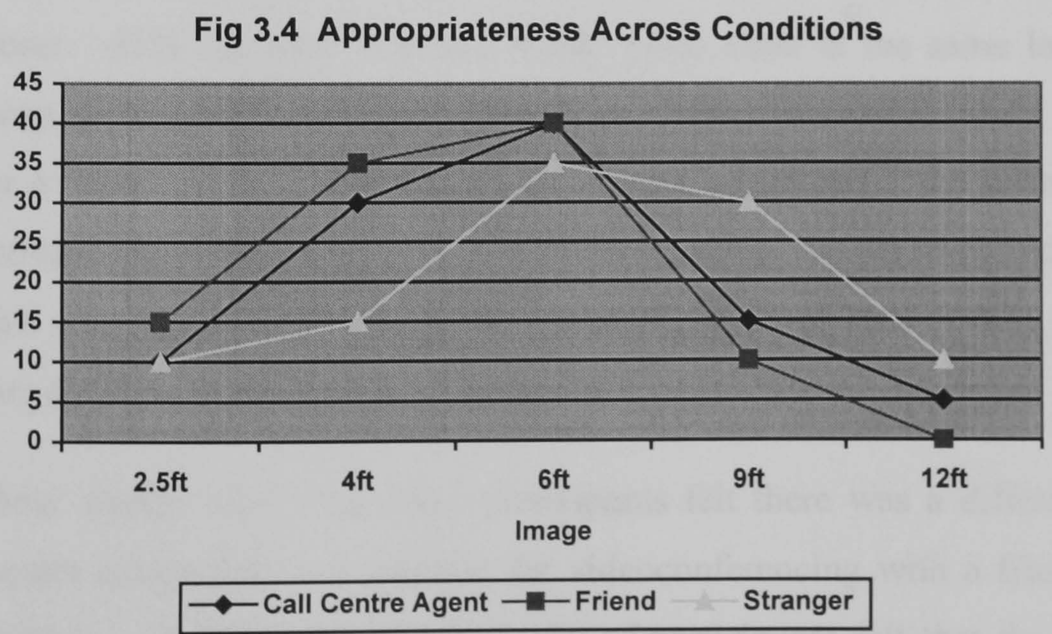
The question which was closest to statistical significance was the one which asked at which distance the call centre agent should be placed at [Chi-square = 5.00, critical value = 9.49, ns]. This is accounted for by the small amount of data. Each condition contained only 10 participants and there were 5 options to choose from.

Although not a statistically significant difference, it can still be seen from viewing the data that there is a trend towards responses being more distributed across the entire range of possibilities in the small condition. This trend hints at greater variability in participants' perception of the most appropriate image when they viewed smaller images, compared to participants viewing larger images.

Table 3.2 presents the "most appropriate distance" data collapsed across conditions.

	Scenario		
Distance	Call centre agent	Friend	Stranger
2.5 ft	10%	15%	10%
4ft	30%	35%	15%
6ft	40%	40%	35%
9ft	15%	10%	30%
12ft	5%	0%	10%

Table 3.2. Percentage of respondents who responded that the image in a particular distance was the most appropriate for each of three remote individuals to use in a video-mediated interaction.



Due to the fact that comparative analysis of these distributions involves within subjects data, a Chi Square test cannot be used. The data are thus presented merely for an illustrative purpose, and are only suggestive of where differences may occur.

It is noticeable that the distribution of the friend and the call centre agent are markedly the same, while that of the stranger appears to follow a slightly different pattern. The results suggest a trend towards a difference between the most appropriate distance for a friend to interact at, compared with a stranger. Notably, these participants also felt that the call centre agent should be placed in the friend category, rather than the stranger.

The next analysis was to study where participants would position a friend relative to that of a stranger. Table 3.3 shows the percentage of participants who reported that the image they would use for video conferencing with a friend was either closer, the same as, or farther than that they would use for a stranger.

Friend closer than Stranger	50%
Friend the same as Stranger	45%
Friend farther than Stranger	5%

Table 3.3 Percentage of respondents who would encourage a friend to use a different image for videoconferencing to that which they would encourage a stranger (%).

The results show that while only one participant would place a friend at a further zoom than a stranger during a video conference, but 50% would place the friend closer. 45% reported that they would place them at the same level. The overall mean score for which image participants would use was 3.15 for a stranger and 2.45 for a friend (when 1 represented the closest image and 5 the furthest image). This corresponds to a friend on average being placed at the 4ft level and a stranger at 6ft. This also corresponds exactly to the conclusions of Hall (1966) who showed that friends talked at 4ft and strangers at 6 ft.

These results show that many participants felt there was a difference between the camera image that was suitable for videoconferencing with a friend and that for a stranger. Of these, the vast majority of participants felt that the friend should be placed closer than the stranger.

3.7.1.2 Could an Image be Imposing?

One item on the questionnaire asked participants whether they ever felt that the image on the screen felt imposing. The results of this question are presented in Table 3.4.

Image	Large Window	Small Window
2.5ft	90%	90%
4ft	20%	40%
6ft	0%	0%
9ft	0%	0%
12ft	0%	0%

Table 3.4 Percentage of participants responding that they considered each image imposing.

The results presented in Table 3.4 show that 90% of participants reported that they found the closest image imposing. 30% of participants also felt that the image at 4ft was imposing. There was no trend towards a difference between the two conditions. 10% of participants found no image imposing.

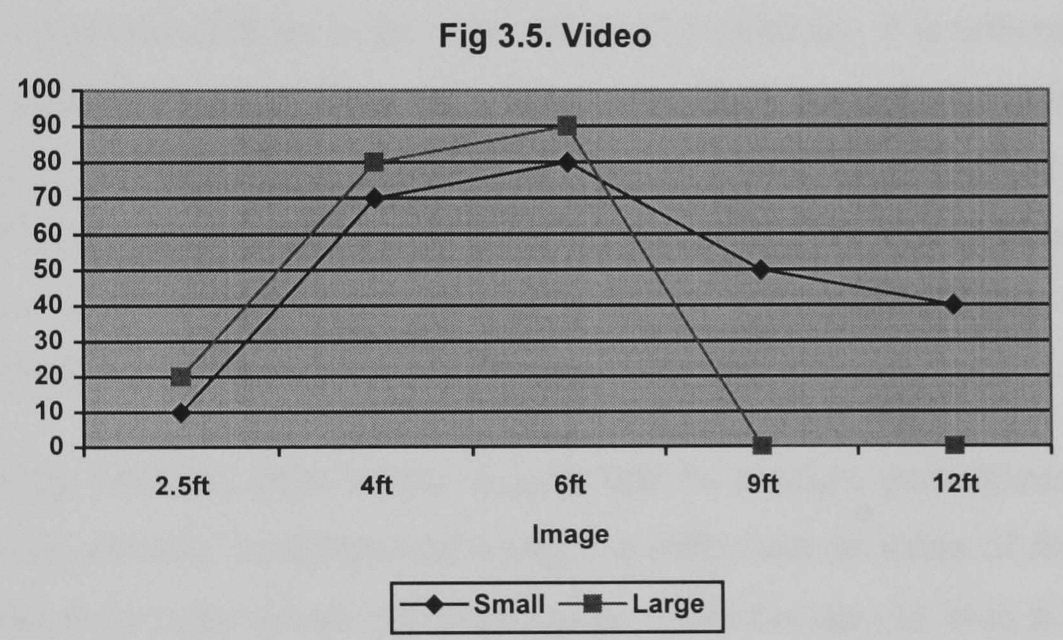
These results show that the large majority of individuals did consider that an individual represented by a video-mediated image could be considered imposing.

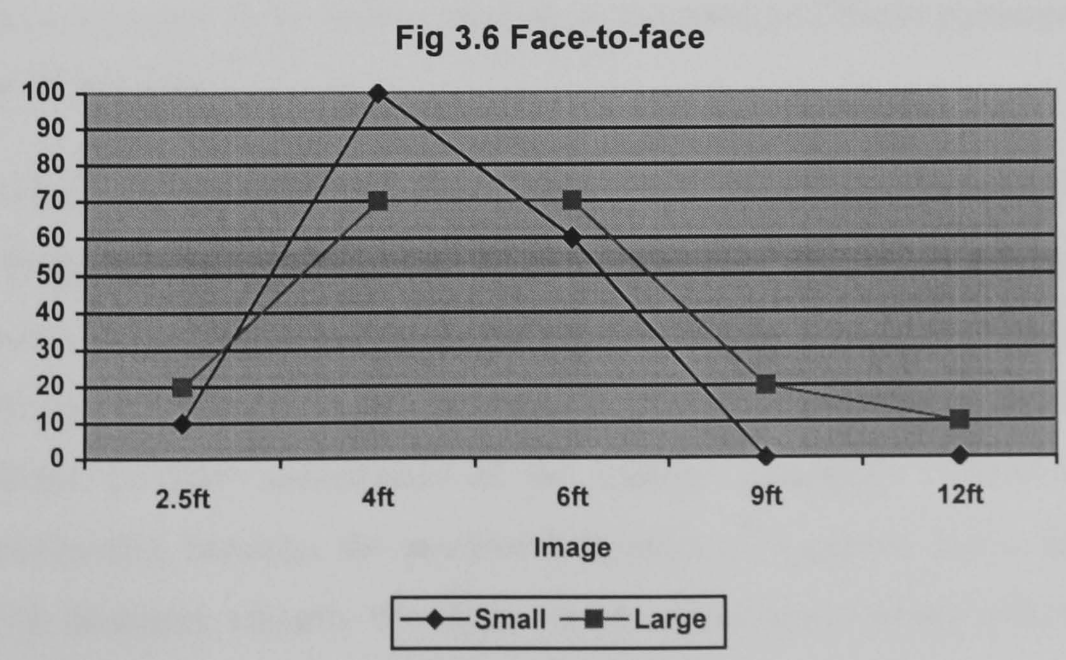
3.7.1.3 Comfort in an Interaction

The final two items on the questionnaire asked participants to report whether they would or would not be comfortable talking to a stranger at each distance. This applied to both the face-to-face distance and the video-mediated image. The results are summarised in Table 3.5 and are graphed in Figs 3.5 and 3.6.

Image	LARGE WINDOW		SMALL WINDOW	
	Video	Face-to-face	Video	Face-to-face
1 (2.5 ft)	20%	20%	10%	10%
2 (4ft)	80%	70%	70%	100%
3 (6ft)	90%	70%	80%	60%
4 (9ft)	0%	20%	50%	0%
5 (12ft)	0%	10%	40%	0%

Table 3.5 Percentage of participants in each condition who reported for each distance that they would feel comfortable talking to a stranger either across a video link or face-to-face.





The results presented in Fig 3.5 and 3.6 show that overall, the majority of participants felt that they would feel comfortable talking to a stranger at distances of both four feet and six feet, in both a video-mediated interaction or face-to-face. Participants generally felt uncomfortable at farther distances, particularly for a face-to-face interaction.

When considering only the data referring to the video condition, there is a trend towards more participants feeling comfortable talking to a stranger at the farthest conditions if they had viewed the smaller window. No participants who had viewed the large window reported that they would feel comfortable talking to a stranger using either the 9ft or the 12ft distance, while at least 40% of those who had viewed the smaller images reported that they would feel comfortable.

When considering the face-to-face data, there is little difference between the distributions of those in the large and small conditions. It is noticeable however, that marginally more of those who viewed the large window rated the far distances as being comfortable for face-to-face interaction than those who had viewed the small window. The relevance of this finding is that it indicates that participants who used the small window were not simply more comfortable talking to people at far distances than those who used the large window.

Taken together, these results suggest that on average, participants who viewed the small window considered the images to differ less in terms of feelings of comfort than those who viewed the large image. This can also be seen in the results of the question asking about the most appropriate distance for placing different people at – there appeared to be more variation in the results of those participants who used the small window.

Overall the questionnaire results show that participants generally felt comfortable talking to others at either the 4ft or 6ft distance, whether video-mediated or face-to-face. No participant reported that they would feel comfortable talking to a person at all the different images, and those that felt comfortable at the far distances also tended to feel comfortable at the middle distances. These results suggest a relationship between the perceived distance of a person and a feeling of comfort, with distances towards the closer end of social space being preferred. It should be noted, however, that this was a measure of expected comfort rather than actual comfort during an interaction. Participants did not actually have a conversation with someone at these distances, they were solely asked to imagine how they would feel.

3.7.2 Measurements of Images

Before each snapshot image was taken, the confederate was asked to place a ruler on their forehead. The ruler had on it a clearly marked area 100mm long. This provided an objective measure of the level of camera zoom for each selected image. Subsequent analysis of the images involved measuring the size that the ruler appeared on the screen in each image. Results were recorded for each participant and compared across conditions for each distance. The results were also compared with the size of image that would occur if the ruler had been placed at each of the five distances, as opposed to changing the zoom of the camera. This was called the "life-size answer".

3.7.2.1 Finding the Life-size answer

In order to have a life-size answer, initial measurements were taken by placing the ruler at the position of the confederate and zooming in until the image of the ruler reached 100mm (life size). From this point then the ruler was moved to positions corresponding to each distance used in the experiment. At each distance an image was taken and the length of the ruler image measured. This resulted in an objective "life-size answer" for any size of window.

This "life-size answer" is based upon adjusting the image size to be as large as in real life. As discussed earlier, there is the possibility of distance being mediated by a ratio between the relative sizes of the image and background. Should this hypothesis prove to be true then the "life-size answer" as we have stated it clearly does not represent the answer that would be expected. An answer based on the relative amount of background visible in the field of vision would be more appropriate. This is, of course, far harder to measure and was beyond the scope of this experiment.

3.7.2.2 *The order of the measurements*

In order for the experiment to be run as smoothly as possible and to avoid unnecessary distress for participants the decision was taken to run the trials in a specific order. This meant the trials going from furthest away to closest. This also aided participants when filling out the questionnaire as some of the questions referred to specific images and by having them in order ensured participants remembered them correctly. The downside was that it did mean that participants always (with one exception) went from a smaller image to a bigger image and were very much guided by where their own starting point was. Although this was not ideal, it can also be argued that a participant will always be affected by the positioning of the previous images and there seemed to be no way to ensure participants judging partly based on previous images as well as the actual feeling of distance at the time.

It is worth noting however, that the vast majority of participants did take their time over each trial individually and did not appear to be simply zooming in the image separate amounts. This was also ensured by having the distances unequal distances apart.

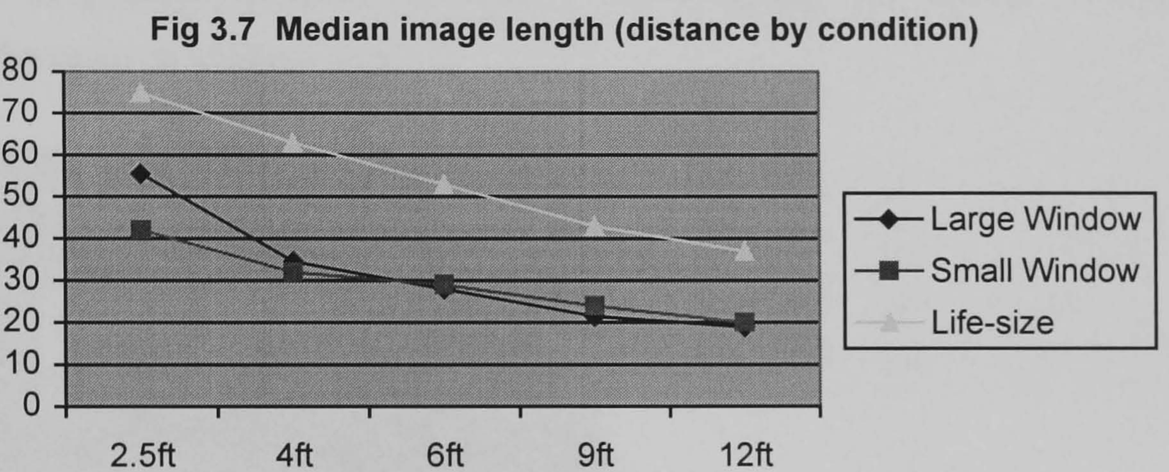
3.7.2.3 *Statistical analysis of length measurements*

This experiment was a mixed design with 2 independent variables, one within subjects (distance; 5 levels) and one between subjects (condition; 2 levels). The dependant variable was the length of the ruler in each saved image (length).

Due to the small number of participants involved, the data was slightly skewed, for this reason it is more appropriate to look at the median scores rather than the mean scores for each condition. Table 3.6 shows the median length of the ruler image for each condition and results are graphed in Fig 3.7.

Distance (feet)	Large Window	Small Window	Life-size
2.5	55.5	42	75
4	34.5	32	63
6	28	29	53
9	21.5	24	43
12	19	20	37

Table 3.6. Median ruler lengths (mm) for saved images in each window condition.



A two way (5x2) ANOVA (repeated observations on one factor) was carried out. This showed there to be a significant main effect of distance [$F(4, 72) = 66.26$; $p < 0.05$] and a significant interaction of distance and condition [$F(4, 72) = 3.06$; $p < 0.05$]. The main effect of condition was not significant overall [$F(1, 18) = 1.91$; $p > 0.05$].

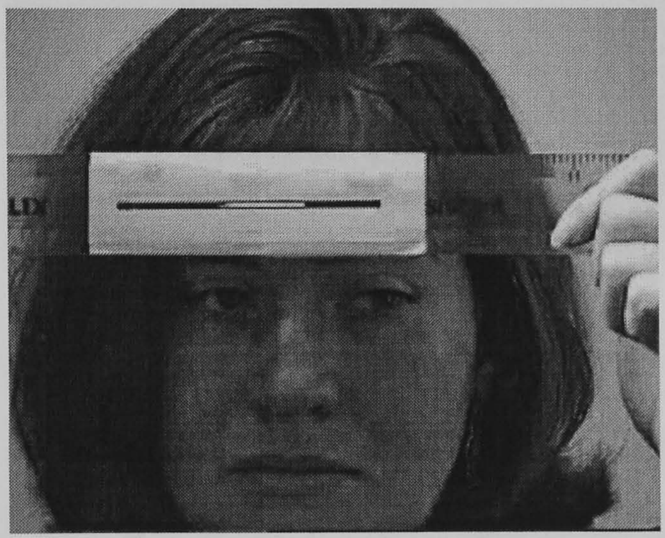
Post hoc tests of simple main effects showed that there was a significant effect of condition only at the closest distance [$F(1, 18) = 5.21$; $p < 0.05$). This result shows that to represent 2.5 feet in a video image, participants using the larger video window chose larger images than those who used the smaller window. There was no significant difference between conditions at any of the other distances.

The ANOVA also showed a strong within subjects effect of distance [$F(4, 72) = 66.26$; $p < 0.05$]. This indicates that participants chose smaller images to represent distances further away. As the live person got closer, participants zoomed in to get the feeling of closeness.

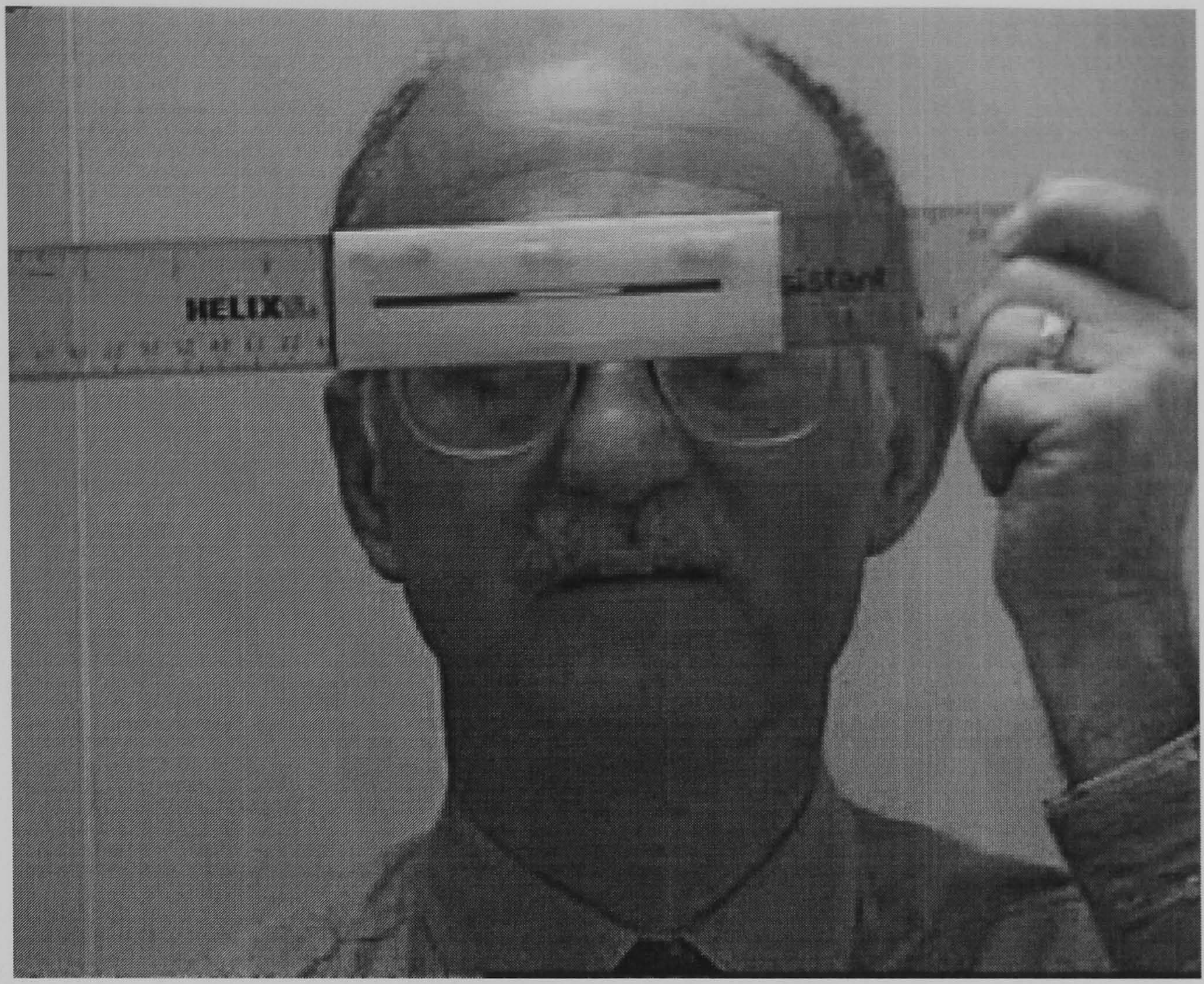
It is clear that participants estimated the image size to be under that of the "life-size measurement". Only one participant estimated images higher than the life-size measurement - effectively "larger than life" and this participant was the outlier which skewed the data. Full discussion of the implications of this result can be found later in the discussion section of this report.

3.7.3 Sample images

The following screenshots represent the median images for particular distances. It should be noted that although the image size on paper differs from that on the screen, the proportions remain the same. These images are merely to illustrate the differences in window size.



Picture 1. Median 2.5ft image for small window



Picture 2. Median 2.5ft image for large video window.

Pictures 1 and 2 represent the median images chosen for 2.5ft for each window size. Although Picture 2 has a larger image size, it can be seen that picture 1 has a larger image to background ratio.

Pictures 3 and 4 represent the median image for 4ft away on the small and the large window. This was considered to be the most comfortable image by most people.

Picture 3. Median image for 4ft with a small window.

Picture 4. Median 4ft image for large video window.

Inspection of Pictures 3 and 4 shows that the image size (length of ruler) does not change, despite the difference in the size of the window. There is of course much greater background visible in Picture 4 (the large window).

3.8 Discussion

The first aim of this research was to investigate whether or not changes in the level of camera zoom reliably caused changes in the impression of proximity of the remote individual. Whilst running the experiment it was clear that although participants found the task challenging, they had little difficulty in considering camera zoom in terms of making the individual appear closer or further away. Only one participant did not have a totally increasing image size related to increased proximity.

A related consideration is how well can individuals judge distance across a video link? The results show that there is a wide variability in participants' responses (up to 38% of the size of the window for the closest distance in the smaller window), referring to what size of image represents a particular distance. This would seem to indicate that the feeling of distance per se in a video-mediated image is not the same for all people, although the trend of appearing closer when zoomed in and farther away when zoomed out appears generally applicable.

It is important to note however, that there is no correct answer. People were asked to make the image feel as far away as the copresent person. Each individual's interpretation of this may indeed be different. Although it was possible to come up with a life-size answer of what an image should look like if it was life-size, no participants were close to selecting a life-size image as representing any particular distance. These results show that participants did not base their judgement of distance on an attempt to make an image the same size as it would be if co-present. It appears that individuals accepted that presentation by video leads to a shrinking of an image. The visual field presented in a video window is, of course, much smaller than that available to the human eye.

3.8.1 What are the determinants of perceived distance across a video link?

This question refers to the issue of which analogy people use to perceive distance within a video link. As outlined earlier there are two probable ways of considering the video window. If the video is seen as the equivalent of a shrunken (or possibly enlarged if larger than life size) field of vision (theory 1) then people should perceive distance as a function of the amount of the visual field an object takes up. If on the other hand the video window is considered to offer a partial view of a visual field (theory 2) then distance should be perceived as directly relating to the size of the object image and unrelated to the amount of visible background.

The results show that with the exception of the closest distance, there was no difference in the median object image size chosen to represent any particular distance between the large and small video windows. This suggests that perceived distance of an object is related to the absolute size of the object image, as opposed to relating to the amount of background visible. The results also showed that those images which gave the impression of any particular distance were smaller than life-size. This suggests that individuals consider the video image that they see as a blinkered vision of a visual field but also that this visual field is shrunken - the camera essentially making things look smaller than they actually are. Following this hypothesis, the size of the video window would not make any difference to the size of the image selected but the sole judgement would not be based upon the real life size.

This theory however, may require a caveat. It seems doubtful that the size of the video window does not matter at all. The relatively small range of video window sizes which could be manipulated here did not show effects. People may be sensitive to much larger variations in size. One suggestion may be that if the window is of a smaller size than can normally be seen within a certain amount of the visual field (for instance central vision) then individuals consider that the image is shrunk and shrink their perception of the size images should appear accordingly. If, on the other hand, the size of the window reaches a certain threshold, for instance the size of central vision, then individuals may perceive that images should be life size. Finally, should a video window be larger than a certain threshold, well into peripheral vision for instance, then people may perceive images of objects as requiring to be larger than in real life.

The finer points of this hypothesis such as the exact number of levels, sizes images should appear and so on are all speculation at present but the basic model of the theory - of a perception of video being a representation of a full but shrunken field of vision, and distance being judged primarily by perceiving the size of the object image - appears to be supported by the data collected in this experiment.

3.8.2 How does the size of the video window affect people's perceptions of distance?

The size of video window has been shown to affect people's perception of distance in terms of the amount of zoom. As explained above, perceived distance appears to be regulated by the image size of the object in focus. This in turn is a function of the size of the video window and the amount of zoom on the camera. To keep image size constant, as the size of the window decreases, the zoom on the camera must increase.

Other effects of the size of the video window were apparent in the questionnaire results. Those who had experienced the small window were more mixed in their opinions as to where was most appropriate to place a call centre agent. Those who used the small window were more likely to place the agent on one of the two far distances (9ft and 12ft).

Similarly on the issue of comfort when talking to a stranger across a video link, those who used the small window were significantly more likely to feel comfortable talking to a stranger using the far images (9ft and 12ft) than those who had used the large window.

Both of these results suggest that the larger video window leads to a greater feeling of presence from the video image. This is in line with research by Lombard & Ditton (1997) who found that screen size was positively correlated with feelings of presence in television. The results also suggest that the proximity effects (feelings of comfort, feelings of appropriateness) are far less pronounced in the small window. Consequently a window of 100mm may be too small to elicit the proximity effects reported in the literature on face-to-face interaction.

3.8.3 Do people feel a video image to have similar proxemic effects to face-to-face interaction?

The distances used were deliberately chosen to relate to the levels of space proposed by Hall (1966). Hall claims that personal space extends to 4ft, so the closest distance (2.5 feet) falls well within personal space, where people will usually feel uncomfortable when talking to strangers, 4ft is on the borderline, 6ft is towards the close edge of social space, 9 feet is towards the far edge of social space and 12ft is at the very far edge of social space.

The results support this taxonomy of space and its relative effects. When questioned about how comfortable they would be talking face-to-face to a stranger at each distance - at the closest distance, most (85%) would feel uncomfortable talking to a stranger as would 90% of people feel uncomfortable talking to a stranger if they were sitting at either of the two far distances. Most would feel comfortable talking to a stranger at both the 4ft (85%) and the 6ft (65%) conditions.

When participants were asked the same question of comfort but using each of the video images rather than the face-to-face situation, the results followed the same pattern but were less pronounced. The results showed that 75% of participants would feel comfortable using the 4ft or 6ft condition, and only 15% would feel comfortable using the closest video image. Of the far image conditions, only 10% considered a conversation with a stranger would be comfortable using the 9ft or 12ft conditions. When using the video image, the results were dependent on the condition used. Fifty percent of those using the small window thought they would feel comfortable using the 9ft video image (and 40% for 12ft), whereas no-one did using the large window.

One other result from the questionnaires supports the previous proximity work of Hall. More people said they would video conference with a friend using a closer image than they would with a stranger, than said they would use the same image or place the stranger closer. This suggests that even across a video link, friends would wish to interact at a "closer" distance than strangers.

Ninety percent of people also stated that they considered the closest image (that which they equated to 2.5 feet) imposing. This agrees with Reeves et al. (1992) who stated that personal space could be invaded by a video mediated image.

These results all suggest that video-mediated images can indeed transmit information which appears similar to the proxemic information which exists in face-to-face communication. It appears that merely zooming the video camera in may lead to the impression of closeness. The questionnaire data suggest that people would feel differently about talking to individuals portrayed in different video-mediated conditions. If people feel differently about the interaction then there is also the possibility that they may behave differently.

3.9 Conclusions

This preliminary study investigated the effects of changing the video camera zoom in a general way. Participants were not interacting across a video link and so no conclusions can be drawn about how people would behave when conversing with a remote collaborator depicted with various levels of camera zoom. This study has however shown that proximity as a form of non-verbal communication can be transmitted by changes in camera zoom. Consequently the potential exists for proximity to affect behaviour in a video-mediated interaction.

The conclusions that can be drawn from this study are that individuals can interpret changes in camera zoom as altering perceived distance. Judgements of perceived distance appear to be related to the physical image size of the object per se. The effects do not seem to relate to the relationship between the amount of visual space occupied by the object and the amount of background visible.

A larger video window seems to lead to a greater sense of presence. Although the image size of the individual does not change across window conditions, having more background has shown to lead to more consistent feelings of comfort or discomfort. The feelings of comfort and discomfort caused by the video images are also similar to those caused by a face-to-face individual.

This study has shown that when using video-mediated communication technology, users may experience differing impressions of proximity. The next question to be explored is whether such differences in perceived distance may lead to differences in the manner in which individuals interact when using a video link.

4 Chapter 4. Study 2: The Effects of Visual Proxemic Information in Video Mediated Communication

As was outlined in Chapter 2, manipulating proximity in a face-to-face interaction can result in differences in the way strangers communicate with each other. These differences manifest themselves both behaviourally, in the way people interact, and attitudinally, in the impressions individuals form of each other. Since Study 1 (reported in Chapter 3) showed that changing the zoom on a camera can lead to different feelings of distance of a remote individual, there are two questions which the next Study seeks to answer. Can changes in the perceived distance of a remote interlocutor affect the way in which an individual will interact with them? And if so, do these differences correspond to the effects demonstrated in face-to-face interactions?

4.1 Face-to-face proximity

Social psychologists have known for years that proximity is an important source of non-verbal communication information. Social rules appear to exist for how closely people interact, depending upon their relationship (Allegier & Byrne, 1973), culture (Vaksman & Ellysman, 1979), personality (Leary, 1983) and situation (Alexander & Rudd, 1981). If an interlocutor infringes these social rules, research has shown that there are certain well-defined compensatory behaviours being performed by the other individuals in the interaction (reviewed by Capella, 1981, 1983, and 1994).

4.1.1 Categories of Space

Hall (1966) categorised interaction distances for seating as intimate space (6-18 inches), personal space (18 inches - 4 ft), and social space (4-12 feet). For each particular type of interaction there is an appropriate distance which corresponds to one of these distances. For example, talking to a close friend may be done within personal space, whereas talking to a stranger will usually be done within the social space.

Of these categories of space, the most heavily researched is personal space. The term was actually first proposed by Sommer (1959), and has been regarded as a type of territory (Scheflen, 1970). Sommer defined personal space as:

> "[Personal space is] an area with an invisible boundary surrounding the person's body into which intruders may not come." Sommer (1959).

4.1.2 The Invasion of Personal Space

There is a large body of research that has investigated how individuals' personal space can be invaded and the compensatory behaviours that occur when such an invasion happens. This research has demonstrated that personal space is roughly circular but with more space in front (Argyle, 1988). Cook (1978) reported that people would consider their personal space being invaded at a greater distance when in a head-on orientation. Vine (1975) found the same individual space shape exists for animals.

An example of a compensatory behaviour is the adjustment of body position. Several researchers have shown that when someone stands closer than would normally be expected then individuals tend to adjust their body position, either by retreating or by changing their orientation (Becker & Mayo, 1971; Efran & Cheyne, 1973; Sawitsky & Watson, 1975). This has been explained by researchers as a method of keeping others from intruding into one's personal space.

Researchers have also shown a link between eye gaze and personal space. Researchers have shown that as a person approaches, individuals decrease the amount of eye gaze they use (Patterson, 1977; Argyle & Dean, 1965). Other compensation behaviours which occur when an individual approaches too close include: decreasing the average duration of their responses (Johnson & Dabbs, 1976), decreasing the number of affiliative responses in a conversation (Capella, 1986), increasing time taken for an individual to leave the situation (Poyatos, 1975).

All of the findings mentioned above consist of the behavioural effects of closer than normal proximity. As well as knowing what happens however, it is necessary to find out why such behaviours happen. To investigate this, researchers have studied the subjective effects of breaking the normal rules of social proximity.

Paterson (1968) found that confederates who stood closer were perceived as warmer and liking the subjects more. Rosenfeld (1966) also showed that people who wish to seek approval will also stand closer than normal. When considered alongside the result that people stay longer in an interaction with a person who stands closer, it would appear that the effects of breaking the social rules of proximity can be rather positive.

The reason for this can only be speculated but could possibly be explained by applying the theory of cognitive dissonance (Festinger, 1957). The dissonant situation is that although you do not know the person who stood near you, you let them stand close. This dissonance can be resolved by thinking "I like this person".

There is also reason to believe that close proximity may increase powers of persuasion. Mehrabian & Williams (1969) showed that when people stood closer than normal, others perceived them as attempting to be more persuasive. The study also showed a high correlation between intended and perceived persuasiveness. Other factors that are known to increase persuasion include liking or involvement (Argyle 1988, p.260), louder faster voices, looking more and touching during an interaction. As well as all being linked to increased persuasion, these all bear a similarity to what happens naturally at a closer proximity. Although more persuasion is perceived to be being used by someone who chooses to stand close, this is not a guarantee of success, Albert and Dabbs (1970) found that persuasion attempts worked better at less close distances.

4.1.3 Culture

Support for cultural differences in spatial behaviour comes from several studies (e.g. Watson & Graves, 1966; Forston & Larson, 1968; Noesjirwan, 1977) all of which show that there are specific cultural rules for how close people should be when they interact. Hall (1968) hypothesised specific differences between "contact cultures" (such as Arabs and Latin Americans) and "non-contact cultures" (such as British and Northern Europeans) citing evidence such as closer distances for interacting in contact cultures.

Sussman & Rosenfeld (1982) carried the work further and showed that proximity of individuals while speaking is both culture specific and situation specific.

Venezuelan students tended to sit close together while Japanese sat further apart when talking in their own languages. When they spoke English however there was no difference between the two groups and they sat at a comparable distance to American English speakers. This suggests that proximity is part of the culture and individuals are adaptable, rather than something which is learned by an individual and kept as an individual trait.

So far all the studies reported have been carried out in co-present situations. In this thesis Study 1 showed that video-mediated communication technology had the potential to relay images so that people appeared at different distances. It is not known however, whether this would lead to the same behavioural and psychological differences as face-to-face proximity. The advent of video-mediated communication produces possibilities of new types of interaction occurring that may have the potential to transmit proxemic information without it in fact being "real". If proxemic information can be transmitted then it may be possible to simulate one-way proxemic effects; one person can appear close while another appears far away. Depending upon the effects of transmitting proxemic information it may be possible to tailor video-mediated technology for particular purposes for particular interactions.

4.1.4 Designing a task for studying proximity in Video-Mediated Communication

The aim of the present study was to investigate the effects of an alteration of proxemic information transmitted across a video link. To do this a task had to be developed that would be suitable for analysing both the behaviour and attitudes of participants taking part in a video-mediated interaction. A number of other factors were required to be present in the design of the task, corresponding to the factors which are known to be affected by proximity in face-to-face interactions.

The task had to have a clear goal, so that there were definite start and end states which participants worked towards. The task should also have a social element to it and involve some measure of persuasion. It was also considered to be desirable to make the task as realistic as possible.

The resulting design was the development of the Financial Advice task. This involved a banking scenario in which participants were role-playing the part of a customer seeking investment advice from a financial advisor. The image of the financial advisor was varied by altering camera zoom so that the image either appeared very close or far away, while the financial advisors view of the customer was kept constant. This was considered to be a realistic task with the developments in modern banking.

4.2 Method

4.2.1 Participants

There were 30 participants (15 males and 15 females; age range 17-32, median age = 19) all of whom were undergraduates and graduates from the University of Glasgow. All participants were native speakers of English and came from a Western culture. No participant had ever used any banking by video conferencing equipment before, and no-one rated themselves as very inexperienced with computers. Participants were told they were going to take part in a psychology experiment investigating banking by videoconferencing. It was explained that they would be audio and video taped for research purposes but that these would be treated as strictly confidential. Participants were also informed that the experiment had no time limit and that at the end there would be a questionnaire. Participants were paid £5 for their participation. All participants signed their consent.

4.2.2 Scenario

Participants were told to role-play the part of a customer seeking financial advice on the best way to invest £5000 that they had just inherited. The customer had gone to the local branch of their bank to seek advice. The bank was a fully automated one and they had sat down at one of the kiosks and used the video conference link application to find the part offering financial advice.

It was explained to the participant that they would be able to see, hear and speak to a financial advisor throughout the task via a live video link. There were four investment options available to the customer, each of which was accompanied by a page of information, presented on-screen. Participants were informed that there was no time limit or correct answer and the task would end when they had chosen an option to invest their money.

4.2.2.1 Investment Options

There were four options for participants to choose between. These were all based on real investment possibilities and the information about them was taken from genuine banking leaflets. Options differed in terms of risk, potential returns, availability of the capital and whether it was taxable or not.

BONUS ACCOUNT: A high interest but taxable bank account offering saving with 120 days notice required for penalty free withdrawals. No minimum time limit for investment but an annual bonus paid if fewer than two withdrawals made in a year. Minimum investment of £5000 pounds.

TESSA: A 5 year, tax free savings plan offering access only to 80% of the interest each year but guaranteed high interest with no risk. Yearly limits to the amount that can be invested - only £3000 in the first year.

SHARES: A high-risk option capable of making the greatest capital gains but guaranteeing nothing. Subject to capital gains tax. No limits on time, although recommended as a long-term savings plan.

PEPs: A lower risk option than the shares but still based on the stock exchange. Opportunity for higher gains than the no risks accounts but lower than the shares. Tax-free. No time limit, but considered a medium to long term investment. Subject to handling charges.

The information that accompanied all these options was more extensive and can be found in Appendix D.

4.2.3 Roles

4.2.3.1 Financial Advisor (FA)

A postgraduate psychology student played the role of the financial advisor. She was instructed to attempt to subtly persuade customers towards accepting the bonus account on the basis that this would make the most money for the bank. The advice must appear to be independent however, as it was most important for the customer to be happy with the advice given. The FA was also asked to look into the camera as much as possible in order to give the impression of looking at the customer. The financial advisor did not know the purpose of the study or which condition they were taking part in for any particular trial. To allow the FA to get used to the experiment, 4 practice trials were run before the experiment was started.

4.2.3.2 The Customer (C)

Participants played the role of the customer; they received instructions immediately prior to the experiment outlining the scenario. They were a 24-year-old industrial designer, reasonably comfortably off, with a stable job who had just inherited five thousand pounds and was looking to invest it. The remainder of their character's situation was left up to them. The reason for leaving this open ended was to stop participants searching the instruction sheet for answers if the FA asked them a question and also to make the task more interactive for them.

4.2.4 The Conditions

The experiment was a between subjects design with one independent variable. The independent variable was perceived distance (proximity) of the financial advisor to the customer, with two levels, close and far. The close condition involved zooming the camera until the view of the FA was virtually just the head, with a minimal amount of background visible. In the far condition, the camera was zoomed until the view was of the FA's head and upper torso, which allowed plenty of background to also be visible.

It is worth noting that it was only the customer's view of the FA that differed. The advisor always saw the customer at a distance in between that of the two conditions.

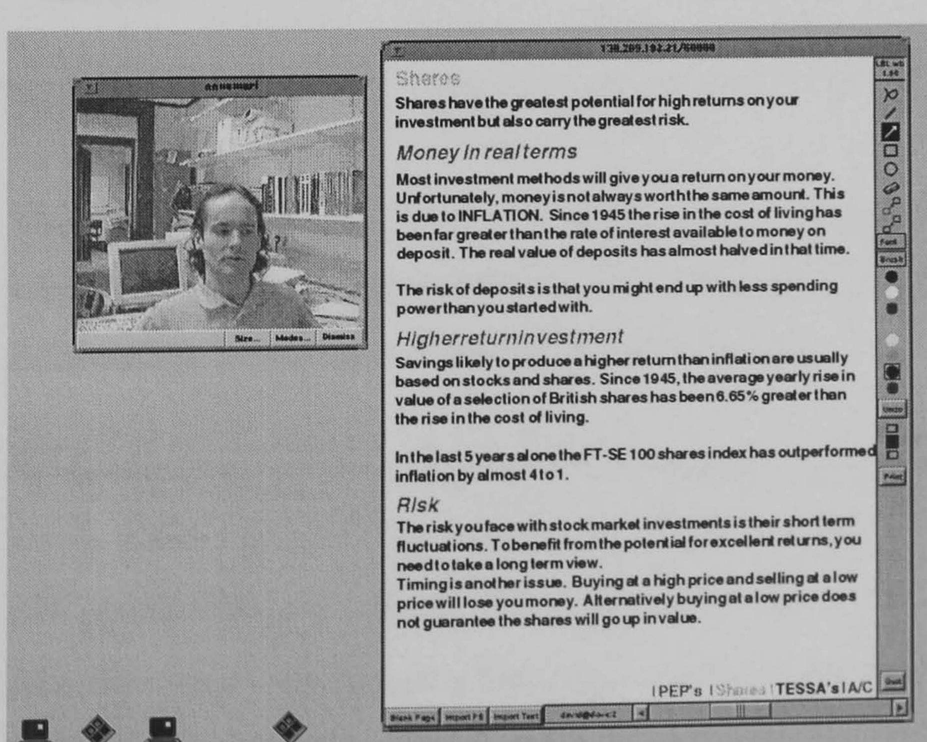
The reason for this was that the study was only interested in the effects of perceived distance on the customer, and so only visual information given to the customer was altered. Had the visual information been altered for both parties then the financial advisors behaviour may have differed between conditions and thus any effects found could be due either to the visual information directly or as a result of the reaction to changes in each others behaviour.

It is true to say that there may be an issue of generalisability to real world interactions, based upon the fact that we are creating an asymmetrical interaction in terms of proximity. It was, however, felt to be more important to isolate the effect on one individual of the other appearing a certain distance. The FA is effectively working to a script and it is therefore unlikely that their behaviour will alter much depending on proximity. Were this work to be generalised to compare with spontaneous social interaction as opposed to this more formalised scenario then this issue would require to be addressed. It should also be pointed out that experiments using confederates also work to scripts, rather than acting spontaneously so the effects caused by reacting to behaviour changes are not being investigated.

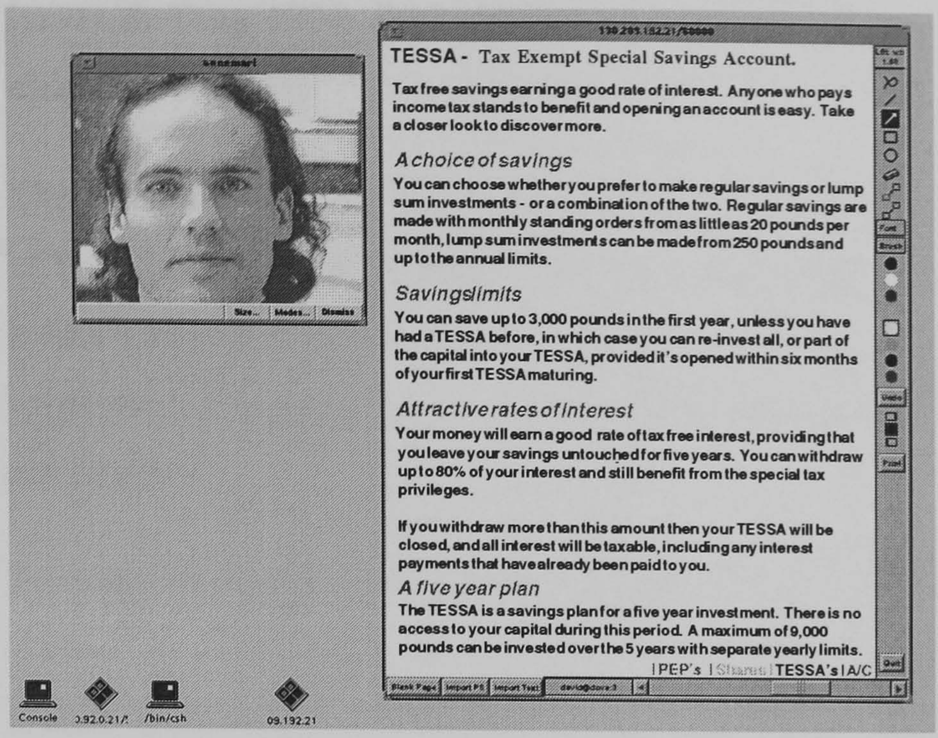
4.2.5 Screenshots

The following screenshots were presented on a 17-inch monitor.

Far condition



Close condition



4.2.6 Technology

Program Name	Colloq. Prog Name	Version No.	Source
wb	White Board	1.60	Lawrence Berkeley Labs (LBL)
rtvc_display (adapted)		1.2.1, REV=4.0.7	SUN Microsystems; code by E. Grattan of MCG, University of Glasgow.

Table 4.1 Programs run during experiment

The customer saw a screen with four pages, which was provided by the shared whiteboard. The whiteboard also allowed the financial advisor to add to the information screens and highlight areas in the text.

Situated to the left-hand side of the whiteboard in the upper half of the screen there was a video window that used the SunVideo system. The video window was 384 x 288 pixels (115mm width x 85mm height) in size. Video and audio were transmitted directly rather than across a network, ensuring broadcast quality frame rates with no appreciable audio lag. Frame rate and audio lag have been shown in the past to have severely deleterious effects on the communication abilities of participants in a videoconferencing system (e.g. Cohen, 1982; Johannsen, 1984; Tang & Isaacs. 1993: Isaacs & Tang 1994; Whittaker, 1995) and thus high quality video was considered essential.

The cameras used were JVC videomovie GR-AX 55E. They were placed to the side of the screen as close as possible to the video window in order to get as close to eye contact as possible. Realistic eye contact is not possible in most commercial video conferencing applications. Audio was provided via headphones with boom microphones.

Perceived distance was altered by changing the zoom on the camera. Very close was defined by cropping the video window around the face and "very far away" allowed a view of the whole of the upper torso. It is worth acknowledging that there is more to perceived distance than simply visual cues, but in the experiment no differences in sound were made, and thus only visual proxemic information was altered. While it would be scientifically interesting to investigate altering the sound as well to give a greater proxemic effect it was still considered worthwhile to investigate the purely visual cues.

4.2.7 Questionnaire

A questionnaire was completed by participants at the end of the experiment and was designed to investigate three areas – prior experience, and participants' impressions of the task and the technology. Most questions were designed on a five point Likert-type scale, although there were also some open-ended questions and some questions that involved the ranking of options.

Prior experience of both videoconferencing and computers in general was monitored to ensure any effects found were not simply due to level of experience with the equipment. Questions in this section included how experienced they were using computers in general, and video conferencing equipment.

The section of the questionnaire asking about the task included questions about how the individuals felt about the realism of the task, qualities of the financial advisor (e.g. likeability, sincerity, helpfulness), and quality of the advice given.

The section on the technology investigated how participants felt about the technology itself, including whether they considered they would be willing to use it in real life, and how it compares to other methods of banking.

A copy of the questionnaire can be found in Appendix C.

4.2.8 Dialogue Analysis

The main aim of this study was to gain an insight into how customers would react to video links which manipulated the image of the remote advisor. To measure the effects on participant behaviour in terms of communication processes, dialogues were recorded, transcribed and analysed. This provides a microanalysis of the nature of the interaction. Of specific interest within the dialogues were:

TURNS: A turn is defined as the time from when one person speaks until another speaks. A pause, no matter how long, does not automatically end a turn. Turns are the most widely researched feature of dialogue and have in the past - like proxemic information - been shown to be related to familiarity. Boyle, Anderson & Newlands (1994) found that when interacting with a familiar person, individuals will tend to use both more turns and shorter turns. Backchannels (see below) are considered turns but do not end the speaker's turn when they are uttered by listeners. Overlaps on the other hand do end the speaker's turn and are counted as turns themselves. Thus a period of speech by one person during which there are 2 backchannels uttered by the other person has a total of 3 turns in it. Had the backchannels been interruptions however, there would have been 5 turns.

BACKCHANNELS: Backchannels are brief responses indicating agreement, understanding or attention, which do not disrupt the flow of the speaker. In this study they were taken to be "mhm", "uhhuh", and "mmm" when said on their own without disrupting the speaker. It should be noted though that sometimes these same words could be used as answers to questions and then would not be considered backchannels.

OVERLAPS: Overlaps were considered to have taken place if one or more words of the second speaker's contribution were perceived to have overlapped the first speaker's contribution. This is the same definition as Boyle et al (1994) used, although they point out that the definition of overlaps and interruptions has been far from consistent across the literature.

WORDS: Words were counted and included all utterances, including "uhhuh", "mhm", and "mmm".

These measures include all of the dialogue aspects that Capella (1994) summarises as being affected by proximity in face-to-face situations.

4.3 Experimental Hypotheses

This study addresses two separate issues. The main experimental question is whether or not changes in the zoom of the camera lead to differences in the behaviour of individuals interacting via video conference technology. If there are behavioural effects of camera zoom then these are hypothesised to be similar in nature to those documented in the face-to-face literature. There is also a practical aspect to this study that is interested in how far away someone should appear to be for tasks such as giving financial advice across a video link.

If video-mediated communication does preserve proxemic information and participants treat it in the same way as they would in a face-to-face interaction, then it is hypothesised that the same pattern of effects as documented for face-to-face interaction will occur. The following hypotheses can then be made.

HYPOTHESIS 1. When the financial advisor appears close then the interactions will take longer.

HYPOTHESIS 2. When the financial advisor appears close then they would be perceived to be more persuasive by the customer.

HYPOTHESIS 3. When the financial advisor appears close then the customer will form a more positive impression of the financial advisor.

HYPOTHESIS 4. There should be more turns in the close condition.

HYPOTHESIS 5. Turns should be shorter in the close condition.

4.4 Questionnaire Results

Two questionnaires (both in the far condition) were only partially completed. Percentages are given for all the data that were available. This corresponds to between 13 and 15 responses in the far condition and 15 in the close condition. A full list of results from the questionnaires is in appendix D.

The questionnaire had three sections, each of which consisted of various numbers of question and question type. Each question was analysed separately to investigate whether there was any significant difference between participants' responses in the close condition and those in the far condition. No question showed a significant effect of condition. Some of the questions did however give an interesting insight into how participants experienced the task. The results are summarised by section.

4.4.1 Experience

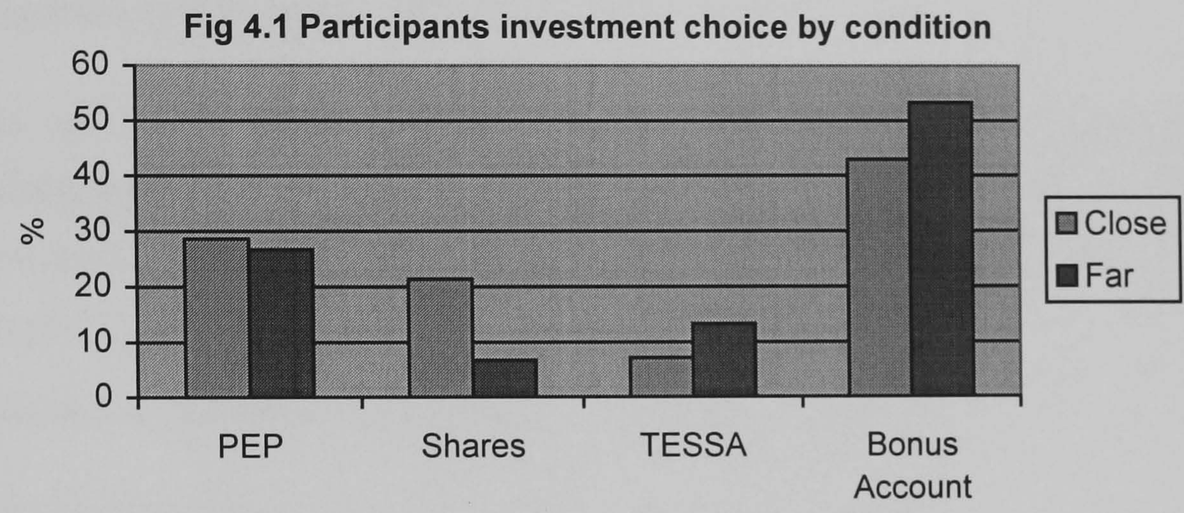
The results of the questionnaire showed that there was no significant difference between conditions in the prior experience of participants with technology. Overall 46.7% of participants rated themselves as having some experience with computers, while the same number rated themselves as either experienced or very experienced. Only 3.3% considered themselves inexperienced. Only 13.8% of participants had ever used desktop video conferencing before, these individuals were split evenly between conditions. No participants reported any technological difficulties in completing the task.

4.4.2 The Task

The goal of the task was for the customer to choose an investment option. The advisor attempted to subtly persuade the participant to choose the bonus account. It was expected that persuasion may have been one of the areas affected by the difference in perceived distance. The results are summarised in Table 4.2 and Figure 4.1.

	Close	Far	Overall
PEP	28.6	26.7	27.7
Shares	21.4	6.7	13.3
TESSA	7.1	13.3	10.0
Bonus Account	42.9	53.3	46.7

Table 4.2. Percentage of Participants making each choice in each condition.



The results show that all of the options were chosen at least twice in each condition. The bonus account was chosen most often (46.7% compared with 27.7% for PEP) but this did not differ between conditions.

These results suggest that the difference in proximity did not lead to any difference in the successfulness of the advisor's persuasion. Other questions asked participants how persuasive they believed the advisor was trying to be. Table 4.3 summarises the responses to the question "How much persuasion did you feel that the finical advisor used?"

	Close	Far	Overall
A lot	6.7	0.0	3.3
A little	60.0	40.0	50.0
None	33.3	60.0	46.7
Don't know	0.0	0.0	0.0

Table 4.3 How much persuasion did the FA use?

The results show that two-thirds of all the participants in the close condition perceived some persuasion being used compared to only 40% in the far condition. Although this failed to reach significance when using a Chi-Square test, it does suggest a trend towards the perception of persuasion being greater in the close condition. It is possible that with a larger rating scale for responses this may have shown a significant effect of condition. The trend is in the same direction as the results of Mehrabian & Williams (1969) who showed that close proximity led to an increased perception of persuasion.

As well as being asked about their perception of the amount of persuasion used, participants were also asked how much they trusted the advice given to them. Table 4.4 summarises the results of the question "How much did you trust the advice given to you?"

	Close	Far	Overall
Trust a lot	40.0	20.0	30.0
Trust a little	46.7	66.7	56.7
Neither trust nor distrust	6.7	6.7	6.7
Distrust	6.7	6.7	6.7
Distrust a lot	0.0	0.0	0.0

Table 4.4 How much did you trust the advice you were given?

Table 4.4 shows that the majority of participants trusted the advice that they were given (86.7%). Of most interest is that in the close condition there appears to be a trend towards the advice being trusted slightly more than in the far condition. As with the persuasion results, this failed to reach statistical significance and thus any speculation must be treated with caution. When put together with the persuasion result presented in Table 3 however, an interesting pattern appears to emerge. Although there is a trend towards greater persuasion being perceived, there is no corresponding drop in the level of trust – if anything there appears to be a related increase in trust. Taken together it appears that the impression may be one of participants perceiving the advisor trying to be more persuasive but for positive reasons and thus trusting the advice given more.

Of the other questions which related to the task there were no trends suggesting any effect of condition. They do however give an insight into how the advisor was perceived by participants. Other overall results of interest included that only 1 participant found the task of making their choice difficult. Two-thirds of participants considered the task realistic or very realistic with only 10% considering it unrealistic. This suggests that participants had little difficulty in completing the task and understanding the nature of the scenario.

The financial advisor was considered to be helpful (80%), unbiased (93%) and likeable (70%). All of these results suggest that the impressions of the financial advisor were extremely positive. Again, none of these results were affected by condition. It had been expected that more positive evaluations of the financial advisor would be found in the close condition than the far condition but this was not the case. It is possible that the failure of the analysis to reveal any difference between conditions was due either to a ceiling effect or the small Likert-type scales. Until that can be tested however, it must be concluded that the level of the zoom on the camera does not lead to any appreciable difference in the appraisal of one's conversational partner.

Taken together, the questionnaire results seem to suggest that the different conditions led to little effect on the impressions the advice or the advisor. There may be a trend towards more persuasion being perceived in the close condition but this does not lead to any corresponding decrease in the level of trust that participants have in that advice. Participants also had positive evaluations of the advisor regardless of the condition they experienced.

4.4.3 The Technology

The questionnaire was also used to assess participants' impressions of the technology. Items on the questionnaire asked individuals to rate how well suited the technology would be for carrying out particular tasks, and what aspects of the simulation were good or bad.

The chosen conditions represented extremes of camera zoom, it is difficult to conceive the camera being zoomed in any more for an interaction. and the camera could not be zoomed out further. Participants were asked whether they felt the image was too close or too far. As expected no participants in the close condition rated the image as appearing too far away or too close in the far condition. The questionnaires showed that in 6 of the 15 close trials, the agent was considered to be too close and in 1 of the 13 far trials, the agent was considered to be too far away. This suggests firstly that participants were comfortable with considering the image in terms of close and far away, and secondly that the range of images that participants consider acceptable may be quite large. For a sizeable minority of users, appearing very close (probably invading personal space) would appear to be less acceptable than being very far away in an image.

Participants were asked whether or not they felt that they would have come to the same decision had they conducted their discussions using a different medium. The results showed that 40% of participants felt it unlikely that they would have come to the same decision had the advice been given over the telephone. Only 7% believed it unlikely to have come to the same decision had the advice been given face-to-face. These results seem to suggest that people felt that the video was an appropriate medium for this type of task but that the telephone was not.

The other questions about the technology, despite again yielding no significant differences between proxemic conditions did offer interesting overall results about people's readiness to use such technology to do their banking.

Participants were asked to rank different media according to their suitability for seeking the type of advice offered in this experiment. Face-to-face was the first choice for 76% of all participants, with leaflets (no human contact) the first choice for 10% of respondents. Video tended to be rated second and consistently outscored telephone banking. This suggests that video is less attractive than speaking to someone face-to-face, although some people prefer to make decisions on their own.

When asked how far people would travel to get face to face advice as opposed to video advice being available in their local branch, 24% responded that they would be unwilling to travel further to get face-to-face advice. Of the others, 21% would travel for up to 15 min, 35% up to half an hour with 21% willing to travel for more than half an hour to obtain face-to-face advice. These results show that face-to-face does remain the ideal choice for most of the participants but that video could prove to be an acceptable substitute for many people if it was more convenient.

When asked about what types of transactions they would be willing to carry out using the technology they had just experienced, some transactions were felt to be more appropriate than others are. The results are shown in Table 4.5.

Transaction	**Close**	**Far**	**Overall**
Apply for a loan	64.3	40.0	51.7
Apply for a mortgage	35.7	6.7	20.7
Discuss an overdraft	71.4	66.7	69
Open an account	85.0	80.0	82.8
Close an account	100.0	73.3	86.2
Customer complaints	57.1	60.0	58.6
Discuss financial advice	85.7	73.3	79.3

Table 4.5 The percentage of respondents willing to carry out financial services using video-mediated communication technology.

It appears from Table 4.5 that customers consider applying for mortgages as fairly inappropriate for video whereas opening and closing accounts were seen as the most appropriate. An interesting feature to note is that for all transactions with the exception of customer complaints, more of those in the close condition were willing to use the technology. This did not represent a statistically significant difference however.

4.4.3.1 Problems with the technology

The final item on the questionnaire was an open ended one for respondents to give any other comments on the simulation. Despite it never having been mentioned anywhere in the questionnaire, 27.6% of respondents commented that there was a problem with eye contact. This is an extremely high level of response considering it was unprompted and shows that the lack of eye contact is still a major issue with this type of system. Even though every effort was made to make the camera as close to the screen as possible and the advisor was asked to look into the camera as much as possible, this appears to have caused occasional problems.

One participant who commented that the advisor "was never paying attention" provides anecdotal evidence of the problem with eye contact. When pressed further on this comment it transpired that the participant had been making thumbs-up gestures to answer questions but that the advisor hadn't noticed them "even though she was looking straight at me". Only when it was explained that when she appeared to be looking straight him, she was in fact looking at the camera so couldn't see him, did the participant understand what had been going on. This illustrates how powerful the notion of being looked at can be.

Overall the questionnaires give a useful insight into the feelings of participants with various parts of the task and the technology but has failed to show significant differences between the attitudes of participants who interacted with a financial advisor who appeared zoomed in close and a financial advisor who appeared zoomed out.

4.5 Results of Dialogue Analysis

All the dialogues were recorded and transcribed. Due to technical difficulties, two of the transcripts had to be rejected. This left fifteen in the close condition and thirteen in the far condition. Once transcribed, analysis of the dialogues was carried out on the following structural measures of dialogue – turns, words, backchannels, overlaps, words per turn, time.

There were two variables in the experiment. Proximity was a between subjects independent variable with two levels – close and far. Role was a between subjects independent variable with two levels – financial advisor or customer. The analysis was complicated by the fact that the FA was always the same person, but it was the effect of condition on the customer that was of most interest.

Each dialogue measure that was considered as a dependent variable was measured in three ways. The effect of proximity on the total dialogue, the effect of the proximity on the customer, and the effect of proximity on the financial advisor, with most dialogue measures providing a score for each of these (exception being time which is not divisible between roles). Since the financial advisor was always the same individual and was not directly influenced by the change in proximity, the primary source of interest was the effect of proximity on the customer. For this reason, the effect of proximity on each dependent variable for the customer only was analysed separately. Data were analysed using independent samples t-tests.

4.5.1 Time Taken

The time taken to complete the task was measured in seconds. Each interaction began by the financial advisor saying "Hello I'm your financial advisor, how can I help you?" and ended when the financial advisor said, "Thank you very much, goodbye". An independent samples t-test was computed to compare the mean length of interaction between conditions and showed a trend towards trials taking longer in the close condition but this fell just short of significance [$t(26) = 1.80$, $p < 0.1$]. Table 4.6 shows the mean time taken for each condition.

Mean time (s)			
Close	Far	t(26)	Sig.
609	477	1.80	ns ($p < 0.1$)

Table 4.6 Time taken (seconds) to complete task (by condition).

The overall mean time to complete the task was 548 seconds (9.08 minutes). It had been hypothesised that interactions would last longer in the close condition. The trend shown by the results is in this direction but the results must be considered inconclusive due the failure to reach statistical significance.

4.5.2 Customer Turns

	Mean No. of turns (FAR)	Mean No. of turns (CLOSE)	t(26)	Sig.
Customer	52.9	73.8	2.06	$p < 0.05$

Table 4.7 Mean number of customer turns per dialogue by condition.

The mean number of turns spoken by the customer was 52.9 in the far condition and 73.8 in the close condition. An independent samples t test showed this to be a significant difference [$t(26) = 2.06, p < .05$] indicating that the close condition led to the customer taking more speaking turns than the far condition.

Turns have been used by some researchers as a measure of the length of an interaction (e.g. Boyle, Anderson & Newlands, 1994). In the present study however the length of an interaction and the involvement in an interaction are very different things, due to the nature of the task. Boyle et al. used a task that involved a constant exchange of information between the two participants (the Map Task), the present task can be completed equally well by someone who just listens and reads as someone who interacts a lot.

4.5.3 Backchannels

	Mean No. of b/c (FAR)	Mean No. of b/c (CLOSE)	t(26)	Sig.
Customer	5.00	7.67	0.85	ns

Table 4.8 Mean number of customer backchannels between conditions

An independent samples t-test indicated no significant difference between conditions for the number of backchannels spoken by the customer [$t(26) = 0.85, p > .05$].

Backchannels are turns in their own right, but do not attempt to end the other speaker's turn. They are very short expressions of feedback such as "mhm". As there were significantly more turns in the close condition, it is appropriate to standardise the dialogues for length. To do this the mean percentage of turns that were backchannels was analysed. The results showed that in the close condition, backchannels accounted for 9.00% of customer turns compared with 8.03% in the far condition. An independent samples t-test showed that this difference was not significant [$t(26) = 0.33, p > .05$]. The results also showed that the percentage of turns that were backchannels ranged from zero to 26.5%, indicating a wide variation in participants' use of backchannels.

These results indicate that the number of backchannels that the participant said during the interaction was not affected by changing the appearance of the financial advisor. This is true whether considering the number of backchannels per se or the number of turns spoken which are backchannels (rate of backchannel production).

4.5.4 Overlaps

The mean number of customer overlaps was recorded and analysed. A customer overlap is an instance of the customer beginning a turn whilst the financial advisor is still speaking. Table 4.9 summarises the mean number of customer overlaps in each condition.

	Mean No. of overlaps (FAR)	Mean No. of overlaps (CLOSE)	t(26)	Sig.
Customer	6.23	13.53	2.47	$p < 0.05$

Table 4.9 Mean number of overlaps between conditions.

The results of the independent samples t-test summarised in Table 12 show that there was a significant effect of condition on the mean number of customer overlaps. Participants' turns overlapped those of the advisor more frequently in the close condition than in the far condition [$t(26) = 2.47$, $p < .05$].

Since there were more customer turns in the close condition, it is important to know whether this increase in the number of overlaps is merely a function of the increased number of turns in the close condition. To investigate this, the number of overlaps as a percentage of the total number of turns spoken was analysed and is summarised in Table 4.10.

	FAR	CLOSE	t(26)	Sig.
Customer	10.63	17.15	2.78	$p < 0.05$

Table 4.10 Mean % of turns that were overlaps between conditions.

Table 4.10 shows that significantly more of the customer turns in the close condition were overlaps than in the far condition [$t(26) = 2.78$, $p < 0.05$]. This indicates that participants overlap more often in the close condition than in the far condition. Furthermore, this increase in the number of instances of overlapping speech is not merely a factor of the increase in the number of speaking turns that the customer makes.

4.5.5 Words

While an analysis of the number of turns shows how often a participant spoke, the total number of words spoken by the customer gives a guide to how much they said during a dialogue. Table 4.11 summarises the results of an independent samples t-test comparing the mean number of words spoken by participants in each condition.

	FAR	CLOSE	t	Sig.
Customer	252	453	2.29	$p < 0.05$

Table 4.11 Mean number of words spoken per trial between conditions.

Table 4.11 shows that there was a significant difference between conditions [$t(26) = 2.29$, $p < 0.05$]. This indicates that the customer said significantly more when the advisor appeared close than when the advisor appeared far away.

This result has shown that as well as speaking more often - as indicated by an increase in the number of turns - participants also say more on average during the task in the close condition. In order to investigate if these results are caused by a function of one another it is necessary to analyse the mean number of words per turn spoken by the customer. Calculating the mean number of words per turn effectively standardises the dialogues for the number of turns spoken, thus showing if the increase in words is merely due to the greater number of turns being spoken. The results are summarised in Table 4.12.

	FAR	CLOSE	t	Sig.
Customer	4.79	5.82	1.72	ns ($p < 0.1$)

Table 4.12 Mean number of words per turn spoken by the customer between conditions.

The independent samples t-test summarised in Table 4.12 indicates that there was a trend towards participants produced longer turns in the close condition but it fell short of significance. This suggests that in general the increase in the number of words being spoken by participants in the close condition is due primarily to the increase in the number of turns being spoken by the customer in the close condition.

4.5.6 Share of the dialogue

Considering the number of words spoken by one individual as a percentage of the total words spoken gives a measure of the share of the dialogue spoken by each individual. This was analysed and is summarised in Table 4.13.

	FAR	CLOSE	t	Sig.
Customer	20.97	26.11	2.38	$p<0.05$

Table 4.13 Share of the dialogue (%) by role between conditions.

Table 4.13 shows that the customer increased their share of the dialogue from just under 21% in the far condition to just over 26% of the total words spoken in the close condition. This increase was significant [$t(26) = 2.38$, $p < .05$] and shows that the customer has a greater share of the dialogue in the close condition than in the far condition.

4.5.7 Summary of Results of Dialogue Analysis for Customer

Table 4.14 summarises the results of the dialogue measures for the customer.

	FAR	CLOSE	Sig.
Mean time taken (seconds)	477	609	ns ($p < 0.1$)
Mean No. of Turns	52.9	73.8	$p < 0.05$
Mean No. of backchannels	5.00	7.67	ns
Mean % of turns that were backchannels	8.03%	9.00%	ns
Mean No. of overlaps	6.23	13.53	$p < 0.05$
Mean % of turns that were overlaps	10.63	17.15	$p < 0.05$
Mean No. of words	252	453	$p < 0.05$
Mean words per turn	4.79	5.82	ns ($p < 0.1$)
Mean share of dialogue	20.97%	26.11%	$p<0.05$

Table 4.14 Summary of the Dialogue Measure for the Customer.

The results of the dialogue analysis on the participants part of the dialogue show that when the advisor appears close, participants say more, speak more often, and overlap with the financial advisor's speech more. Interactions also tend to last longer than when the advisor appears further away.

Taken together, these results show that participants appear to be more interactive in the close condition than in the far condition. This general impression was also created when listening to the dialogues. In the far condition there were many more interactions where the customer was sitting and listening to the financial advisor without speaking much - basically speaking when spoken to. This compared to interactions in the close condition which were often very interactive, with customers giving a lot of feedback and appearing more engaged in a two-way discussion.

4.5.8 Analysis of Transcripts by Role

Since this is a new task, and the scenario differs largely from the majority of experimental tasks used, it is important to explore the nature of the task communication. To investigate the difference between the two roles in the task. data was analysed with role being the independent variable, across conditions of proximity. Table 4.15 summarises the results.

	Fin. Advisor	Customer	t (54)	Sig.
Turns	57.64	64.11	0.92	ns
Backchannels	1.00	6.43	3.44	$p < 0.01$
Overlaps	9.11	10.15	0.47	ns
Overlaps as percentage of turns	14.72	14.12	0.29	ns
Words	1083	360	7.36	$p < 0.01$
Words per turn	19.04	5.34	14.34	$p < 0.01$
Percentage of dialogue	76.29	23.72	31.90	$p < 0.01$

Table 4.15 Communication breakdown by role across conditions.

Table 4.15 shows that the differences inherent in the task are that the financial advisor speaks more overall, in terms of both the number of words spoken and the share of the dialogue. The financial advisor also indulges in longer speaking turns but does not speak more often (turns). This is because in a dyadic interaction the number of number of turns is virtually bound to be equal. The customer makes more backchannels (feedback that does not interrupt a turn) than the financial advisor. There is no difference in the overall number of incidence of overlapping speech made by the financial advisor or the customer.

4.5.9 Effect of proximity on the financial advisor

The present experiment was designed to investigate the effect of proximity on the customer. It was expected that there would be no difference between conditions in the manner in which the financial advisor interacted with the customer. This is because the financial advisor always had the same view of the customer (midway between the close and far conditions) and that the financial advisor was effectively working to a script. One individual's contribution to a dialogue can not be considered independently of the other person however. Although the financial advisor did not experience any difference in the appearance of the individual they were talking to, they did experience a difference in the conversational style of the customer, caused by the difference in the customer's view of the financial advisor.

To investigate whether the difference in the customer's behaviour affected the financial advisor, analyses were conducted which considered the difference between the mean scores for the financial advisor in the close condition and in the far condition. For each set of data, a t-test for independent samples was carried out. The reason an independent samples test was selected was the interactive nature of a dialogue. The nature of any one individual's speech may be largely dependent upon the way that the other individual speaks, thus although the financial advisor was the same in each trial, there may have been different influences which were equivalent to making the advisor behave as an individual in each trial. The results of these analyses on each of the dialogue measures are presented in Table 4.16.

	Close	Far	t (26)	Sig.
Turns	66.27	47.69	2.17	$p < 0.05$
Backchannels	1.53	0.38	1.95	ns
Overlaps	10.80	7.15	1.23	ns
Overlaps as percentage of turns	14.83	14.60	0.07	ns
Words	974.08	1177.20	1.18	ns
Words per turn	18.35	19.83	0.81	ns
Percentage of dialogue	73.90	79.04	2.39	$p < 0.05$

Table 4.16 Effect of condition on the dialogue of the Financial Advisor.

The results in Table 4.16 show that the financial advisor has a greater share of the total dialogue in the far condition, and that they make more turns in the close condition. No other dialogue measure shows any effect of condition for the financial advisor.

Both of the measures which do show an effect of condition for the FA are those which are structurally linked to the dialogue of the customer. In a two-person dialogue, the percentage of words spoken by one individual is directly related to the percentage of words spoken by the other individual. Similarly, although the number of turns need not technically be even, there can only be differences caused by backchannels (the only instance where a turn does not interrupt another turn) and since these are rare, the number of turns will always be fairly even. Since this is the case then an increase in the number of turns spoken by the customer inevitably leads to an increase in the number of turns spoken by the financial advisor.

These results have indicated that the difference in condition did affect the participants - who experienced the difference in appearance first hand. The change did not lead to any major impact on the style of the financial advisor who retained a similar style of speech throughout interactions in both conditions.

4.5.10 Questionnaire and Dialogue Analysis

Although the questionnaire data failed to show any significant differences between the attitudes of participants between conditions, there was some variation between participants across the whole experiment. This variability can be used to investigate whether there are any determinants other than the appearance of the advisor over the course of the experiment that may lead to differences in behaviour of participants. To do this, the results of some questionnaire items were compared with the major dialogue measures in which perceived proximity appeared to have an effect - turns, words and overlaps.

The item of most interest was the question concerning whether participants felt that the financial advisor appeared to be at an inappropriate distance. It is conceivable that should a participant feel unhappy with the appearance of their conversational partner that they would then change their behaviour. There were 6 participants who considered the advisor to be too close, one who considered them too far away and 21 who reported feeling the advisor appeared neither too far nor too close. An independent samples t-test was computed using "happiness" as an independent variable. As it was likely that considering the advisor too close or too far would be likely to cause opposite results with regard to an effect on the dialogue, only those participants who stated that they considered the advisor to appear too close were considered in the unhappy group. This group was compared with the other 9 participants who took part in the close condition. The mean results of this analysis are presented in Table 4.17.

	Unhappy (too close)	Happy (Close)	t(13)	p
Turns	81.33	68.78	0.756	ns
Overlaps	14.00	13.22	0.149	ns
Words	473.17	439.67	0.207	ns

Table 4.17 Results of independent samples t-tests for happiness.

The results of Table 4.17 show that there were no significant differences between the unhappy (those that considered the FA too close) or happy participants for any of the main dialogue measures which were examined.

Although any conclusions that may be drawn from this analysis must be cautious due to the small sample of unhappy participants (6) they do not suggest a link between a conscious dislike and a behavioural effect. Those participants who considered the advisor to be too close did not show any behavioural difference to those that did not consciously regard the advisor as too close.

This was the only questionnaire item that was analysed with respect to the dialogue measures for this study. Future studies would attempt to investigate the link between subjective impression and behavioural performance more clearly by designing questions that may differentiate participants on more scales.

4.6 Discussion

The results of this study indicate that altering the appearance of a remote individual in a videoconference can lead to differences in the manner in which participants converse. While the questionnaire appeared to show little effect of perceived proximity on participants' subjective impressions of the task or technology, the dialogue analysis has shown that there were significant differences in the way that people behave.

When looking at these results it is important to remember that *the perceived proximity of the other person only differed for the customer*. The view seen by the FA remained the same in all conditions and was in between that of the far and close conditions of the customer. Bearing this in mind, the fact that only the customer was showing significant differences in the dialogues is further evidence that they are caused by effects of perceived proximity.

Results initially can be summed up by saying that when the financial advisor appeared very close, the customer made more interruptions, overlapping speech and indeed said more than when the financial advisor appeared far away. This in turn tended to increase the length of the interactions without altering subjective feelings about the individual or the interaction (as measured by the questionnaire).

Although the only change between conditions was the appearance of the advisor, to be able to conclude that the cause of the behavioural differences was one of proximity, the results of the experiment must be compared to the established face-to-face literature. This literature led to five hypotheses being made, which must be re-evaluated in light of the results.

HYPOTHESIS 1. When an individual appears closer then the interactions will take longer

The study by Poyatos (1975) showed that when people stand closer than is normally considered appropriate then individuals take a longer time to leave the situation. In the present study there was a trend towards interactions taking longer in the close condition, although this did fall just short of significance. The results showed that the mean length of interaction was nearly two minutes shorter in the far condition than in the close condition. The trend was in the predicted direction, and although not statistically significant due to reasonably high variability, the trend suggests that this hypothesis may yet remain inconclusive.

HYPOTHESIS 2. In the close condition customers would choose the bonus account more often as the FA is more persuasive.

The results showed no difference between conditions of proximity for which option was chosen so consequently this hypothesis must be rejected. The results also showed that only 46% of participants in total chose the bonus account, which would seem to rule out any floor or ceiling effect. It would appear that although the bonus account was selected more than any other option, the appearance of the financial advisor did not aid or inhibit the financial advisor's attempts at persuasion. Persuasion has not been shown to be any more successful when the persuader appears closer than when they appear further away.

Although the success of the persuasion attempt showed no significant difference between conditions, there is also a related question of the perception of persuasion. Many of the face-to-face studies which have investigated persuasion have showed that people who stand closer are perceived as attempting to being more persuasive (e.g. Mehrabian & Williams, 1969) but make no comment about the actual success of the persuasion attempt. In the present study, the results of the questionnaire hinted at a difference in the level of persuasion being perceived by the participants. The questionnaire data showed that 60% of participants in the far condition perceived no persuasion compared with only 33% in the close condition. Future experiments should aim to extend this work by increasing the rating scales for perceiving persuasion and recording both effects of persuasion and the perception of persuasion - two very different things.

It should also be noted that if there was an increase in the perception of persuasion in the close condition, this was not linked to a lack of trust of the advice being given. The levels of trust followed a similar relationship to the levels of perceived persuasion. It is possible that in the close condition participants perceive more persuasion being used but believe it is for the right reasons so trust the advisor more. Further investigation using more sensitive questionnaire items would be required before firm conclusions along this line could be drawn.

HYPOTHESIS 3. The participants should form a more positive impression of the advisor when she appears closer.

The results from the questionnaires showed no significant differences between conditions on any of the questionnaire questions. The questionnaire included items about likeability, helpfulness, pushiness, and the independence of the advice. None of these showed a significant effect of condition and consequently this hypothesis must be rejected. It appears that either perceived distance does not affect the impression of the remote individual or that in this study either the task or the questionnaire was not sensitive enough to capture any differences that may exist.

If it is accepted that the proxemic information in this study has not led to any differences in the conscious impressions formed by participants, then the question of why not must be addressed. It is plausible that for attitudes to be altered would require more proxemic information than that needed to alter behaviour. Hall (1966) suggested that one of the reasons that personal space differed from social space was the fact that at the closer distance, other senses come into play such as the ability to touch. In this experiment there was only visual proxemic information available - no olfactory or audio cues to proximity were present and touch was certainly not possible. It seems probable that the decrease in proxemic information may lead to a decrease in the strength of the proxemic effect.

HYPOTHESIS 4. There should be more speaking turns in the close condition.

Results of the dialogue analysis clearly showed that when the advisor appeared close then participants spoke more turns than when the advisor appeared further away. In some experiments the number of turns has been used to define the length of an interaction rather than time (e.g. Boyle, Anderson & Newlands, 1994; Doherty-Sneddon et al., 1996). In the present study, participants could gain information in two different ways – they could read the information on the screen or they could talk to the financial advisor about it. In this situation, the number of speaking turns could be regarded as a fairer reflection of the amount an individual was engaged in the conversation because of the option to sit and read the whiteboard information rather than talk. It is possible to have two totally different styles of interaction lasting the same length of time where one individual talked little and took most of their information from the whiteboard while the other individual talked a lot but read little. This difference would be shown up by the difference in the number of turns spoken. This result therefore could be considered evidence in favour of the hypothesis that a remote individual appearing closer may lead to an increase in the length of an interaction (where length is defined in terms of the amount of speaking done).

HYPOTHESIS 5. Turns should be shorter in the close condition.

Analysis of the mean number of words per turn showed a trend towards customers saying more words per turn in the close condition. This result leads us to reject the hypothesis that turns are shorter when one's partner appears close.

When considering why this should be the case, again the ability to gain information without speaking may be important. In tasks based solely on conversation, there is likely to be a trade-off between the length of the speaking turns and the number of speaking turns made. An increase in the number of turns made is likely to be accompanied by a decrease in the mean number of words per turn made. In the present task, however, an increase in the number of speaking turns is more reflective of an increase in the overall amount of speech being spoken and consequently does not need to be accompanied by a decrease in the average number of words per turn.

4.6.1 An Increase in Speech

The results showed that when participants were interacting with a remote individual who appeared to be closer they made more speaking turns, the number of words per turn was unaffected and consequently the number of words spoken by the customer was greater in the close condition than the far condition. Similar results were found by Doherty-Sneddon et al. (1997), when comparing the effect of eye contact in video-mediated communication. They found that when eye contact was preserved, participants used significantly more words and turns but their conversations lasted an equal amount of time.

Doherty-Sneddon et al. concluded that this pattern of results showed a decrease in efficiency in the eye contact enabled condition. At first glance it may appear that the present study has found a similar decrease in efficiency when the financial advisor appeared close. It should be noted however, that no instruction was given to the participants to complete the task as efficiently as possible. The results suggest that naturally close perceived proximity may lead to an increase in speech in conversation. This does not mean however, that in a task where both partners desired efficiency, the close condition would perform less well.

Before any conclusions can be drawn about the nature of the increases in the amount of speech, it is necessary to study the content of the dialogues. At this stage it is unknown whether the differences in the amount of speech are accounted for by social chat, extra acknowledging, more questioning of the advisor, more feedback or simply a general increase across the board. In order to investigate this, a system of dialogue content analysis, Conversational Games Analysis, was applied to the dialogues and shall be described in the following chapter.

4.6.2 Evaluation of the Task

With every new task which is devised it is necessary to evaluate the results of the experiment to see whether the task was successful in meeting its original aims. The main aims of the financial advice simulation were to provide a scenario that was easy to understand, realistic and supported dialogue between the participant and a remote advisor. The results have shown that participants did feel the task to be realistic and easy, with only 10 percent of participants considering the task unrealistic and one participant rating it as difficult. The amount of speech that was generated by the task was also a manageable size. Dialogues had a mean length of approximately 120 turns, which proved to be enough to show significant differences between conditions while not tiring the participants too much.

The experiment was also designed to have a more social feel to it rather than being purely task-oriented. Participants seemed to have no difficulty role-playing and found the task quite enjoyable. Although very few of the dialogues included much non-task oriented talk, several included instances of social chat within the task scenario.

The task itself allows the participant to say a great deal or not depending on whether they prefer to talk or simply listen and read the whiteboard. Consequently the task leads to the financial advisor saying much more than the customer and holding a greater share of the dialogue. This in turn leads to there being more backchannels being spoken by the customer - presumably to provide more feedback during long periods of the financial advisors speech, the like of which are not often present in customer dialogue.

The main advantage of this task appears to be that it provides a framework for studying the effects of aspects of video-mediated communication on a single individuals speech. Because the financial advisor is reactive to the customer but is effectively working to a script, different participants behaviours can be compared without any concerns over the effect of the other interlocutor. Because they are working to a script, however, few conclusions can be drawn as to the effect of the participants behaviour on the confederate.

4.7 Conclusions

The present study has shown that visual proxemic information can cause similar behavioural effects to that of face-to-face proximity. Much of the face-to-face literature regarding the effects of proximity is therefore relevant to modern applications of video-mediated communication. The effects that changing the zoom on a video camera produce are essentially similar to those of face-to-face interactions but on a lower level - most likely due to the unimodal proxemic information available compared to the multimodal information available in a co-present interaction.

A financial advisor appearing very close tends to decrease the formality of an interaction, with the customer taking a greater speaking role. The customer makes more speaking turns, more words and more interruptions when the advisor appears closer. It is possible that this may be in some way connected to the similar results found with research into familiarity.

This study has also provided evidence of the advantages to be gained in employing a multifaceted approach to data collection. This research shows the worth of analysing dialogue as a tool for examining subtle psychological effects, which may be difficult or impossible to examine using questionnaires.

Since it is now known that when the financial advisor appears close, the manner of participants' speech changes, the next stage is to investigate whether the structural changes in dialogue relate to consistent differences in the content of the dialogue.

5 Chapter 5 . Analysing Dialogue Content: The Application of Conversational Games Analysis to Study Two

Chapter 4 showed that a difference in the presentation of an image of a remote individual can lead to differences in the way that a video-mediated interaction will occur. It was shown that when the financial advisor appeared zoomed in there was a tendency for participants to speak more and more often, than when talking to an advisor who appeared far away. There were also more instances of overlapping speech in the close condition than the far condition. These results refer to structural differences in the manner in which participants participated in a dialogue, but offer no insight into what function the extra talk is providing or why there are more interruptions.

The aim of the present study was to investigate the content of the dialogues to examine whether any of the structural differences found corresponded to any particular functional differences in content between dialogues.

5.1 Analysing Content

In the previous chapter, communication analysis was restricted to structural measures of dialogue. What individuals actually said was of no consequence as to the analysis, as measurements were made only of the manner in which it was said. The purpose of verbal communication however, is the transfer of meaning, and consequently it seems reasonable to make some analysis of the content of the verbal communication. Both style and the outcome of communication are mediated by content (Rutter, Stephenson & Dewey, 1981).

Researchers interested in interpersonal interactions have used a wide variety of coding schemes to deal with the content of verbal communication. The diversity of meaning which occurs in speech, and the wide range of situations in which speech can be used and studied accounts for the diversity in schemes. Each scheme seeks to categorise speech, but is strongly influenced by the purpose for which they were developed.

The first decision that had to be made regarding the present study was whether or not to use a scheme which had already been developed or to devise a new coding scheme specifically for this task. Before considering some of the possible schemes it is useful first to consider the main elements of a coding scheme.

There are several characteristics that combine to make a good coding scheme. Pavitt (under review) outlines 5 requirements for a good categorical coding scheme. The requirements that Pavitt identifies are simplicity, exhaustive and mutually exclusive categorisation, validity and reliability.

5.1.1 Simplicity

A coding scheme needs to be as simple as possible. It should be so, however, without sacrificing the significant distinctions that the researcher intends it to address. Coding schemes are simple if they have relatively few categories, and also if those categories are defined clearly. This is because few categories are relatively easy for researchers to learn and to use. Clear definitions allow users to know exactly which behaviors do and do not fit each category.

Like transcription, coding is an extremely labour intensive task. For this reason, it is important that coding is as simple as possible. In general, an easy to use coding scheme will remove as much of the subjective element to coding as possible, by having clearly defined categories, removing as much ambiguity as possible.

5.1.2 Exhaustive and Mutually Exclusive Categorisation

In a good coding scheme, categories should be both exhaustive and mutually exclusive. This means that every behaviour must fit into one category or another, but also that each behaviour should fit into only one category, never into two or more.

Some coding schemes permit multifunctionality within their coding schemes, allowing a researcher to put a statement into more than one category (Pavitt, under review). In these schemes, exclusivity is maintained by creating a new category for when two specific behaviours are occurring. For instance, there may be three categories to cover two behaviours.

It is assumed that everything that is said in a dialogue is said with a purpose, and consequently carries some communication information. For this reason, everything in a dialogue should be able to be coded in one of the categories. Each utterance should also only be able to be coded into only one of the categories or else the reliability and validity shall both suffer.

5.1.3 Validity

As with all experimental methods, both validity and reliability are important. A scheme is valid if it measures what the researcher intends it to measure, however there are several types of validity.

Face validity. Face validity is the need for the scheme to make intuitive sense, both to its users and those who evaluate it. In order to make sense, a scheme should define its categories clearly, and these categories should measure what the researcher intends them to measure.

Construct validity. A system has construct validity if the categorization of behaviours within it is consistent with a theory that the system intends to test. Kerlinger (1964) discussed both face and construct validity. For reasons of construct validity, Bakeman & Gottman (1997) recommend that coding schemes embody the hypothesis under investigation.

Representational validity. A coding system may also have representational validity if it makes the same types of distinctions among categories that the average person would make (Poole & Folger, 1981). In other words, the average person and the coding scheme would agree on how to classify particular behaviours.

Any coding scheme requires face validity. The other kinds of validity apply to specific types of coding schemes. For instance, if a classification system measures behaviours that people naturally interpret, it needs to have representational validity. If a scheme is testing a theoretical perspective then it must have construct validity (Pavitt, under review).

5.1.4 Reliability

A scheme is reliable if the categorisations made by one coder match those made by others. Further, a scheme is reliable if one coder can categorize the same behavior the same way at two different times. In order to achieve this it must be able to be shown that the coding distinctions can be understood and applied by people other than the coding developers (Carletta et al., 1998). A classification system needs to have clarity and ease of use, and those who use the system must also have a sufficient amount of training and practice in order for the system to be reliable.

Guetzkow (1950) distinguished between two types of reliability that a coding scheme must have. The first type he distinguished was categorizing reliability. This term means that two coders, or the same coder at different times, can place the same statement in the same category. Guetzkow's other type of reliability was unitizing reliability. This means that two coders, or the same coder at different times, agree on the "size" of the behavior. By the word "size," Guetzkow meant where the behavior begins and where it ends. If two coders do not agree on the size of the unit, their separate codings of the same conversation will yield vastly different results.

Thus, in order to be successful, behavioral coding requires that coders be reliable both in their categorization and in their unitizing.

5.1.5 Generalisability

One element that Pavitt does not mention, but is important for our purposes is generalisability. Although inevitably developed primarily for a specific type of communication situation, a good coding scheme will be generalisable to dialogues created in other situations. The suitability of a coding scheme for any particular task will be largely dependent on its ability to generalise from the task for which it was designed.

Comparison of coding schemes can be done with consideration to these factors, although the relative weight that should be given to each factor will differ depending on the context of use. The advantages of devising a new scheme will be a high level of specificity and presumably ease of use but proving reliability and validity will be more difficult than using a pre-existing scheme.

The goals of the present study were to gain an insight into how the function of participants' utterances differed between conditions in the financial advice task. Specifically, a coding scheme was sought that could explain why there was more conversation by participants when interacting via a video link where the interlocutor appeared zoomed in. Several coding schemes were considered for this purpose, as was the idea of devising a coding scheme specifically for this study.

5.1.6 Designing a new scheme

As stated earlier, many researchers investigating the content of dialogue devise their own method of doing so. The main advantage of this is that the coding scheme can be designed specifically for a particular task or to test a specific hypothesis. There is little or no pressure on the researcher to develop a classification which is particularly generic.

The disadvantages of devising a new scheme, are that it is much harder to be able to demonstrate high levels of reliability and validity. Consequently, any conclusions based on the results as found by this new scheme are open to questions about interpretation. There is a risk of merely finding what you set out expecting. For this reason it was considered ideal to use a pre-existing scheme, if a suitable one – with proven reliability and validity could be found.

5.1.7 Examples of Coding Systems

5.1.7.1 Interaction Process Analysis

Interaction Process Analysis (Bales, 1950) was developed as a way of measuring leadership in face-to-face groups. He devised a general purpose scheme for encoding dialogues in small groups no matter what the composition of the groups. Although this method of coding was designed for, and is consequently most suited for, the analysis of group communication, there is no reason why this cannot be applied to dyadic interactions. The scheme consisted of 12 measures, categorised on three levels. These are presented in Table 5.1.

Socio-Emotional Leadership Measures		Task-Oriented Leadership Measures	
Positive Reactions	**Negative reactions**	**Questions**	**Attempted Answers**
Showed solidarity	Showed antagonism	Asked for a suggestion	Gave suggestion
Showed tension release	Showed tension	Asked for an opinion	Gave Opinion
Agreed	Disagreed	Asked for orientation	Gave orientation

Table 5.1 Categories of utterance in Interaction Process Analysis (Bales, 1950).

The basis behind this scheme was that Bales believed that the success of a group depended on two factors: How well it can solve the tasks facing it (the task function) and how well it can keep its members satisfied with the group (the socio-emotional function). Socio-emotional communication was considered to consist of positive or negative comments, while task oriented communication was based on either asking questions or attempting to provide answers. The third level of distinction attempted to capture the ways in which these types of communication can be presented.

The emphasis in the scheme is based on the effect of the communication on the group. Coding is done from the perspective of the group – essentially "what did an utterance add to the group?". Other coding schemes are based from the perspective of the speaker – "what did the individual intend that utterance to be?".

This distinction between the aim of an utterance and the impact of an utterance is crucial for the successful analysis of communication. This is returning to the point raised by Bakeman & Gottman (1997) that the coding scheme should embody the hypothesis under investigation. If the aim of the analysis is to understand the effects of communication, then the coding scheme should be focussed on the impact for communication, as in the case of Interaction Process Analysis. Alternatively, if the aim of the analysis is to understand how an individual's communication is affected by a variable, then the scheme should be based on the intention of the communication by the individual.

For the present study, the aim of the analysis was to understand how perceived proximity affected the communication behaviour of an individual. For the reason outlined in the previous paragraph, Interaction Process Analysis was considered less than ideal, as it is focussed on the impact of a variable on the communication of the group.

The advantages of using IPA are that it has been widely used and is fairly straightforward, alternatively however there are also a couple of disadvantages. Although still widely quoted and used, the reliability of this scheme has been variable (Waxler & Mishler; 1966). It is also worth noting that categories are not mutually exclusive, an utterance which conveys both emotional information and task based information must either be double coded, or placed in one or other category.

5.1.7.2 Conference Process Analysis

Conference Process Analysis was developed by Morley & Stephenson (1977). This was applied to the effects of technology on human communication and was used in a clinical setting. Utterances are coded along three scales resource, mode and referent. Resource is the activity being performed of which nine were outlined (procedure, settlement point, limit, positive and negative consequences of proposal, other statements about outcome, positive and negative acknowledgment, and information). Mode referred to the method of performing the activity which could be one of four (accept, reject, seek, and offer). Finally the referent was to whom the activity applies (including: none, self, person, other, party, opponent, both persons, and both parties). A similar method was used by Labov & Fanshel (1977) who analysed every utterance in terms of its propositional, dialogue and social function.

CPA has a high level of reliability and construct validity (Pavitt, under review) but has not been widely used. Consequently its suitability for generalising to the financial advice task is questionable. CPA was designed for a negotiation situation, and although the financial advice task was designed to have a degree of negotiation involved, the financial advice task would not require all of the referent categories due to its dyadic nature. Such a system seems overly specific for coding a dyadic interaction.

Other coding methods which were rejected included Discourse Analysis (Sinclair and Coulthard, 1975) and Conversational Acts (Traum and Hinkleman, 1992) which do not have proven reliability and generalisability. Speech Act Theory (Austin, 1962; Searle 1969, 1976) has made an important contribution to linguistics by analysing speech in terms of the actions it performs. However, Speech Act Theory was only intended to be applied to constructed, not spontaneous utterances (Schiffrin, 1994), so was considered unsuitable.

Olson, Olson, Carter & Storrosten (1992) developed a system for use to code design meetings. Their scheme involved classifying each utterance from one of 10 categories. They also used a binary scale for each to note whether the utterance was a clarification or not. This scheme had high reliability but its goals are a long way from those of the present task. Many categories would become obsolete and many may need to be altered to suit the financial advice task.

5.1.7.3 Conversational Analysis

A method of discourse analysis which has focussed on the way individuals carry out conversational activities is Conversational Analysis (Sacks, 1972; Sacks, Shegloff and Jefferson, 1974; Taylor & Cameron, 1987). This is not so much a coding scheme, with rigid categories into which the coder allocates utterances, as a style of coding. It is more an approach to analysing dialogue and has been used to investigate the processes which occur in conversation.

Examples of the processes which have been studied include turn taking (Sacks et al, 1974), opening and closing sequences of turns in conversations (Schegloff & Sacks, 1973) and the making of self-corrections in conversation (Schegloff, Jefferson & Sacks, 1977). More recently it has also been used in an applied setting to aid the teaching of novice dieticians in taking a successful history from patients (Tapsell, 2000).

One of the problems with Conversational Analysis is that there are few explanations of how to do good CA from a methodological perspective (ten Have, 1986). This is primarily because the methods need to be based around the situation and problem being studied, rather than specified beforehand.

Conversational Analysis is a useful tool for understanding the nature of conversation. The present task however was interested less in the general aspects of the type of conversation that occurs in the financial advice task per se, but is concerned more with the differences between conversations in each condition. For this reason, and the fact that CA is a general approach rather than a coding scheme, Conversational Analysis was rejected for the present study.

5.1.7.4 Conversational Games Analysis

Conversational Games Analysis is a coding scheme which aims to provide a general system for coding conversation. The scheme should be applicable to many different situations and thus be less influenced by the nature of the task. Whereas some researchers have been concerned with what is being spoken about in any utterance, CGA is concerned more with how the information is being shared. The CGA scheme is based upon the premise that the dialogue is an interactive process whose aim is to share information. Much of the work that preceded CGA was in the field of Artificial Intelligence (Power, 1979; Houghton & Isard, 1979) where the aim was to create computational models of software robots attempting to achieve goals by coordinating information.

A key difference between CGA and IPA is that CGA is concerned with the aim of the speaker as opposed to the understanding of the listener. This assumes dialogue to be proactive as opposed to reactive. Intention is more important than the belief of what was intended.

This system of analysis has been used quite extensively to explore dialogue content for problem solving tasks is Conversational Games Analysis (CGA). The scheme was developed by Kowtko, Isard and Doherty-Sneddon (1991) to code the functional use of utterances in task-oriented dialogues. There are two functional levels, moves and games. A move is an utterance or sequence of utterances from one speaker that embodies the same function. This can take the form of a single word such as "yes" to answer a question, or it could consist of several sentences - to describe a building for instance. A game is a sequence of moves which are used to accomplish some goal. While a move involves only one speaker, a game usually will involve more than one. A game may for instance involve one speaker asking a question e.g. "do you have a watch", the other answering e.g. "yes", and the original speaker acknowledging the answer e.g. "right". This game consists of three moves.

There are two levels of analysis in coding a dialogue. Conversational games can include any number of utterances but refer to the sharing of a piece of information. A conversational game begins when a speaker attempts to achieve a new communicative goal such as giving an instruction or asking a question, and ends when that goal is achieved or abandoned. Conversational moves, on the other hand, are utterances with a particular purpose. A number of conversational moves may make up a conversational game. This is similar to the ideas of Clark & Schaefer (1987) who talk about conversational contributions which include both presentation of information and reaction to it.

Carletta et al (1996) also make provision for a third level of analysis called transactions, which are "sub dialogues which accomplish a major step in the participants' plan for accomplishing the task". For the financial advice task, analysis at the level of transactions was not conducted. This decision was taken because the task itself was a fairly heavily structured one. Participants accomplished the task in a fairly well defined manner and always went through the same stages, often led to a large extent but the financial advisor. The transactions involved in the conversation would be to discuss each option in turn and move on to the next one, there was little opportunity for any differences in the number of transactions that would occur in a dialogue. The aims of the present experiment were to investigate how conversations differed under different circumstances. As such, the task was more of a framework, which served to encourage conversation than the focus of the research in itself. Analysis at the level of both moves and games was considered to be of interest because they embodied the hypotheses in question. It is analysis at the level of moves and games that will be most likely to demonstrate the differences in the manner of participant's communication.

CGA consists of 12 different categories of move, which are divided into two categories, based on their functional level - initiating moves and response moves. Initiating moves begin a game and response moves occur during a game. A brief description of the communicative function of each move is given below with examples taken from the original studies from which CGA was designed (Kowtko et al., 1992). These studies were using the Map task.

5.1.7.4.1 Initiating moves

INSTRUCT. A direct or indirect instruction to be done immediately or shortly.

Example: "You then go down south two inches."

CHECK. Checks self-understanding of a previous message or instruction by requesting confirmation, either directly or indirectly.

Example: "So you want me to go down two inches?"

QUERY-YN. Yes-No question asking for new or unknown detail about task (not clarifying instructions)

Example: "Do you have a rockfall"

QUERY-W. As Query-YN but an open ended question.

EXPLAIN. Describes status quo or position in task, with respect to goal. Freely offered, not elicited.

Example: "I've got a cairn"

ALIGN. Check the other persons understanding or accomplishment of a goal; checks alignment of both participants plans or position in task with respect to goal; checks attention, agreement or readiness.

Example: "Okay?" meaning "are you with me?"

5.1.7.4.2 Response Moves

CLARIFY. Clarifies or rephrases what has previously been said; usually repeats given or known information.

Example: "South, two inches"

REPLY-Y / REPLY-N. Affirmative (REPLY-Y) or negative (REPLY-N) response to a Query-YN, CHECK or ALIGN move. Elicited by other person.

Example: "Yes I have", or "No I can't do that"

REPLY-W. An elicited response that is not categorisable as either positive or negative.

Example: "Down"

ACKNOWLEDGE. Vocal acknowledgement of having heard and understood. Announces readiness to hear next move.

Example: "All right" or "Oh right, I see what you mean"

READY. Indicates intention to begin new game and focuses attention on oneself, in preparation for the new move.

Example: "Now" or "Right"

It is worth noting that in the original publications which describe CGA, the authors restrict their categorisation to verbal signals. Clearly, communication does not only take this form, backchannels for instance, can be visual, such as head nods, as well as verbal (Krauss, Garlock, Bricker, & McMahon, 1977). In this dialogue analysis however, to include non-verbal information would require studying video analysis of individual's communication as well as merely audio recordings of the speech. Although it has not yet been done, non-verbal communicative behaviour could potentially be included in an analysis using CGA, were such data to be available.

Even when only verbal communication information is used, this does include more than simply words. All of the verbal communication information is used to aid the correct categorisation of utterances. Such information includes anything audible, including utterances such as mhm, uhhuh, which are considered word tokens, intonation, semantic content, noises with an expressive function (a sharp intake of breath, for example).

Similarly, it is worth noting that Conversational Games Analysis is not restricted to a literal interpretation of the surface structure of the utterances made by interlocutors, the analysis takes into account the manner in which the content may be shared. Although semantic content can be useful as a clue towards correct classification of an utterance, the coder is not constrained by what is said – they are concerned with function. The advantage of this is that it permits the coder to interpret the inexplicit or apparently ambiguous contributions which are typical of spoken dialogue. An example of this given by Kowtko was the treatment of the word "Right". This can function as a statement, question or a confirmation depending upon intonation, dialogue context, and the participants' current shared and unshared knowledge of the map world they are discussing. All these sources of information are used to categorise each utterance in CGA. To return to the point raised by Eighmey (1993) in Chapter 2, that the process model of communication is liable to underestimate the importance of ideas and context. CGA explicitly encourages the coder to take into account the context of an utterance to infer the meaning of the communication that occurs between interlocutors.

The coder is concerned with categorising all utterances with reference to the speakers' communicative goals. This is the intended meaning of the utterance, and not necessarily the received meaning of an utterance. An instruction does not cease to be an instruction if the intended receiver perceives that instruction as a question. Conversational Games Analysis codes meaning, not understanding.

Conversational Games Analysis is clearly not the only dialogue coding scheme available but was chosen for this study because of its emphasis on the function and content of the utterances. The aim of this study is to discover whether the functional content of dialogues differed between participants interacting with a remote interlocutor who appeared zoomed in, compared with those interacting with an interlocutor who appeared zoomed out.

5.1.7.4.3 Prior Applications of CGA

Since CGA was developed it has been used to code conversational content in a wide range of studies. It has been used to examine the structural differences between dialogues in task oriented situations. Doherty-Sneddon, Anderson, O'Malley, Langton, Garrod & Bruce (1997) completed two experiments, which compared two-person dialogues for the map task which were conducted either face-to-face, video-mediated or audio-only. The results of their first analysis compared participants who completed the task in either a face-to-face or an audio-only situation. The analysis showed that participants who could not see each other (audio-only) used on average 53% more conversational games to complete the task than participants who could see each other (face-to-face). When the frequency of each type of move was analysed, the results showed that in the audio-only condition, participants used 131% more ALIGN games than those in the face-to-face condition and 66% more CHECK games. These results show that when participants were deprived of the visual channel, they checked both their own understanding (CHECK games) and that of the listener (ALIGN games) more often over the course of the interaction than participants completing the task face-to-face.

In their second study, Doherty-Sneddon et al. compared the dialogues of participants who each completed the map task in one of three remote conditions. These conditions were remote audio-only, which had no visual channel, and two video-mediated conditions, which differed depending on whether eye contact was possible or not. Eye contact was provided by the videotunnels technology described in Chapter One. In this study, the results showed that participants in both of the video-mediated conditions used more speaking turns to complete the task. Conversational Games Analysis showed that those participants who used a video-mediated system (with or without eye contact) used more EXPLAIN games – showing that they explained themselves more. Participants who interacted remotely by using a video link with eye contact used more CHECK games than those in any other condition, showing they were checking their own understanding more often. Participants who interacted remotely without any visual channel made more ALIGN moves than those interacting with the aid of a visual channel. This indicates that speakers without the use of a video link were confirming the understanding of the listener more often than those interacting with a visual channel.

Taken together, the results of the Doherty-Sneddon et al. studies showed that participants needed to elicit verbal feedback less often when talking face-to-face than when interacting either remotely or without visual signals. They also check their understanding of messages less often, a finding the researchers account for by suggesting listeners are more confident in a face-to-face interaction. Participants who interact remotely by video link do not show the same pattern of dialogue as in a face-to-face situation, they volunteer information more frequently and also check their own understanding more. Again confidence is suggested to be a factor, with remote participants less confident that they have reached a state of common ground than co-present participants.

Veinott, Olson, Olson & Fu (1999) used a modified version of Conversational Games Analysis to study the content of dialogues, using a version of the map task. The aim of their experiment was to investigate cross cultural differences in the way individuals complete a task remotely using either a video link or an audio-only link. They compared pairs of speakers who were either native speakers of English or non-native speakers of English. Their results show that when communicating by audio alone, non-native speakers make fewer instructing moves and fewer acknowledgements than native speakers. This difference was not apparent for those interacting using a video link. Participants using a video link also used more ALIGN and CHECK moves than those using audio alone – as was found by Doherty-Sneddon et al. (1997). Finally, Veinott et al. found that non-native speakers asked more clarification across the course of an interaction than the native speakers. This was not affected by the media they were using.

CGA has also been used to code an extensive number of dialogues, most notably the HCRC map task corpus (Anderson et al, 1991), leading to high levels of reliability and validity. Carletta et al. (1997) report that the classification of moves has been found to be reliable between coders and across task domains. Carletta et al. analysed a transcript of a discussion between a sales assistant and two customers and demonstrated that CGA could be reliable in this alternative situation.

Being a new task, no dialogues from the Financial Advice task had previously been coded. There were however, a number of reasons to expect that conversational games analysis would be appropriate for the task. Most of the studies using CGA in the past were using the Map Task, which has some similarities to the Financial Advice task. Both are tasks which promote a dyadic interaction, involving discussing information in front of them, with an aim of one participant successfully completing a particular goal. The Map Task however has a more defined outcome. Participants are aware that there is a very well defined correct answer in the Map task. The Financial Advice task is less focused on achieving a right answer, participants are made aware that the aim of the task is to find the option that they feel suits them.

The tasks also differ in terms of the status of the two interlocutors. In the Map task, participants are peers, both collaborating to achieve a goal. They also have broadly equivalent amounts of information and the emphasis of the task is on the exchange of each individuals information to the other. Consequently both participants have the same invested in the outcome of the task. In the financial advice task, the interlocutors are not of equal participatory status – one is an expert and the other is seeking their advice. The financial advisor posses more information, and information of a different kind to that of the customer. The financial advisor is attempting to help the customer but has little invested in a successful outcome – at the end of the day the customer is always right. In the financial advice task, the customer is also at liberty to add information of their own to the task, which they are unable to do in the Map Task. This, along with the emphasis on opinion and conjecture, leads to the financial advice task having more of a social aspect to it than the Map task.

CGA has been used to explore a number of problem solving domains. The Financial Advice task represents a new form of problem solving dialogue. There is an empirical question of whether CGA can be used both reliably and in a way that illuminates the hypothesis for the present task.

5.1.7.4.4 Applicability of CGA to the Present Task

Because of the labour intensive nature of dialogue coding it was essential to test whether the scheme would be applicable to the present research before actually going ahead and coding all the dialogues. In order to do this, CGA was applied to two of the dialogues from the present study, selected at random. These were compared with two dialogues taken from the Map Task corpus (Anderson et al., 1991) which coded by the same coder at the same time.

These 4 dialogues were then compared to investigate whether CGA seemed appropriate for use with the financial advice task. The results of this comparison showed that every category was used in at least one of the dialogues from both the financial advice task and the Map Task. This suggests that the application of coding to the financial advice task can be done in a similar way to that of the Map task. Examples of each type of conversational move from the dialogues of the Financial Advice task are given in Table 5.2.

Conversational Move	**Definition**	**Example of use in Financial Advice Task**
INSTRUCT	A direct or indirect instruction to be done immediately or shortly	*5.1.7.5 "Take me back to the page about shares"*
CHECK	Checks self-understanding of a previous message or instruction by requesting confirmation, either directly or indirectly.	*5.1.7.6 "Did you say the shares page"*
QUERY-YN.	Yes-No question asking for new or unknown detail about task (not clarifying instructions)	*5.1.7.7 "Do I have to pay charges every year"*
QUERY-W.	As Query-YN but an open ended question.	*5.1.7.8 "What are the charges for"*
EXPLAIN	Describes status quo or position in task, with respect to goal. Freely offered, not elicited.	*5.1.7.9 "I am looking for a long term savings plan"*
ALIGN	Check the other persons understanding or accomplishment of a goal; checks alignment of both participants plans or position in task with respect to goal; checks attention, agreement or readiness.	*5.1.7.10 "Are you happy with that"*
CLARIFY	Clarifies or rephrases what has previously been said; usually repeats given or known information.	*5.1.7.11 "The charges are applicable every year"*
REPLY-Y / REPLY-N.	Affirmative (REPLY-Y) or negative (REPLY-N) response to a Query-YN, CHECK or ALIGN move. Elicited by other person.	*5.1.7.12 "Yes I am"*

Conversational Move	Definition	Example of use in Financial Advice Task
REPLY-W	An elicited response that is not categorisable as either positive or negative.	*5.1.7.13 "The charges are for managing the account"*
ACKNOWLEDGE	Vocal acknowledgement of having heard and understood. Announces readiness to hear next move.	*5.1.7.14 "Yes, go on"*
READY	Indicates intention to begin new game and focuses attention on oneself, in preparation for the new move.	*5.1.7.15 "Right"*

Table 5.2 Examples of types of conversational moves in the Financial Advice task.

Also of importance was the fact that no utterance was discovered which appeared to require the development of a new category. Similarly, although there were clearly differences in the relative frequencies of the categories, none appeared to be overwhelmingly common in comparison to others, which could have suggested a need to subdivide categories. Consequently, CGA appeared to be suitable for the analysis of the Financial Advice task dialogues. Thus it was decided to use CGA to analyse the dialogues of Study 2.

5.1.7.15.1 Inter-Coder Reliability

The other important issue was whether the coding scheme would provide sufficiently high levels of inter-rater reliability to justify its use. Reliability was assessed by having an experienced CGA coder, who had never experienced the financial advice task before, code the two test dialogues described above. To measure the agreement between the two coders, apart from the probability of chance agreement, the Kappa statistic was used (described by Carletta, 1996). The coefficient of agreement (K) is a number between 0 and 1, where 0 means no agreement apart from chance and 1 means perfect agreement. It is possible to test whether K differs significantly from chance but it is also possible to interpret the scale of agreement.

Taking both dialogues together, the analysis for the financial advice task revealed a kappa coefficient of 0.94 at the level of the conversational games, and 0.88 at the level of conversational moves. A coefficient of 0.8 or higher is generally considered to show high reliability (Flammia & Zue, 1995; Carletta, 1996) so these results were taken to show a high level of inter-coder reliability. These figures are comparable with those for coders using CGA to analyse the Map Task, e.g. Carletta et al. (1997) had an interjudge reliability at the move level of 0.92 for selecting when a move begins, and 0.82 on agreement of the actual classification of moves.

5.2 Summary of the Results of the Structural Dialogue Analysis in Study 2

The aim of the present research is twofold. Firstly to discover the typical distribution of communicative goals displayed by speakers in the financial advice task. Secondly to examine any systematic differences between the content of dialogues in each of the two conditions. Such differences may include how frequently individuals request information, provide feedback or clarify their understanding. The two conditions were close (participants saw an image of the financial advisor which was zoomed in so that only the head was visible) or far (participants saw an image of the financial advisor zoomed out, showing the head and upper torso, as well as a lot of background).

The results presented in Chapter 4 showed that there were structural differences between the dialogues of those participants in the close condition and the far condition. The main findings from the structural analysis are presented in Table 5.3. It is the aim of the present analyses to determine what kinds of communicative functions underpin these differences.

Dialogue Measure (Mean number per dialogue)	Condition		Net difference (far → close)
	Far	Close	
Total turns	100.6	138.6	+39
Customer turns	52.9	73.8	+21
Financial Advisor turns	47.7	66.3	+19
Backchannels	Not significant		
Turns which included overlaps	14.83%	17.15%	+2.32%
Customer overlaps	6.23	13.53	+7
Customer words	252	453	+201
Financial Advisor words	974	1177	(+203; ns)

Table 5.3 Summary of results of the structural analysis of dialogues in Study 2.

These measures illustrate the mean structural differences that exist between dialogues in each condition. By using Conversational Games Analysis to analyse the typical differences in content it is hoped to explain what function underlies the differences between typical dialogues.

As a rough guide, the results show that the present analysis seeks to explain a difference between dialogues averaging 201 customer words and 20 customer turns. It should be remembered that the definition of turn taking in a dyadic interaction means that it is nearly inevitable that when the number of turns of one interlocutor rise, so the other interlocutor will show a similar increase. This explains why the increase in the number of turns for the financial advisor did not necessarily lead to an increase in the number of words spoken.

5.3 Method

5.3.1 Transcription

The dialogues had all been fully transcribed for the structural analysis carried out in Study 2. There were 28 dialogues in total, 13 in the far condition and 15 in the close condition. Transcription involves recording every word spoken by each interlocutor in chronological order. Parts of the dialogue when both individuals are speaking at the same time are noted, as are pauses in the dialogue.

The result is an accurate record of everything that was said by all interlocutors but without any guide to the pragmatic function of the speech.

The dialogues had been recorded in stereo with the customer's contributions on one channel and the financial advisor's on the other channel to aid the transcription process.

5.3.2 Coding

The aim of coding is to take the transcript and to produce an accurate record of the function of all of the speech in a dialogue. The most successful way of doing this is to read the transcript whilst listening to the original dialogues. The transcript is then coded, dividing all speech into moves. All the moves are also grouped into conversational games.

The coder was trained by a coder experienced in CGA until they could achieve an inter coder reliability, measured by the kappa coefficient, of above 0.80 at the move level on three separate dialogues chosen at random from the HCRC Map task corpus (Anderson et al., 1991). To check the validity of the training, the two coders also compared a sample dialogue from the financial advice task and also achieved an inter-coder reliability of 0.88.

It should be noted that both transcription and coding are extremely labour intensive processes. Coding a 15 minute dialogue will take on average about 3 hours, after having already been fully transcribed.

5.4 Results

Coded dialogues were analysed using 2-way (2x2) ANOVAs for independent samples. The independent variables were condition (close or far) and role (customer or financial advisor). The dependent variables were the number of each particular type of move. For ease of analysis, the results shall be presented grouped by those moves that are initiating moves (those that also represent games) and response moves.

A 2-way ANOVA comparing the total number of moves made by each participant in a dialogue showed a significant main effect of condition. The mean number of moves in the close condition was 105.9 compared with 71.9 in the far condition [$F (1, 52) = 75.26$; $p < .01$]. This indicates that dialogues where the advisor appeared closer to the customer contained significantly more conversational moves than those when the advisor appeared further away. There was no significant interaction of role and condition, indicating that both the customer and the financial advisor were making more conversational moves.

5.4.1 Initiating Moves

The first analysis compared the total number of conversational games initiated by participants over whole dialogues. This type of analysis can give an indication of the relative number of goals in a interaction and the overall interactivity of the participants. The mean total number of games was compared by both condition and role and is presented in Table 5.4.

		Condition		
		Close	Far	Total
Role	Financial Advisor	71.33	62.15	67.07
	Customer	22.00	12.77	17.71
	Total	93.33	74.92	84.78

Table 5.4 Mean total number of games initiated by role and condition.

Results of a two-way ANOVA (2x2) on the total number of games initiated showed no significant main effects of proximity or any significant interaction between proximity and role. There was a significant main effect of role [$F(1, 52) = 122.7$, $p < 0.01$] which showed that the financial advisor initiated more games (mean = 62.91) than the customer (mean = 13.32).

It should be concluded that changes in proximity did not affect the total number of games initiated by either the financial advisor or the customer during the interaction. The financial advisor initiated more games than the customer in both conditions.

5.4.1.1 Total Number of each particular type of Initiating Move

The frequency of participants initiating each individual type of game was also analysed separately. Table 5.5 presents the mean number of games initiated by interlocutors in each condition.

	Close		Far	
	Financial Advisor	Customer	Financial Advisor	Customer
INSTRUCT	4.53	0.93	5.00	0.31
CHECK	4.40	4.93	4.69	2.54
QUERY-YN	12.53	2.80	11.77	1.23
QUERY-W	5.13	2.60	5.15	1.62
EXPLAIN	34.20	10.33	26.62	7.00
ALIGN	10.53	0.40	8.92	0.08
TOTAL	71.32	21.99	62.15	12.78

Table 5.5 Summary of mean number of games initiated by role and condition.

A 2-way (2x2) ANOVA was computed for each type of initiating move comparing both role and condition. Table 5.6 summarises the results of all of the ANOVAs

Move	Effect	F (1, 52)	p
INSTRUCT	Role	75.26	p<.01
	Condition	0.03	ns
	Role x Condition	1.31	ns
CHECK	Role	0.48	ns
	Condition	0.81	ns
	Role x Condition	1.32	ns
QUERY-YN	Role	174.37	p<.01
	Condition	2.31	ns
	Role x Condition	0.28	ns
QUERY-W	Role	34.75	p<.01
	Condition	0.88	ns
	Role x Condition	0.95	ns
EXPLAIN	Role	57.32	p<.01
	Condition	3.61	ns (p<0.1)
	Role x Condition	0.55	ns
ALIGN	Role	122.79	p<.01
	Condition	1.27	ns
	Role x Condition	0.56	ns

Table 5.6 Summary of ANOVA results for individual initiating moves.

The results of the ANOVA's reported in Table 5.6 show that the only significant effects involving individual initiating moves, were all main effects of role. There was a significant main effect of role for ALIGN [$F(1, 52) = 122.79$; $p < .01$].

EXPLAIN [$F(1, 52) = 57.32$; $p < .01$], INSTRUCT [$F(1, 52) = 75.26$; $p < .01$], QUERY-W [$F(1, 52) = 34.75$; $p < .01$] and QUERY-YN [$F(1, 52) = 174.37$; $p < .01$] games. Indeed the only initiating move for which there was no significant effect of role was CHECK, indicating that CHECK moves were evenly distributed between the two different roles. There was no significant main effect of condition for any type of conversational game. Nor were there any significant interactions between condition and role.

Table 5.7 shows the means for each initiating move by role.

	Financial Advisor	Customer	Sig.
INSTRUCT	4.78	0.6	**
CHECK	4.52	3.71	
QUERY-YN	12.14	2.0	**
QUERY-W	5.14	2.1	**
EXPLAIN	30.41	8.62	**
ALIGN	9.73	<0.5	**

Table 5.7 Means for individual initiating moves by condition.
** $p < 0.01$

Examination of the means for each type of initiating move shows that the financial advisor makes more of each type of initiating move than the customer. The only initiating move that this does not apply to is the CHECK move, which shows no significant effect of role.

5.4.2 Response Moves

As for the initiating moves, 2-way (2x2) ANOVAs were computed to compare the frequency of occurrence of each individual response move between role and condition. Table 5.8 presents the mean number of each type of response move comparing between role and condition.

	Close		Far	
	Financial Advisor	Customer	Financial Advisor	Customer
CLARIFY	0.60	0.40	0.31	0.08
REPLY-Y	8.47	17.87	5.00	16.77
REPLY-N	0.80	1.87	0.23	2.38
REPLY-W	3.13	5.40	1.38	4.54
ACKNOWLEDGE	29.33	35.40	20.62	19.54
READY	1.80	0.87	1.00	1.08
TOTAL	44.13	61.81	28.54	44.39

Table 5.8 Mean number of individual response moves by condition and role

Table 5.9 summarises the results of all of the ANOVAs analysing the frequency of the response moves.

Move	Effect	F (1, 52)	p
CLARIFY	Role	0.69	ns
	Condition	1.40	ns
	Role x Condition	0.01	ns
REPLY-Y	Role	41.28	p<.01
	Condition	1.92	ns
	Role x Condition	0.52	ns
REPLY-N	Role	16.84	p<.01
	Condition	0.01	ns
	Role x Condition	1.92	ns
REPLY-W	Role	20.98	p<.01
	Condition	4.86	p<.05
	Role x Condition	0.56	ns
ACKNOWLEDGE	Role	0.40	ns
	Condition	9.61	p<.05
	Role x Condition	0.81	ns
READY	Role	1.06	ns
	Condition	0.50	ns
	Role x Condition	1.48	ns

Table 5.9 Summary of ANOVA results for individual response moves.

The results of the two-way ANOVA's on the mean number of each type of response move revealed no significant interactions between condition and role for any type of response move.

The results did show significant main effects of role for REPLY-N [$F(1, 52) = 16.84$: $p < .01$], REPLY-W [$F(1, 52) = 20.96$; $p < .01$] and REPLY-Y [$F(1, 52) = 41.28$; $p < .01$]. The means for each of these moves are presented in Table 5.10.

Move	Financial Advisor	Customer	Sig.
CLARIFY	0.46	0.25	ns
REPLY-N	0.5	2.1	**
REPLY-Y	6.9	17.3	**
REPLY-W	2.4	4.9	**
ACKNOWLEDGE	25.29	28.04	ns
READY	1.43	0.96	ns

Table 5.10 Means for individual response moves by role.
** $p < 0.01$

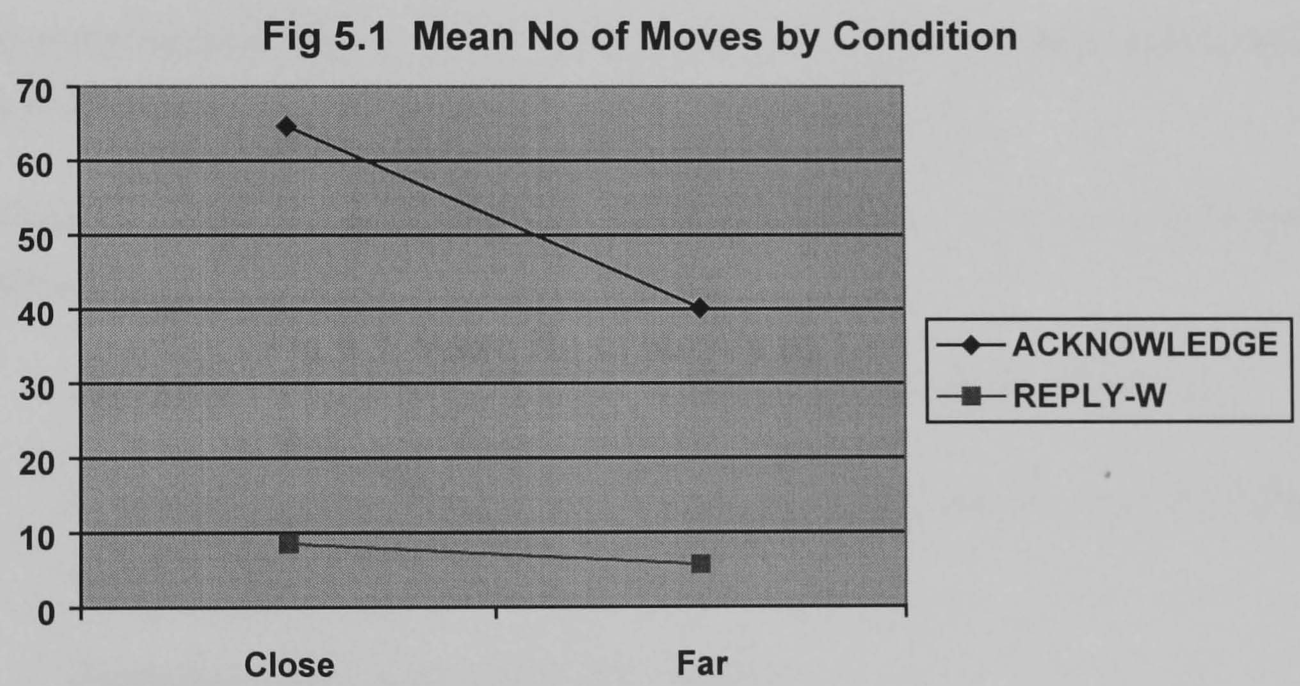
By studying the means it can be seen that the customer made more of each of these types of response move than the financial advisor.

The 2-way ANOVA also showed a significant main effect of condition for both REPLY-W [$F(1, 52) = 4.86$; $p < .01$] and ACKNOWLEDGE [$F(1, 52) = 9.61$; $p < .01$] moves. The mean number of each type of response move in each condition are presented in Table 5.11.

Move	Close	Far	Sig.
CLARIFY	1.00	0.38	ns
REPLY-N	2.67	2.62	ns
REPLY-Y	26.33	21.00	ns
REPLY-W	8.53	5.69	*
ACKNOWLEDGE	64.73	40.15	*
READY	0.38	2.08	ns

Table 5.11 Means for individual response moves by condition.

* $p < 0.05$



It can be seen from Table 5.11 and Fig. 5.1 that significantly more REPLY-W and ACKNOWLEDGE moves were made in the close condition than in the far condition. The results suggest that on average, when the advisor appears close the dialogue contains a total of 24 more ACKNOWLEDGE moves and 3-4 more REPLY-W moves. Because there was no significant interaction of role by condition for these moves, it can be concluded that both the customer and the advisor are making the extra moves. Both interlocutors are therefore acknowledging more and giving more responses which would not be categorised as either affirmative or negative.

ACKNOWLEDGE moves are defined as "vocal acknowledgement of having heard and understood. Announces readiness to hear next move". This tended to be done by using a short piece of dialogue, which may or may not constitute a full turn in itself. Examples of typical uses of acknowledgements in dialogue are given in examples 1 and 2.

Example 1 – using an ACKNOWLEDGE move to achieve closure.

> **Customer**: Are there charges on the bonus account (QUERY-YN)
>
> **Financial Advisor**: No (REPLY-N) the bonus account is free of charges (EXPLAIN)
>
> **Customer**: Ok, I see. (ACKNOWLEDGE)

This is the most common use of an acknowledge move in the Financial Advice task. It is a whole turn and represents closure of a conversational game.

Example 2 – using an ACKNOWLEDGE move as feedback, without seeking to gain control.

> **Customer**: What are the advantages of the bonus account? (QUERY-W)
>
> **Financial Advisor:** The bonus account is free of charges and has instant access to funds (REPLY-W)
>
> **Customer**: Right, ok. (ACKNOWLEDGE)
>
> **Financial Advisor:** and it has a yearly bonus payable if you make less than two withdrawals in a year (REPLY-W)

In this second example, the ACKNOWLEDGE move is used to indicate understanding but without closing the game. The point of this move is to indicate understanding but without closing the game. The Conversational Game is still open for the financial advisor to continue their explanation.

When reviewing the dialogues, there appeared to be a tendency for the dialogues in the close conditions to contain more of the latter type of acknowledge move. This led to the dialogues feeling more interactive. There were fewer long periods of talk when the financial advisor was effectively lecturing the customer.

5.5 Discussion

The results of the present analysis have demonstrated that the content analysis can illuminate how different forms of videoconferencing displays impact on users. The results show that the increase in the number of turns spoken in the close condition is mirrored by an overall increase in the mean number of conversational moves spoken by participants. This increase is not across the board however, with only the number of ACKNOWEDGE and REPLY_W moves being significantly higher in the close condition. It is notable also, that the increase in the number of acknowledge moves in the close condition is greater for the customer than the financial advisor.

It is useful to consider the results in terms of the overall increase in dialogue that was observed from the structural analysis of study two. By studying the mean number of moves in each condition, differences can be measured on average. Overall in the close condition, dialogues consisted of an average of 105 moves compared to 72 in the far condition. This corresponds to an average of 33 moves of "extra" dialogue in the close condition. Of these, 16 can be traced to ACKNOWLEDGE moves by the customer, 9 ACKNOWLEDGE moves by the financial advisor, but only 1 REPLY-W move to the customer and 2 REPLY-W moves by the financial advisor. It seems reasonable to suggest therefore, that the major difference in the dialogues is the number of ACKNOWLEDGE moves being made.

These results show that both the customer and the financial advisor made more ACKNOWLEDGE and REPLY-W moves in the close condition than the far condition. The size of the effect we were seeking to explain from the previous Study was roughly 200 customer words and 20 customer turns. ACKNOWLEDGE moves tend to be short, while REPLY-W moves can differ widely in verbosity. Considering that the ACKNOWLEDGE and REPLY-W moves were spread across both roles, it would unrealistic to claim that they accounted for all of the extra speech caused in the close condition, but it is fair to say that they account for most of it.

The aim of this analysis was to attempt to explain the structural differences in the dialogues that were investigated in the previous chapter. The key findings from the structural analysis were that there were more turns in the close condition, the customer spoke more, and interrupted more in the close condition. The functional analysis has shown that most of the extra dialogue which is present when the advisor appears close is in the form of acknowledgements. When it is considered that the customer says much less than the financial advisor in the dialogue, it is noticeable that the increase in the number of acknowledgements is higher for the customer than the financial advisor.

ACKNOWLEDGE moves tend to be short utterances of verbal feedback. Their function is to communicate understanding, but they often do not seek to take the floor for their speaker. When listening to the dialogues from both conditions, it was possible to get a feeling for a definite pattern emerging regarding the style of the dialogues in both conditions. Typical dialogues in the close condition were often characterised by the customer in particular giving a great deal of verbal feedback while the financial advisor was saying something. There would be a lot of "mhm", "yup", "uhhuh", I'm with you", "carry on" etc. when the financial advisor was speaking. By contrast, typical far condition dialogues were characterised by long passages where only the financial advisor spoke. In these dialogues, the customer seemed content to let the financial advisor speak, wait for them to finish and then say something else, whereas in the close condition, the customer kept acknowledging that they were understanding what the financial advisor was saying. The difference in styles could be described as conversational style or lecture style and was akin to talking with or being talked to. This difference in style may also account for the increase in overlapping speech in the close condition.

During the experiment, only the customers' view of the advisor was different between conditions, not the advisors' view of the customer. For this reason, it would be fair to expect differences to occur only for the customer and not for the financial advisor. This was not the case with the acknowledgements, despite the fact that some of the dialogue measures - most notably the number of overlaps – did indeed illustrate this one-way effect. A possible reason for this is that whereas incidences of overlapping speech are undoubtedly examples of individualistic behaviour, dialogue is reciprocal and reactive. An individual may change the way they speak depending upon the nature of the conversation, whereas overlapping the other persons speech is perhaps more a function of mistiming when someone will pause or wanting to jump in. It is possible that a high level of acknowledging by one person encourages and leads to more acknowledging than usual by the other, whereas overlapping may be less likely to be "contagious".

It is interesting to note that the number of conversational games did not differ significantly between conditions. This suggests that the number of games required to complete the interaction did not differ according to any changes in style afforded by the participants. Despite there being more turns and more moves, because there were no more games, there are several conclusions which can be drawn. These include that in the close condition the customer is probably not asking more questions or making more social comments, more dialogue irrelevant to the task. Both of these were possible explanations when looking simply at the structural data.

The other domain in which this study has been useful is that the results gave a clearer indication of what was going on for the different roles of the task. The financial advisor initiated significantly more games than the customer, indicating their dominance in driving the conversation along.

5.5.1 Implementation of CGA

This was the first time that any type of coding of dialogue content had ever been done for dialogues in the financial advice task. The use of Conversational Games Analysis would seem to have been generally successful in the present study. The coding scheme is well suited to analysing dyadic interactions and there were few problems in applying the scheme to a task which had never been coded before.

There were a couple of issues which came up during the course of the analysis which are worthy of note. Firstly it is possible for an utterance to have more than one meaning. Consider the following interaction:

FA: “Would you like a new car?“

Customer: “Yes” (phonetically – “yeeeeeees”)

FA: “I was wondering if you might be needing access to the money”

In the above example, the answer that the customer gives is not only an affirmative reply to the original question, but also a query as to why the question is being asked. In the rules for coding using CGA (Carletta et al. 1997), no mention was specifically made of such a case. It is worth stating though, that utterances can have two meanings, both intended by the speaker, and in such a case, both should be coded and analysed. There is no reason for either meaning at the game level or the move level to take precedence over the other.

5.6 Conclusions

This study has used analysis of the content of dialogues to illuminate structural differences in dialogues. The application of Conversational Games Analysis for this purpose was successful. The results of the coding have provided a detailed analysis of the speech of both the participants and the confederate in each dialogue. The results have shown Conversational Games Analysis to be useful in exploring differences in style of dialogues for the present task, without modification.

This study has been illuminating in terms of explaining the data from experiment 2 which examined the structural differences in the dialogues between conditions. Participants playing the role of the customer in the financial advice task have been shown to make a greater number of ACKNOWLEDGE and REPLY_W moves when the image of the financial advisor appears zoomed in close than when they appear far away. This is concluded to translate to a tendency to give a lot more verbal feedback in the close condition than in the far condition. This in turn may lead to the advisor changing their dialogue style to also offer more verbal feedback. These findings are in accord with the suggestion in the previous chapters that dialogues in the close condition could be considered as being more interactive than those in the far condition.

6 Chapter 6 . Study 3: The Effect of Image Size and Background on the Perceived Distance of Individuals in Photographs

Studies one and two have shown that impressions of proximity of remote individuals presented via a video link can be manipulated by changing the level of camera zoom used. In an interactive situation this alteration in perceived distance has been shown to lead to differences in participants' conversational behaviour. As the behavioural effects show a similarity to those caused by changes in proximity in face-to-face proxemic studies it is hypothesised that the crucial factor in the changing of camera zoom is to change the impression of distance of the remote individual. Changing the level of camera zoom used to portray an individual actually changes more than one factor within the image that may regulate impressions of proximity.

Visual cues to distance are plentiful, and consequently there may be many factors that affect the impressions of visual proximity of mediated individuals. In order to understand more about the subtleties of creating impressions of perceived distance, an experiment was constructed to investigate the role of visual distance cues on the perception of distance. The key question that this research sought to answer was "Which factors in an image regulate impressions of proximity in an image?"

Study One showed that changing the zoom on the video camera affected the feeling of perceived distance of the other individual. As was noted in Chapter 3, altering the camera zoom changes both the image size of the object in focus (in this case the individual) and the amount of background visible. Study 1 showed that image size was the key factor in perceived proximity using video-mediated technology, as opposed to the relation of the size of the image to the amount of background visible. What this research did not show however was whether this finding is specific to video-mediated technology or if it would also generalise to photographs of individuals.

6.1 Estimating Distance

In face-to-face situations the problem of estimating distance is straightforward. Humans are very good at estimating the distance another person or object is from them, as they do it every day. An experiment by Witmer & Sadowski (1998) tested participants in real and virtual environments. The task was to view a target for 10 seconds and then walk to it without vision. This task was completed in both a real-world situation and a virtual environment. The results showed that people tend to slightly underestimate the distance of the target, real-world judgements average 92% of the distance, while in virtual environments the judgements average 85%.

Virtual environments represent an intermediate stage between pictures and face-to-face situations. In a virtual environment both pictorial cues and binocular disparity cues are available, losing only minor face-to-face distance cues such as olfactory information. Witmer & Sadowski (1998) note that the performance deficits seen in the Virtual Environment could be due either to poor binocular disparity cues or from a distortion of pictorial depth cues. It is also probable that the novelty of the technology is also likely to play a part - even amongst experienced users.

Modern research into distance perception has indicated that there may be two types of perceptual system, the two visual systems framework (Bridgeman et al., 1979; Milner & Goodale 1995). This proposes that there is a motor visual system which guides movement, and a cognitive visual system which is involved in conscious perceptual judgements (Mon-Williams & Tresilian, 1999). In video-mediated encounters and in still photographs, judging how far away someone appears will be governed by the cognitive system. Mon-Williams & Tresilian (1999) showed that in reduced cue environments, participants were much better at motor tasks (pointing to a target) than verbal tasks.

In still photographs, only pictorial cues to depth are available. To judge distance in a photograph is to assign three-dimensional properties to a two dimensional image. Unavailable are the physiological cues to depth of convergence and accommodation, binocular disparity is identical for every photograph being viewed, regardless of subject, provided the actual photograph is held at the same distance. Studies of depth perception have highlighted several pictorial cues to depth, all of which may be present in either a video-mediated image or a still photograph of an individual (for an overview of depth perception see any introductory perception text e.g. Gregory, 1998, pp.60-66; Bruce & Green, 1992, p141-159). As only pictorial cues to depth are available, the motor visual system is redundant when it comes to viewing photographs with a view to eliciting impressions of distance. Similarly, impressions of distance within video-mediated communication must also primarily use the cognitive system. This may primarily be where differences lie between the impression of proximity in a face-to-face environment and a mediated environment.

Research has shown that altering camera zoom does alter people's perceptions of how far away a video-mediated individual appears (Reeves & Nass, 1996). This is not surprising, and is the way that the majority of people would seek to make someone appear closer or farther away without actually moving them – hence the terms "zooming in" and "zooming out". How zooming creates this feeling of distance is less clear, however. What "zooming in" actually does is to effectively decrease the total physical area seen by a particular size of field of vision. The net result is to see fewer objects but for them to appear larger, taking up a greater percentage of the total field of vision - just like moving closer to an object being focused on. There are two possible reasons for why objects appear closer in a "zoomed in image". The first reason is that they appear larger – quite simply the image size of the object is bigger. Alternatively, the reason may be based on the amount of the field of view the object image takes up – the object image to background ratio.

Study 1 implicated the former of these explanations to be the most critical. Participants had to zoom a camera until the video-mediated individual on the screen produced the same feeling of distance as a co-present individual in the room. Results showed that when a larger video window was used, less zoom on the camera was required to produce a particular feeling of proximity. The object image size chosen to represent a particular distance was constant across video window size, but the object image to background ratio differed.

It is important to understand whether these results generalise to still photographs – indicating applicability to human perception in general, or merely to the medium of video-mediated technology. The present experiment sought to answer the question of what the main factors governing feelings of perceived distance are and whether still photographs are treated in the same way as live video images.

The task devised was for participants to guess how far from the camera individuals were sitting in a variety of photographs. They were told a normal camera was used, it never moved and subjects in photos were always sitting at one of 4 distances. The factors being manipulated were image size (small, medium or large) and background (objects or blank). Background was included because it was felt that the presence or absence of depth cues may increase or decrease the difficulty of the task.

It should be noted that although interesting, it is not the primary aim of the present study to investigate how good people are at making judgements of absolute distance per se. Rather, the experiment investigates how the manipulation of factors in a photograph affects individual's judgement of distance, relative to each other. This is because it is virtually impossible to accurately judge distance from the camera of an object without knowledge of the zoom function of the camera, size of the object in the image, etc. Performance measures should indicate which variables affect the perception of distance relative to each other.

During the experiment, colour photographs were used to preserve as closely as possible the comparison with the video experiments that used live colour video. Similar reasoning explains why the images were of seated adults. Participants had just completed an experiment involving video conferencing with a seated adult. This has the added advantage of nullifying the height variation between the people in the images; thus ensuring excessive differences in size between characters was not an issue for participants when estimating distance.

6.2 Method

6.2.1 Participants

61 participants (25 Female, 36 Male; Age range 16-60 years, Median 22 years), recruited from around the University of Glasgow took part. All had just completed a separate experiment using video conferencing. Participants were paid £5 for taking part in both experiments and signed their consent after the procedure of the experiment was explained to them.

6.2.2 Materials

32 photographs were taken using an Apple QuickTake 100 digital camera. Photographs were taken on two separate days, within a 2-hour period in the afternoon. On the first day, all photographs with the blank background were taken and on the second day the photographs using the object filled background were taken. This ensured the lighting remained constant across photographs.

The camera was mounted on a tripod 12 feet 4 inches from the rear wall of the room. The wall was covered in white, matt finish boarding to provide a blank background. Particular effort was made to avoid excessive glare from the flash of the camera and attempts to reduce the amount of shadow were also made by shining anglepoise lights directly on to the rear wall. Some shadow was inevitable, however this was not considered to be a major problem. When being photographed, individuals sat on a chair, which was placed at each of 4 distances from the camera (2.5ft, 4ft, 6ft and 9ft). 2 photographs were taken at each distance and the best one of each selected for use later, ensuring that no photos were spoiled by the subject having their eyes closed etc. Subjects were asked to sit up straight with a neutral expression.

The photographs were downloaded onto an Applemac PowerPC and converted to high quality JPEG format. Adobe Photoshop version 4.0 on a PC running Windows 95 was used to convert the images to the appropriate size. Pictures were printed on to EPSON photo quality ink jet paper (formerly EPSON special coated paper for 720 dpi printing) using an Epson Stylus Color 400 ink jet printer, mounted on purple A4 card and covered in self-adhesive plastic to ensure durability. All pictures were in colour.

There were 96 images in total, encompassing 4 different variables:

CHARACTER: 4 different individuals were photographed - all were adults of similar height (range 5ft 4 → 5ft 10).

BACKGROUND: All photographs were taken against either a blank background or a background containing a number of objects, abundant in depth cues. The objects used to provide the depth cues were: a poster, filing cabinet, door, desk, anglepoise light, table, folders, sticky tape dispenser, flowers. These were placed at different distances to the camera, as opposed to all along the back wall.

SIZE: The images were converted to 3 different sizes, small (285 x 214 pixels; 100 x 80mm); medium (566 x 425 pixels; 200 x 160mm); and large (285mm x 215mm). Small and medium sizes were achieved using Adobe Photoshop version 4.0. To get the large images, medium ones were photocopied and enlarged 141% (A4 → A3), this was done without any noticeable decrease in quality. All pictures were in colour.

DISTANCE: Participants were photographed at each of four distances from the camera, 2.5ft, 4ft, 6ft and 9ft.

6.2.3 Procedure

Participants were told they were going to take part in an experiment investigating how well people can judge distance from photographs. It was explained that they would see a series of photographs of individuals, all of whom were sitting at one of four distances from the camera. Their task was to say at which distance each individual was sitting. Answers would be given by ticking a box on a simple answer sheet.

The experimenter then sat in a chair at each of the 4 distances from the participant, in order to demonstrate how far each distance actually was. The participant then began looking at the pictures and choosing the distance that they thought each picture corresponded to. Participants were instructed to go at their own pace but not to go back to any pictures after they had moved on. This was to prevent participants comparing answers with previous ones. After 48 there was a break and the experimenter repeated the sitting to demonstrate the distances, before the participant judged the other 48 photographs. The photographs were shown in the same order each time, only swapping the order of the two sets of 48. Within that order, the photographs were randomised.

6.3 Results and Discussion

Data was analysed in two ways. The primary measurement for each condition was the number of correct answers, which showed which conditions were guessed correctly most often. Another measurement involved studying the errors made, for each condition the direction of the errors would show whether the subject was generally appearing either closer or further away than was in fact the case. Overall, the mean score correct was 59%, above chance and ensuring a large amount of both correct answer data and error data.

6.3.1 Error Analysis

If either of the variables studied do systematically alter the perception of distance, it is expected for this to be evident from the direction of the errors made. If a variable leads to a feeling of closeness, then more errors will be made where the participant guessed too close than guessed too far away. The aim of this analysis was to compare the frequencies of each type of error (too close vs. too far) for each condition.

The only pictures included in this analysis were of those at 4ft or 6ft, as these were the only ones that allowed for an error in either direction to be made. In order to analyse the results, the frequencies of errors were compared by using a 4-way 2x3x2x2 ANOVA for repeated measures [background (2) x size (3) x distance (2) x direction (2)]. The only results of interest here are those that involve direction, as any effects across direction are discussed in the performance data (discussed in the next section). The data used for the direction was the percentage of answers that were errors of too close compared with the percentage of answers that were errors of too far.

6.3.1.1 Main effect of error direction

Of the total number of responses examined across the 4ft and 6ft conditions (2928), 57.1% were correctly identified. Of those in which errors were made, 26.2% of the total responses were errors of too close and 16.6% of responses were errors of too far. The results of the 4-way ANOVA indicated a significant main effect of direction [$F(1, 60) = 8.35$, $p < .01$]. This indicates that overall, errors were more likely to be made by the participant guessing that the subject of the photograph was sitting closer to the camera than was actually the case.

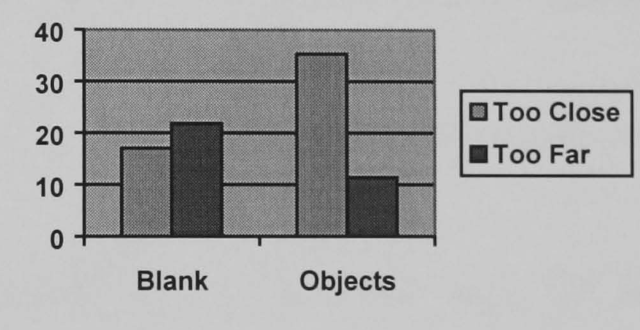
The results of the 4-way ANOVA comparing the direction of the errors with the distance, size and background showed significant interactions with direction of background [$F(1,60) = 105.23$, $p<0.01$], size [$F(2, 120) = 85.49$, $p<0.01$], and distance [$F(1, 60) = 137.74$, $p<0.01$]. There were also significant 2-way interactions of background by size [$F(2, 120) = 8.035$, $p < 0.01$], background by distance [$F(1, 60) = 4.79$, $p < 0.05$] and size by distance [$F(2, 120) = 3.37$, $p < 0.05$]. The three way interaction of background by size by distance was not significant [$F(2, 120) = 1.38$, $p > 0.05$].

6.3.1.2 2-Way Interactions involving error direction

6.3.1.2.1 Background

The ANOVA indicated a significant main effect of background [$F(1,60) = 105.23$, $p < 0.01$]. The mean percentage of errors of each type in each condition is presented in Table 6.1 and Fig 6.1.

Fig 6.1 Errors by background



	Too Close	Too Far
Blank	17.01	21.86
Objects	35.31	11.34

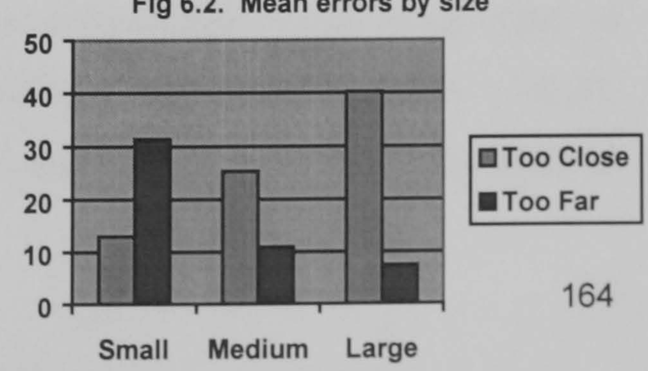
Table 6.1 Mean percentage of answers that were errors for each background condition.

Post hoc tests of simple main effects showed that with the blank background, errors were no more likely to be of too close than too far [$F(1, 60) = 1.87$, $p>0.05$]. When the object background was used, however, errors were more likely to be made by guessing too close [$F(1, 60) = 43.41$, $p < 0.01$].

6.3.1.2.2 Size

There was a significant interaction of size of image and direction [$F(2,120) = 85.49$, $p<0.01$]. Fig 6.2 illustrates the mean percentage of errors by type in each different size condition and they are presented in Table 6.2.

Fig 6.2. Mean errors by size



	Too Close	Too Far
Small	13.01	31.35
Medium	25.31	10.96
Large	40.16	7.48

Table 6.2 Mean percentage of errors by size.

Post hoc tests of simple main effects showed that at each size, there is a significant difference between the frequencies of the two error types ($p < 0.01$). Both the medium [$F(1, 60) = 15.30$, $p < 0.01$] and large photographs [$F(1, 60) = 21.38$, $p < 0.01$] leading to more "too close" errors while the small photographs led to more "too far" errors [$F(1, 60) = 63.40$, $p < 0.01$].

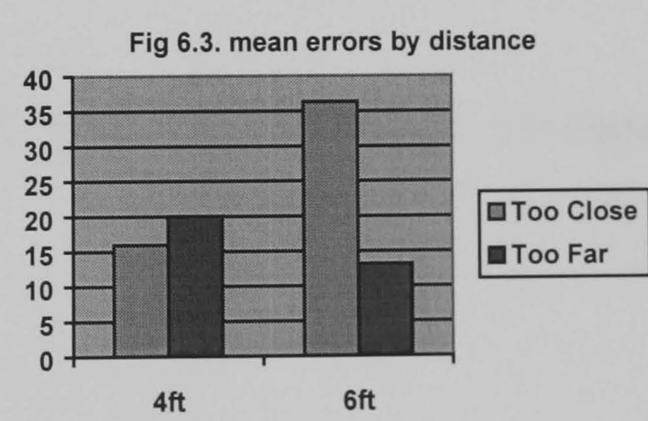
Fig. 2 also shows the number of close errors increases with size between the three conditions, while the number of far errors is greater in the small condition than either the medium or the large conditions.

6.3.1.2.3 Distance

The ANOVA indicated a significant interaction of direction and distance [$F(1,60) = 137.74$, $p<0.01$]. The percentage of errors in each direction is presented in Table 6.3 and Figure 6.3.

	Too Close	Too Far
4ft	15.98	20.01
6ft	36.34	13.18

Table 6.3 Mean percentage of errors by distance.



Post hoc tests of simple main effects indicated a significant simple main effect of direction when the photograph is of an individual sitting at 6ft [$F(1, 60) = 40.33$, $p<.01$]. It can be seen from Fig 6.3 that at this distance the errors are more likely to be made by individuals guessing too close. For the 4ft condition there was no significant simple main effect of direction [$F(1, 60) = 1.44$, $p < .05$]. This shows that there is no difference in the direction of errors made when viewing 4ft photos.

To summarise the results of these effects, the results have shown that there are more errors of too close than of too far overall. When broken down by variable, it can be seen that the object filled background, large pictures, medium pictures and the 6ft distance all lead to more errors of too close than too far. Small pictures lead to more errors of too far than too close.

6.3.1.2.4 3-Way interactions involving error direction

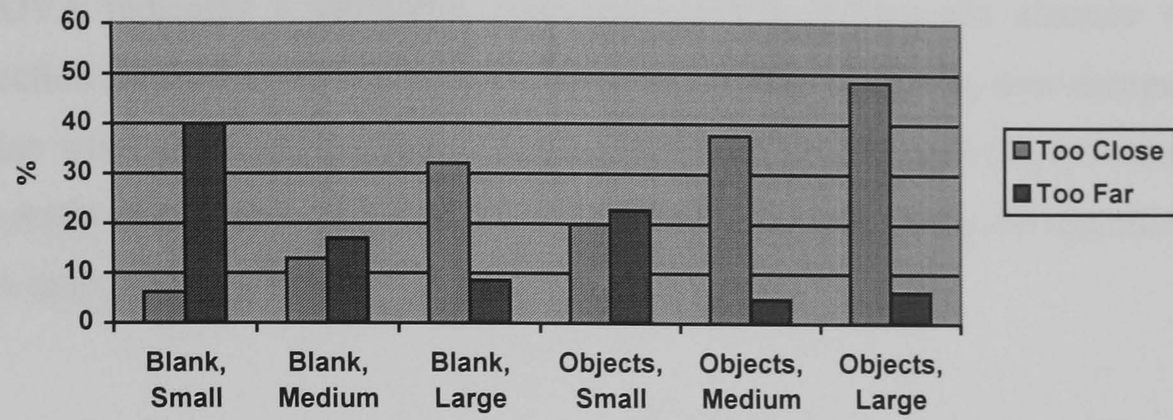
*6.3.1.2.4.1 Background * Size*

The 4-way ANOVA indicated a significant interaction of background and size with error direction [$F(2,120) = 8.035$, $p < 0.01$]. Post hoc Tukey HSD tests were computed to examine whether or not there was any significant difference between the number of errors made of each type in each condition. Table 6.4 summarises the direction of the errors made. Mean percentages of errors made are graphed in Figure 6.4.

	Small	Medium	Large
Blank	Far	ns	Close
Objects	ns	Close	Close

Table 6.4 Direction of errors made in size by background conditions. All effects reported are significant to $p<0.01$.

Fig 6.4. Errors by background and size



The results of the t-tests indicate that when the blank background is used, errors in the small condition tend to be "too far" errors, while errors in the large condition tend to be "too close". If objects are present in the background then there is no difference between the frequency of each type of error when the photograph is small but when medium or large there are significantly more errors of "too close" than "too far".

The difference in effect of the blank and objects background appears well illustrated by the difference between the blank and object conditions at the medium size. When the blank background is used there is no difference, but when the objects background is used, the majority of errors are of guessing too close. These results strongly suggest, as the main effects did, that a blank background has little or no effect on the perception of distance but a background filled with objects tends to make people appear closer in a photograph.

Similarly, the results suggest that small photographs tend to make people appear far away, while large photographs make people appear close. The reason there is no difference between the error types in the small condition with the object filled background may be that the two variables cancel each other out, each pulling in the opposite direction.

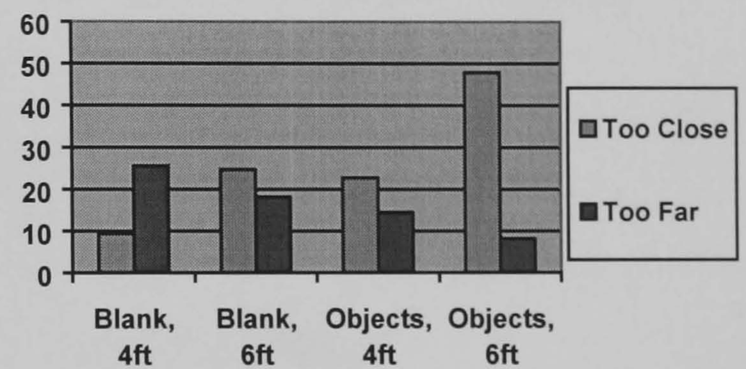
*6.3.1.2.4.2 Background * Distance*

The ANOVA indicated a significant interaction of background and distance with error direction [$F(1, 60) = 4.79$, $p < 0.05$]. Post hoc Tukey HSD tests were computed to examine whether or not there was any significant difference between the number of errors made of each type in each condition. Table 6.5 summarises the direction of the errors made. Fig 6.5 graphs the percentage of errors.

	4ft	6ft
Blank	Far***	ns
Objects	Close*	Close***

Table 6.5 Direction of errors made in size by distance conditions.

Fig 6.5. Errors by background and distance



* $p<0.05$ *** $p<0.01$

When the blank background was used for images of an individual at 6ft there was no difference between the frequencies of the different error types. This contrasts strongly with the significant difference found when using the object background at 6ft, when the frequency of "too close" errors was much higher than that of "too far" errors, indicating again the strong difference between the blank and object backgrounds.

When the picture was of an individual at 4ft the direction of the errors depended on the background – when used with the blank background, errors were likely to be of "too far" but when using the object background, errors were more likely to be of too close.

This pattern of results suggests that at 4ft, without any other variable pushing in any particular direction, errors tend to be of too far. Between 6ft and 4ft the main difference would appear to be that there is an increase in the amount of close errors when the 6ft distance is used and a small increase in the number of far errors made when the 4ft distance is used.

*6.3.1.2.4.3 Size * Distance*

The 2-way interaction comparing the size of the image and the distance of the individual failed to reach significance when using the lower bound epsilon value for degrees of freedom [$F(1, 60) = 3.37$, $p < 0.1$]. Further analysis using post hoc Tukey HSD tests however indicated some significant differences between the directions of errors in specific conditions. Table 6.6 displays the direction of the errors made.

	Small	Medium	Large
4ft	Far***	ns	Close***
6ft	ns	Close***	Close***

Table 6.6 Direction of errors made in size by distance conditions.

Post hoc Tukey HSD tests indicate that at both large conditions, close errors are more prevalent than far errors, although this is greater in the 6ft condition than at 4ft. There are also significantly more close errors than far errors in the medium 6ft condition, while in the small, 4ft condition there is significantly fewer close errors than far errors. It appears clear that large image sizes increase the number of errors made by guessing too close, while small images increase the number of far errors. More errors of too close also appear to be made to the 6ft condition as opposed to the 4ft condition.

6.3.1.2.5 4-Way Interaction (direction * background * size * distance)

Although this interaction failed to reach statistical significance, a Tukey HSD test was used to compare the direction of errors for each condition. Analysis of the error breakdown for the 4-way interaction is displayed in Table 6.7.

	Small	Medium	Large
Blank, 4ft	Far***	Far***	Close*
Blank, 6ft	Far***	ns	Close***
Objects, 4ft	Far***	Close***	Close***
Objects, 6ft	Close*	Close***	Close***

Table 6.7 Mean frequency of errors made by size and distance (%)

* $p<0.05$; *** $p<0.001$

This analysis indicates the relative strengths of each variable in terms of whether they cause errors to be made by guessing too far or too close. This shall be summarised in terms of the whole data.

6.3.1.3 Summary of Error Data

When considering the data, it must be remembered that the percentages refer to the total possible number of errors in any condition. Therefore they are standardised across the whole range of conditions studied and although linked, the frequencies of each type of error for any one condition are not directly correlated (if one increases, it does not necessarily precipitate a decrease in the other).

The data shows that some of the variables do cause differences in the impression of distance in the photographs. The particularly strong effects appear to be those of making the picture either very large or filling the image background with objects, both of which cause individuals to appear closer to the camera than they are. Just as powerful, although in the opposite direction is the effect of making the image small, which leads to the individual appearing further from the camera than they actually are.

Less powerful effects of having a blank background or a medium sized picture appear to exist. The blank background appears to lead to a weak representation of distance, while the medium background may not lead to any direct effect due to it's position directly between small and large.

Results were complicated by the effect of the different distances. There was a tendency for people to guess towards the middle on each distance (too close more common for 6ft and too far for 4ft). These were weak effects however, 6ft being balanced by the blank condition and 4ft being overshadowed by the objects background.

One of the most interesting factors of the results is that the effects of the variables appear to be cumulative or able to cancel each other out. To summarise the variables it is useful to consider them in terms of weak or strong and in which direction they tend to cause errors. Table 6.8 summarises the effect of each condition.

Strength of Variable	Direction of errors	
	Too Close	Too Far
Strong	Object background	Small
	Large	
Weak	6ft	Blank background
		4ft
Nothing	Medium	

Table 6.8 Summary of the relative strengths of each condition with the direction of the most common errors.

Table 6.8 can be used to work out the direction of the majority of errors using any combination of background, distance and size of an image. To do this the following rules must be applied:

1. All conditions are additive

2. A strong condition will not be cancelled out by two weak conditions

3. Each condition must have three compatible variables (one background, one size and one distance). Comparisons across conditions (e.g. only Medium) must take into account all of the possible conditions and treat them additively.

These results have shown that both image size and background have an effect on the impressions of perceived distance of individuals in photographs.

6.3.2 Performance

The task proved to be a fairly difficult one, the mean total score being 57% (range from 33% to 79%, median 56%). This is well above chance of course, and shows that the task was suitably difficult so as to avoid possible ceiling and floor effects.

The data was analysed using a 3-way ANOVA (2x3x4) for repeated measures. The independent variables were background (blank or objects), size (small, medium or large), and distance (2.5ft, 4ft, 6ft, and 9ft). Data was not analysed in terms of the different characters used in the photographs, as all hypotheses were assumed to apply to the perception of all human individuals, any differences caused by the character would be merely coincidental. For this analysis all data was included, the dependent variable was the percentage of correct answers.

Most of the data failed the Mauchly Sphericity Test so the Greenhouse-Geisser (epsilon corrected averaged F) test was again used to make the degrees of freedom more conservative. Results of the ANOVA indicated a significant main effect of distance [$F(3, 180) = 17.88$; $p > 0.01$] and significant 2-way interactions of background by distance [$F(3, 180) = 37.01$; $p > 0.01$], size by background [$F(2, 120) = 18.97$; $p > 0.01$], and size by distance [$F(6, 360) = 70.95$; $p > 0.01$]. The 3-way interaction was not significant [$F(6, 360) = 1.85$; $p > 0.05$].

6.3.2.1 2-Way Interactions

The repeated measures ANOVA indicated that there were significant interactions for all three 2-way interactions. Each of these results shall be discussed in turn.

6.3.2.1.1 Background * Size

The 3-way ANOVA indicated a significant interaction of background and size of image [$F(2, 120) = 18.97$; $p < 0.01$]. The mean score for each condition is presented in Table 6.9 and graphed in Figure 6.7.

	Mean Score (%)	
	Blank	Objects
Small	57.25	64.63
Medium	59.75	58.31
Large	62.50	51.56

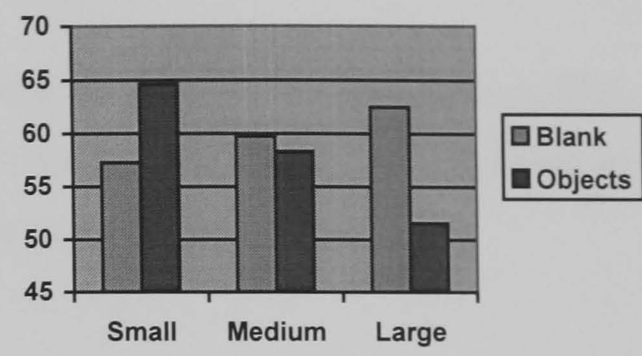


Table 6.9/Fig 6.7 Mean score for size by background (across distances)

Post hoc tests of simple main effects showed that there was no significant effect of size when the blank background was used [$F(2, 120) = 2.02$; $p > 0.05$]. When the object background was used, there was a significant simple main effect of size [$F(2, 120) = 18.50$; $p < 0.01$] scores were highest when the image was small and lowest when the image was large.

The post hoc tests of simple main effects also showed a significant effect of background at both the small size [$F(1, 60) = 6.37$; $p < 0.05$], and the large size [$F(1, 60) = 33.39$; $p < 0.01$]. There was no significant effect of background when the image was a medium size [$F(1, 60) = 0.42$; $p > 0.05$].

These results show that when the blank background is used, the size of the image has no effect on the likelihood of guessing the correct distance of the individual in the picture. When the background is filled with objects however, participants are most likely to guess correctly if the image is small, and least likely to guess correctly if the image is large. The size of an image has more of an effect on the impression of distance if the background is filled with depth cues than if it is blank.

When considered alongside the results of the error data, the reason for this result appears clear. The results of the error data showed that the effect of each variable on the perception of distance was cumulative. The error data also showed that both the large pictures and the object background tended to make the individuals in the pictures appear closer. When both are presented together, then the performance drops significantly. Performance is significantly higher when conflicting variables are present, if one condition pulls one way and another pulls the other then they cancel each other out and the correct answer is much more likely to be chosen, as is the case with small pictures with an object background.

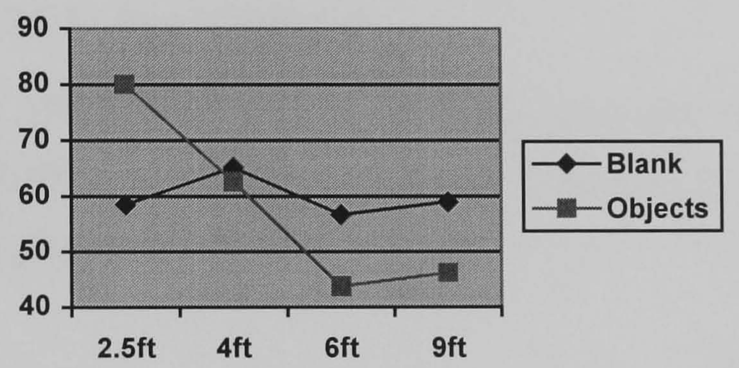
6.3.2.1.2 Background * Distance

The ANOVA indicated a significant interaction of background and distance [$F(3, 180) = 37.01$; $p<0.01$]. Table 6.10 presents the mean scores in each condition by background and distance and these are graphed in Fig. 6.8.

	2.5ft	4ft	6ft	9ft
Blank	58.5	65.17	56.67	59
Objects	80.08	62.67	43.75	46.16

Table 6.10 Mean score (%) for conditions of background by distance (across sizes).

Fig 6.8 Mean score for background by distance



Post hoc tests of simple main effects revealed that distance had a significant effect on performance when the object background was used [$F(3, 180) = 41.76$; $p < 0.05$]. When the blank background was used, there was no significant main effect of distance [$F(3, 180) = 2.26$; $p > 0.05$].

Post hoc tests of simple main effects also showed that there was a significant effect of background at all distances except 4ft [$F(1, 60) = 0.56$, $p > 0.05$]. At 2.5ft performance was significantly higher when objects were present in the background [$F(1, 60) = 87.46$, $p < 0.01$]. At 6ft performance was significantly higher when the background was blank [$F(1, 60) = 14.97$, $p < 0.01$]. At 9ft performance was also significantly higher when the background was blank [$F(1, 60) = 33.54$, $p < 0.01$].

When objects were present in the background, performance was highest in the closest condition (higher than any other condition), and lowest in the two distant conditions (lower than any others were).

To explain these results, it is again helpful to refer back to the results of the error data. The object filled background, made individuals appear closer than they were. This explains why there is such a high score in the 2.5ft condition with the object background. At the 2.5ft condition, there is no opportunity for participants to guess too close, so all those to whom the individual appeared closer would have guessed the correct answer as well as those who would have got it correct anyway.

The closeness effect that the object background caused also explains why the scores for the objects background at 6ft and 9ft are lower than the equivalent for the blank background which appears not to exert much effect. Why this does not occur at 4ft shall require further discussion later.

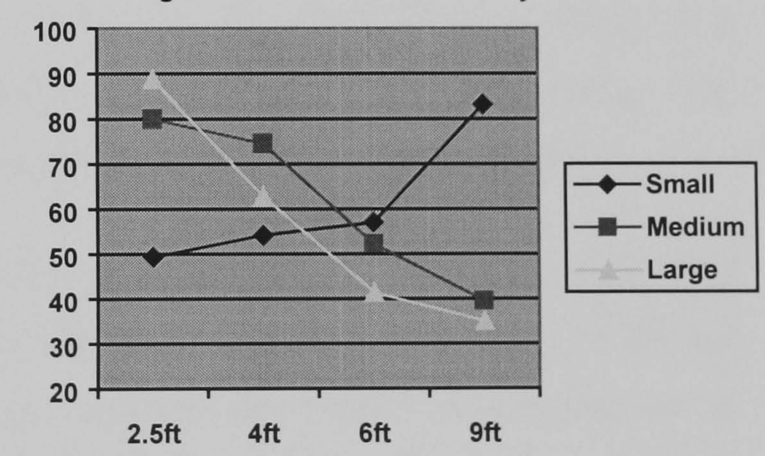
6.3.2.1.3 Size * Distance

The ANOVA indicated a significant interaction of image size and distance [$F(6, 360) = 70.95$; $p<0.01$]. The mean scores for each condition are presented in Table 6.11 and graphed in Fig. 6.9.

	Small	Medium	Large
2.5ft	49.38	80.03	88.75
4ft	54.25	74.62	62.88
6ft	57	52.25	41.38
9ft	83.25	39.5	35

Table 6.11 Mean score (%) for size by distance (across backgrounds).

Fig 6.9 Mean scores for size by distance



Post hoc tests of simple main effects revealed significant simple main effects of distance at each size, and of size at each distance. The results of this analysis are presented in Table 6.12

	df	F	p
Size at 2.5ft	2, 120	70.18	< .01
Size at 4ft	2, 120	13.73	< .01
Size at 6ft	2, 120	11.76	< .01
Size at 9ft	2, 120	116.83	< .01
Distance at small size	3, 180	25.18	< .01
Distance at medium size	3, 180	29.49	< .01
Distance at large size	3, 180	77.13	< .01

Table 6.12 Summary of analysis of simple main effects for size by distance (across backgrounds).

The results of the analysis presented in Table 6.12, when taken along with the results of Fig 6.9 show that performance for the 2.5ft items is significantly decreased by using a small image size. The opposite is true for the 9ft items, when a small image leads to a significant improvement in performance. For the 4ft items, performance is best when the image is a medium size, while for 6ft items the large size leads to significantly lower performance than the small size.

There are also significant effects of distance on the performance to items of a particular size. For small items, performance is improved in the 9ft condition. For both medium and large images, performance improves as the distance decreases.

At the closest distance, large pictures led to the greatest scores, but medium also led to good performance, while performance was poor in the small condition. At 4ft, the poor performance with small pictures remained and was accompanied by a significant drop in the performance of the large pictures, while medium pictures continued to lead to good performance. At 6ft the small images again remained low, but were not significantly different to the performance of the medium pictures, and both gained significantly higher scores than the large pictures. Finally at 9ft, the performance on small pictures increased significantly to give high scores, while both the medium and large pictures led to a further drop in performance, significantly lower than the scores for small images.

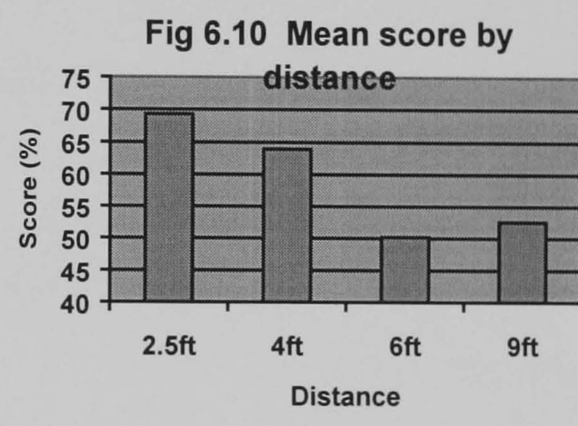
The reason for these results can partially be traced to the error data and the particular pulls of the different sizes of image. Those conditions that tended to lead to feelings of closeness (large and medium) showed highest performance at the 2.5ft condition as expected (although for medium, performance did not drop when the distance changed to 4ft). Results of the small condition were predicted by the fact that small pictures give a feeling of distance. This meant that the highest score was in the 9ft condition (where errors of too far could not be made) and none of the other conditions differed significantly from each other.

6.3.2.2 Main Effects

The analysis indicated that there was no significant main effect of background [F(1, 60) = 1.03; p >0.05] or size of image [F(2, 120) = 2.23; p > 0.05] but that there was a significant effect of distance [F(3, 180) = 17.88; p < 0.01]. Table 6.13 shows the mean score in each distance condition, and these are graphed in Fig. 6.10.

Distance	Score (%)
2.5ft	69.25
4ft	63.92
6ft	50.21
9ft	52.58

Table 6.13 Mean score (%) for each distance.



Post hoc paired samples t-tests showed that there was no difference between the number of correct answers given to pictures at 6ft and 9ft, but both were fewer than the number correct at 4ft which was fewer still than the number of correct answers at 2.5ft. The fact that the 2.5ft distance had the highest score suggests either that this was much easier or that there were more variables, which caused impressions of closeness. As the error data showed that large pictures and the object filled background both made individuals appear closer than they in fact were, compared with only the small pictures strongly making people appear farther away then it would be expected that overall the closest image would have the highest performance score. The reason for the improved score in the 4ft condition is less clear and shall be discussed later.

6.3.2.3 Summary of Performance Results

Some of the results of the performance analysis are directly predicted by the influences of each variable as shown by the error data. Those variables that were shown to have a strong effect of making the individual appear close (object background and large size) had the best performance at the 2.5ft distance. Similarly, the small images, which were shown to lead to an impression of distance, led to the highest performance at the farthest distance. Of the two conditions which seemed to lead to a weak impression of either closeness (medium size) or distance (blank background), neither had significantly improved scores at the extreme ends of the distance scale, suggesting that any impression of distance these caused was indeed a weak one.

There were also some results found that are not adequately explained simply by the influence of condition. Those that require further explanation involve significant differences between conditions not involving the extreme distances. Specifically it is the 4ft condition that tends to score more highly than would be predicted. For all variables except the small sized image, the 4ft distance scores more highly than the 6ft. Also of interest is that for both the large and the medium sized images, scores are higher in the 6ft condition than the 9ft condition.

6.3.2.3.1 Ease of Answer

It is reasonable to expect some distances must be easier than others to recognise. However the likely candidates are the two extremes - at 9ft there are hands visible and at 2.5ft there is only the head visible. The differences in image to background ratio are not equal either - there is a much larger difference between 9ft and 6ft (1:11.1 vs. 1:6.9) than between 6ft and 4ft (1:6.9 vs. 1:5.4) suggesting it is easier to differentiate between the 6ft and 9ft image than the 6ft and 4ft image.

One measure of the ease of recognising an answer is by recording the hit rate. Table 6.13 compares the total number of times each distance was guessed as the answer, with the number of correct guesses for each distance. The hit rate is taken as the percentage of guesses that are correct for each distance.

Distance	Mean Score (%)	Total No. of Guesses	Actual No. of Correct Guesses	% of Guesses correct (hit rate)
2.5ft	69.25	1274	1014	79.59
4ft	63.92	1922	936	48.70
6ft	50.21	1675	735	43.88
9ft	52.58	978	770	78.73

Table 6.14 Percentage of guesses of each answer that were correct.

Table 6.14 shows that the hit rate is substantially higher in the two extreme conditions than in either of the middle conditions. When a participant guessed 4ft or 6ft for an image, they were correct less than half of the time, this compares with being correct nearly 4 times out of 5 in the extreme conditions.

A high hit rate suggests that when a participant guesses a particular distance they are fairly sure and are unlikely to guess that distance when unsure. Conditions with a high hit rate are also easier than those with a low hit rate as it can be assumed that more of the correct hits are accounted for by chance in the conditions with a lower hit rate.

This finding may also explain why for each of the middle images, mistakes tended to be in the direction of the other middle answer rather than towards the extreme. As extreme conditions were easier to spot than the middle conditions, when a mistake was made it was more likely to be by guessing one of the harder conditions.

The hit rate data directly accounts for the increased performance in the 4ft condition. The 4ft answer was guessed nearly twice as often as the 9ft answer. The number of correct hits gained by chance is therefore likely to be much higher in the 4ft condition than the 9ft condition. If those successful scores gained by chance were removed, performance in the 9ft condition would be significantly higher than in the 4ft condition.

Since the extreme distances were easier to get right, it is fair to say that for a variable to cause an effect on perceived distance which made a middle distance appear like an extreme distance (e.g.4ft to 2.5ft) is a much greater effect than to make a middle distance appear like the other middle distance (e.g. from 6ft to 4ft). Since there were more variables which seemed to make the individuals in the pictures appear closer than they were, it would be expected that there would be a greater performance in the 4ft than the 6ft condition as the effect required to make the 4ft image appear like the 2.5ft is greater than the effect required to make the 6ft image appear like the 4ft.

The fact that more guesses were made towards the closer conditions than the farther also backs up the finding that there were more conditions which pushed towards feelings of closeness rather than distance.

6.4 General Discussion

The aim of this experiment was to investigate visual cues that may contribute to impressions of proximity, with specific emphasis on the effect of background and size of image. The results have shown that both of these variables have a significant impact on impressions of distance. A background rich in depth cues provided by objects leads to an impression of closeness, while a blank background would appear to have little effect (although relative to the object background the blank background obviously leads to an impression of distance). Similarly, the size of the picture also affects impressions of distance, large images make the subjects of the photograph appear closer, while small pictures make them appear far away.

Not only do different variables produce similar effects of apparent proximity, these variables have been shown to have a cumulative effect. Large pictures with an object background produce scores significantly lower than those in the large blank condition, as participants made more errors of too close. Equally variables can cancel each other out – as in the case of the small photographs with an object filled background.

The results showed a general bias towards participants underestimating the distance that the subjects of the photographs were sitting at. It is notable that in the perceptual experiments of Witmer & Sadowski (1998), participants also tended to underestimate distance, despite their task being motor based, rather than cognitively based. These results suggest that underestimation in the visual system is more common than overestimation.

The results also indicate which combinations are the best for producing particular impressions. To make someone appear close without altering the camera in any way, they should be presented in a large picture, with a background rich in depth cues. To make someone appear average they should be in a moderately sized picture with a blank background and to ensure someone appears remote then they should be presented in a small picture with a blank background.

6.4.1 Object Image to Background Ratio

The findings of Study 1, presented in Chapter 3 indicated that the dominant determinant of an individual's perceived distance on video was the size of the object image per se as opposed to the object image to background ratio. One of the aims of the present experiment was to investigate whether this translated to distance perception in photographs.

In the present experiment the object image to background ratio increased as the actual proximity increased (9ft → 2.5ft). The object image to background ratio remained unchanged across the different sizes of pictures.

As the size of the picture increases, the size of the object image increases, while the image:background ratio remains unaffected. If people judge distance primarily from the size of the object image, as opposed to the image:background ratio, then changing the size of the image will have an effect on the perceived distance. The predicted effect would be that as the size of the picture (and consequently the size of the object image) increases, the perceived distance decreases. Should the ratio be the key factor then altering the size of the picture will have no effect.

Since the results show that indeed the size of the pictures does have an effect on the perceived distance and it is in the hypothesised direction, the results indicate that the perception of distance is at least partially influenced by the size of the object image per se.

It should be noted that this has not shown the image:background ratio to be totally unimportant but merely that it is not the fundamental basis of distance perception. The present experiment did not systematically alter the image:background ratio while keeping the object image size the same as it was considered more important to compare the sizes of images as the same as those of the video window used in Study 1.

6.4.2 Depth Cues

The effect of the object filled background was to make the individuals in the photographs to appear closer than they in fact were. It had been hypothesised that the presence of depth cues would lead to an increase in the performance of individuals rating the photographs with the object filled background. This hypothesis was based on the fact that these photographs contained a reference point that remained static across all pictures with that background. In fact 46% of responses to the object background were errors.

One reason for the lack of this skill may be due to the fact that people were not viewing pictures together and could therefore only compare from memory. Because individuals were, however, comparing photographs in their mind with previous ones – some of which would be blank and some with objects, then concentrating on the position of the subject in relation to the objects may not have been the chosen strategy, as this was of no help when the photograph had a blank background. It is possible that were this experiment to have been run in blocks – with all the object photographs coming successively then participants would have chosen a strategy based on the depth cues and consequently had higher scores for the object conditions. Similarly, had participants been able to compare pictures alongside each other, rather than only being able to look at one at a time, performance may have been improved in the object conditions.

Alternatively, it may be that the abundance of objects in an image leads to an impression of closeness due to some psychological phenomenon that is beyond the scope of this thesis.

6.4.3 Personal Space

One of the conclusions drawn from the research using video conferencing is that feelings of personal space can be invaded by someone appearing very close (Okada et al., 1994; Reeves and Nass, 1996). It is fair to say, however that in generating these feelings of how close someone feels, an individual is not spending time attempting to work out exactly how far away they look. It may be more useful to consider cues to distance being more psychological than physical. In order to ask how close someone is, perhaps participants rely to an extent on how close they *feel* rather than how close they *look*. How far something feels could be based less on physical factors such as occlusion, and more on a less definable impression of closeness, using such factors as size of image. If this is the case then it is possible that this effect may only apply to human beings and be related to the feelings of personal space experienced when using video-mediated encounters.

6.5 Conclusions

This experiment has shown that the visual perception of proximity can be affected in more ways than simply altering the zoom on a camera. This research has shown that the size of an image is positively correlated to feelings of close proximity, backing up the findings of Study 1. A less expected finding was that the presence of an object filled background also leads to increased impressions of close proximity. The presence of physical depth cues in a photograph was not shown to lead to any increase in the accuracy of perceiving distance.

The consequences of this research are likely to be that impressions of proximity can be manufactured in more ways than simply by the placement and settings of the camera. Any psychological or behavioural effects influenced by apparent proximity may also be induced by more subtle differences between systems, which may not even require technological differences to be effective.

The research does not claim to be exhaustive in terms of all the possible variables in a picture/video link that may cause an effect on the perceived distance of an individual but does indicate that such effects can and do exist.

7 Chapter 7 . Study 4: A Comparison of Image Size and Camera Zoom in a Simulated Service Encounter

7.1 Introduction

Studies 1 and 3 have both highlighted the effect of image size on impressions of distance. In Study 1, participants reported that a remote individual felt closer if the image was larger. Study 3 showed similar results, with the finding that individuals appeared closer in photographic representations if the images were larger. The present study sought to apply this finding to the original interactive study (Study 2) to investigate the specific effect of image size on behaviour in a video-mediated interaction.

The results of Study 2 showed that participants were affected by changes in the amount of camera zoom that was used. The aim of the present experiment was to isolate the effect of image size from the effect of camera zoom - which contains both alterations in image size and image to background ratio – and to investigate it in an interactive task. The experiment also sought to extend the results of Study 3 by studying various psychological factors not studied in the previous experiment. These factors included understanding, body positioning, and a greater examination of trust and persuasion.

7.1.1 Body Position

One of the aspects of behaviour that was not recorded in any of the previous studies is perhaps the most fundamental behaviour regarding proximity, that of body position. The start of an interaction will often be the positioning of individuals prior to starting conversation. In naturalistic settings, both individuals are usually able to control their body positioning and face-to-face studies of proximity in conversation have shown that usually people act to ensure an appropriate personal distance between interlocutors.

Argyle & Dean (1965) put forward a theoretical model which they called the approach-avoidance equilibrium, where people altered their body position to seek a certain degree of proximity. Alterations in body position could be as explicit as stepping forward or backwards or less obvious such as leaning forward or back or altering their bodily orientation. The model states that should individuals fail to attain what they consider to be an appropriate interaction distance then they will feel uncomfortable.

Sundstrom & Altman (1976) expanded on this theory by examining the approach-avoidance equilibrium comparing the behaviour of strangers and friends. In their experiment, Sundstrom & Altman found that when interacting, strangers showed little or no approach behaviour although they did show avoidance behaviour. This research also confirmed the research of Hall (1966) that the distance that friends were most comfortable interacting at was usually closer than that selected by strangers.

When applied to the concept of perceived distance, this model makes predictions about individuals' behaviour in a setting where differences in perceived proximity are present. In Study 2, body position was not recorded, although feelings of comfort were recorded by questionnaire response. The results of the questionnaire showed that there was no difference in feelings of comfort between the two proxemic conditions. The approach-avoidance model of Argyle & Dean (1965) would suggest that there should have been some avoidance behaviour in the close condition and some approach behaviour in the far condition in order to equalise the perceived distance between conditions. Since other behaviours differed between conditions, it is clear that the equilibrium could not have been totally reached, but it is unknown whether participants made any attempt (either conscious or unconscious) to alter the perceived distance of the financial advisor.

The present study will investigate whether participants alter their body position when interacting with a remote individual, in relation to the appearance of the other person. If a remote individual really seems to be appearing close, then it may be hypothesised that participants will alter their body position in the close condition to establish a greater interpersonal distance between themselves and the remote advisor. If, on the other hand, the remote advisor does not feel uncomfortably close then no avoidance behaviour would be expected.

Mehrabian (1972) showed that leaning forwards was linked to proximity, gaze, touching and direct orientation. He put these together to name a dimension of immediacy, which he said was employed by people who liked each other more and by women more than men. Gaze analysis was beyond the scope of the present research, as it is not truly available in a video mediated encounter. Touch was not available either in this type of interaction and therefore, the concept of immediacy is reduced to bodily orientation and distance.

7.1.2 Image Size

In order to determine more fully the relationship between perceived distance and communicative style it is necessary to compare the effects of altering camera zoom with the effects of image size. This should show whether both variables, each of which alter impressions of perceived distance, have similar effects on communicative style. Comparing the two variables should also indicate which is the more important as far as creating impressions of perceived distance in an interactive communication context.

7.1.3 Live interaction vs. photographs

In study 3 investigating the impact of camera zoom and image size on the perceived distance of individuals in photographs, it was found that both the size of the image and the size of the object image to background ratio (level of camera zoom) were important in determining how far away an individual appeared. The results showed that large pictures made individuals appear closer and small pictures made the same individuals appear further away. The presence of objects in the background also made the individual appear closer.

In Chapter 3 it was shown that changing the size of a video window had an impact on impressions of distance. This study showed that the size of the object image (i.e. the head) was the defining feature of perceived distance. The object image to background ratio (level of camera zoom) was less important except at very close distances.

It should be noted that there are obvious differences between a live video image and a static photograph. Equally there are also differences between simply viewing a live video image and interacting with a live video image. The most notable of these differences is the presence of audio. Perceived distance contains more than merely visual cues, most notably audio - changes in volume are certain to enhance or decrease impressions of perceived distance.

7.1.4 Camera Zoom

The effect of increasing the level of camera zoom is to decrease the focal length of a camera. The net effect on the projected image is that a much smaller field of vision is projected, and the overall size of the image is increased, albeit at a lower resolution. The equivalent in the natural world is for an individual to focus on one object and move closer to it. The closer one is to the object, the more of the field of vision it takes up. Just as an individual always has the same size of field of vision, so does any particular camera.

The interesting result from the experiment which investigated the effect of the size of a video image on perceived distance was that individuals did not use the size of the object image to background ratio to determine perceived distance, instead using simply the object image size, except at the very closest distance.

7.1.5 Life-Size Images

One of the most common explanations for a lack of difference between the results of using audio-only versus video-mediated systems is the small image that desktop video conferencing uses. In a review of the video-mediated literature, O'Donnell & Draper (1997) commented that "it seems reasonable to suggest that by reducing the video from life size, will detract from the usefulness of having a visual channel". This is based on reasoning that such factors as expression and gesture may become less noticeable in a smaller image.

For this reason, the present experiment compared a small image size which corresponded to usual desktop video conferencing size and a large image size which was life-sized. It should also be pointed out however, that a decrease in resolution of the image may also decrease from the usefulness of the visual channel in the same way as a decrease in size.

Previous research using life sized images has been sparse, but anecdotal evidence suggests that it may be popular. Okada et al. (1994) used the MAJIC system which used life size video projection for interaction between two people, they reported that participants who used the system at a trade fair were more friendly and outgoing than in usual video conferencing demonstrations. Participants also reported that sometimes they felt pressure when the image was zoomed in very close, but when students who knew the demonstrators used the system, they reported that they preferred the image close, or even larger than life size. Okada et al. suggest that this is because they wanted to talk at personal distances.

For the present experiment, the financial advice experiment was repeated to compare the effects of image size with the effects of camera zoom, using projection equipment to provide a life size image as well as a monitor for the small image. The experimental design was a between subjects design with two independent variables - size (life-size vs. monitor) and zoom (close or far).

The 4 conditions were:

Life-size Close (LC) - Projection screen, zoomed in.
Life-size Far (LF) - Projection screen, zoomed out.
Monitor Close (MC) - Monitor, zoomed in.
Monitor Far (MF) - Monitor zoomed out.

Since large images are known to lead to a greater feeling of closeness, it was hypothesised that the financial advisor would appear closest in the LC condition and farthest in the MF condition. Both LF and MC conditions provided conflicting information and consequently could indicate whether image size or camera zoom was the stronger influence on perceived distance.

Every effort was made to make the quality of each condition comparable but inevitably the projector image was of a lower resolution, leading to a slightly more blurred image. Blurring has been shown to be a pictorial cue to depth perception (O'Shea, Govan & Sekuler, 1997) and thus the projection conditions arguably contained an inherent perception of distance which might partially counteract the impressions of closeness caused by the increase in image size.

7.1.6 Understanding

The results of the previous financial advice experiment indicated that customers were more interactive in the close condition than the far condition. The impact of this may go beyond simply a difference in dialogue pattern as in the past researchers have suggested that there may be a link between interactivity and understanding (Clark & Schaefer, 1987; Schober & Clark, 1989).

The theory, known as the *collaborative view* (Schober & Clark, 1989), is that conversation involves participants trying to establish the mutual belief that the listener has understood what the speaker meant. This collaborative process is called *grounding*, and it requires actions by both the speaker and listener (Clark & Schaefer, 1987). The alternative view is the *autonomous view* of understanding which asserts that merely hearing and seeing all that happens and having the same background knowledge is sufficient to understand fully. Schober & Clark (1989) showed that people who overheard instructions being given did not perform as well on a task as those to whom the instructions were given. Those receiving the instructions could ask questions, clarify etc., those who overheard could only listen to all that was said.

This suggests quite strongly that participation in an interaction aided understanding and memory in comparison to listening. The implications for the present experiment should be clear. Those who have a more interactive conversation may well understand more than those who sit and listen more. Consequently those in the close condition may well have a better understanding than those in the far condition.

A difference between this study and the Schober and Clark study is that they only went so far as to compare participating in an interactive conversation with listening to an interactive conversation but not taking part. It could be suggested that there is actually no evidence that once in an interactive conversation (as all participants in the present study were) the level of interactivity matters.

To determine the nature of the link between understanding and interactivity, the present experiment introduced a surprise recall test at the end of the experiment. Participants were interviewed after the task was complete and asked questions about the four options they had chosen between. It was hypothesised that the more interactive dialogues would lead to a better understanding than more "lecture style" dialogues.

7.1.7 Social Factors

The previous financial advice experiment generally found no significant effect of camera zoom on conscious impressions of the task or the technology. It is possible that this was because the questionnaires five point scale was too blunt in terms of its ability to pick up differences in individuals' impressions. In order to remove this doubt, the questionnaire was extended to include more questions about persuasion and trust, and answers were given on a 10 point scale.

The reason it was still considered important is that there is considerable evidence that individuals are perceived differently depending upon whether they are visible or not (e.g. Argyle, 1978; Harper Wiens & Matarazzo, 1978; Cook, 1979). Morley & Stephenson (1969, 1970, 1977) and Wichman (1970) demonstrated that persuasion is more successful when individuals could see each other, whether by video or face-to-face.

Altering the physical distance between individuals also has been shown to affect individuals' ability to detect intentions (Ekman, 1973). Using video-mediated technology, Reeves & Nass (1992) showed that differences in perceived distance could lead to differences in the participants' evaluations of pictures, attention and memory. Specifically, Reeves & Nass found that participants' evaluations of people in pictures were more extreme with closer images. The questionnaire sought to evaluate if this was the case in the present experiment.

7.2 Method

7.2.1 Scenario

The present study used the same scenario as Study 2. Participants were told to role play the character of a young customer with steady employment who had just inherited five thousand pounds and was seeking to invest it as well as they could. In order to do this the customer had gone to their local automated branch of their bank to seek advice, sat down at one of the kiosks and used the application to find the part offering financial advice.

7.2.1.1 Options

There were four options available to the customer, each of which presented a page of information. The information was taken from real leaflets advertising genuine options available in banks and could only be viewed one page at a time. Both the financial advisor and the customer could alter the page being viewed.

The options offered were as follows:

Bonus Account: A high interest but taxable account offering saving with 120 days notice required for penalty free withdrawals.

TESSA: A 5 year tax free savings plan offering access only to 80% of the interest but guaranteed high interest with no risk.

Shares: A high risk option capable of making the greatest capital gains but guaranteeing nothing.

PEPs: A lower risk option than the shares but still based on the stock exchange. Opportunity for higher gains than the no risk accounts but lower than the shares. Has running charges.

7.2.1.2 Financial Advisor (FA)

The role of the FA was played by a 27 year old actress with experience as an audio call centre agent. The FA was instructed to subtly attempt to persuade customers towards the bonus account on the premise that this would make the most money for the bank. The advice should, however, appear to be independent as it was most important for the customer to be happy with the advice given. The FA was also asked to look into the camera as much as possible in order to give the impression of looking at the customer. The FA was paid for their participation.

This was a different financial advisor to the one used in Study 2. It was hoped that the losses in replicability would be more than compensated for by the advantages gained in the level of generalisability of the results. These advantages and disadvantages shall be discussed later.

7.2.1.3 The Customer (C)

The customer received instructions immediately prior to the experiment outlining the scenario. They were a young industrial designer, reasonably comfortably off, with a stable job. The rest of their situation was left up to them. The reason for leaving this open-ended was to stop participants searching the instruction sheet for answers if the FA had asked them something and also to make the task more interactive for them. In the present study, it was explicitly explained to the participant that they were role playing and they were to imagine themselves in the role of the character they have been given. This extra instruction was given because in Study 2, participants often seemed to forget that they were role playing.

7.2.2 The Conditions

The independent variables in this experiment were the level of camera zoom used to display the financial advisor to the customer (2 levels, close and far) and the size of the image used (2 levels, monitor or life-size).

7.2.2.1 Camera Zoom

The close condition involved the FA being seen as virtually just a face. The image was zoomed until all of the head was visible but no other vertical background. The far condition allowed plenty of background and a head and shoulders view of the FA.

As with Study 2, only the customers' view of the financial advisor changed and not vice versa. The advisor always saw the customer at a distance in between that of the two conditions. The reason for this is that this study was only interested in the effects of perceived distance on the customer, and so only visual information was altered. Had the visual information been altered for both parties then any effects found could be due either to the visual information directly or as a result of the reaction to changes in each others behaviour. The reasons for this approach were discussed in Chapter 4.

7.2.2.2 *Size of Image (system used)*

Participants either carried out the task using a monitor or a projection screen. The monitor was 43.18 cm (17 inches) with a 39.37 cm (15½ inch) viewable area. The video window on the monitor measured 100mm x 74mm. The projector screen showed a video window 400mm x 296mm, which corresponded to a life-sized image. The size of the image of a face in the life-sized far condition was larger than the size of the image in the monitor close condition.

The image on the projector screen was simply a direct projection of the image shown on the monitor. Consequently due to the larger screen size, the life-sized image was of a lower resolution than that on the monitor. Participants in each condition sat approximately 30 cm from the screen. Differences in physical screen size did lead to different viewing conditions, however as far greater head movement was required to read the text and to move attention from the advisor to the text. This was an unavoidable effect of using a larger screen.

In each condition the camera was placed at eye level to the left hand side of the screen (the side of the video window). The reason for placing the camera on this side was so that when the customer looked from the text to the financial advisor they were moving their attention towards the camera. Although this did not permit true eye contact, this did at least ensure that the visual information being given was not paradoxical.

7.2.3 Procedure

Participants were informed that they were going to take part in an experiment investigating new multimedia technology. It was explained that they would role-play the part of an individual who had just inherited five thousand pounds and was looking to choose the best way of investing it. Participants were given an instruction sheet to read which gave them information about their character.

Participants were then led into the experimental room, sat in front of a screen and given a headset microphone to wear. The experimenter then left the room and the financial advisor began the interaction. The interaction began with the financial advisor saying "hello, how can I help you" and ended with the advisor saying "ok well we'll send information on that choice out to you".

It is important to note that at no point beforehand did the participants have any contact with the financial advisor, who was a complete stranger to them.

Following completion of the experiment, participants completed a short questionnaire based on their impressions of the task, following completion of this they were given a surprise interview which asked questions assessing their understanding of the four options they had seen.

7.2.4 Technology

The screen the customer saw consisted of an information screen with four pages provided by the shared whiteboard in Microsoft NetMeeting v2.1. This allowed the presentation of coloured textual information. No animation, graphs or pictures were used. The whiteboard enabled both individuals to change the page being viewed and to highlight text if they so desired. All alterations made were immediately visible to both individuals.

Situated towards the centre of the left hand side of the screen there was a video window. Video and audio were transmitted directly rather than across a network, ensuring broadcast quality frame rate with no appreciable audio lag, and was viewed using the VDOPhone program version 3.02.002(124) (Product S/N: D87ACCC52C302216). Frame rate and audio lag have been shown in the past to have severely deleterious effects on the communication abilities of participants in a videoconferencing system (e.g. Cohen, 1982; Johannsen, 1984; Tang & Isaacs, 1993; Isaacs & Tang 1994; Whittaker, 1995). As technology continues to improve, it was felt to be more useful to test using technology which will soon become available, rather than that which will soon be superseded technically.

The cameras used were JVC videomovie GR-AX 55E. They were placed to the side of the screen as near as possible to the video window in order to get an experience as close to eye contact as possible.

Camera zoom was altered by changing the focal length of the video camera, courtesy of the zoom function. For the close condition the camera was zoomed in until only the head was viewable vertically (allowing some background horizontally). The far condition allowed a view of the whole of the upper torso and plenty of background. At this point it is worth acknowledging that there is more to perceived distance than simply visual cues, but during the experiment no differences in sound were made, and thus only visual proxemic information was altered. While it would be scientifically interesting to investigate altering the sound as well to give a greater proxemic effect it was still considered worthwhile to investigate purely visual cues.

7.2.5 Participants

There were 60 participants (34 males and 26 females; age range 16-34, median = 21) who were all undergraduates and graduates from the University of Glasgow. One older participant (aged 60) also completed the experiment but was included only for comparison, their results were not analysed with the others for fear of skewing the data. No participant had ever used any banking by video conferencing equipment before, and no-one rated themselves as very inexperienced with computers.

Participants were told they were going to take part in a psychology experiment investigating multimedia technology. It was explained that they would be audio and video taped for research purposes but that these would be treated as strictly confidential. Participants were also informed that the experiment had no time limit and that at the end there would be a questionnaire, they would be paid £5 for their participation. All participants signed their consent.

7.2.6 Questionnaires

At the end of the experiment, participants completed a questionnaire designed to investigate their impressions of three areas. A copy of the questionnaire is shown in Appendix C. The questionnaire in Study 2 failed to show any significant differences between conditions for any of the measured scales. Because of this the questionnaire was extensively remodelled from that used in Study 2. Scales were made longer and more questions were used to address each individual issue. The aim of this remodelling was to show whether or not the problem lay in the format of the questions or if the conditions actually caused no differences in individuals conscious opinions of various parts of the task.

Prior Experience: Prior experience of both videoconferencing and computers in general was monitored to ensure any effects found were not simply due to level of experience.

The Technology: This section investigated how participants felt about the technology itself, including whether they considered they would be willing to use it in real life, and how it compares to other methods of banking.

Social Factors: This section investigated how participants felt about the advisor and the advice given in terms of trust, persuasion, likeability etc.

7.2.7 Interview

After participants had completed the questionnaire, (roughly 10 minutes after having completed the task) they were given a short interview to investigate their level of understanding of the options that they had discussed during the task. This took the form of a recall task with 23 questions. All of the answers were available in the information presented so all participants had had the information presented to them, where they differed would be in the amount they had talked to the advisor. The point of the recall interview was to investigate how interactivity affects understanding. A copy of the interview can be found in Appendix F.

7.2.8 Dialogue Analysis

The main results of interest in Study 2 came in the dialogue analysis. The present Study uses the same method of analysis as Study 2. Dialogues were analysed for structure using the same measures as those outlined in Chapter 4. Of specific interest within the dialogues were speaking turns, instances of overlapping speech, words spoken, average words per turn, and backchannels.

7.3 Results

The projection experiment contained four analysable parts. These were:

The Dialogues. These were transcribed and structural measures (turns, overlaps, words, time) analysed.

The Questionnaire. Investigating attitudes towards the task and the technology used, and also the participants' subjective impressions of the advisor and the advice given.

The Interview. Investigating subjects understanding of the different options offered in the task.

The Video tapes. Investigating individuals body position during the experiment.

7.3.1 Results of the Dialogue Analysis

When analysing the dialogues it was clear that there were 2 extraordinarily long dialogues, both more than 3 standard deviations from the mean. For the purposes of analysing the data these dialogues were both considered outliers and were removed due to the extremely skewing effect they had on the data. Data was collected on an interval scale and analysed by using 2-way ANOVAs for independent samples. The independent variables were size (life-size or monitor) and zoom (close or far).

7.3.1.1 Time

It was hypothesised that conditions in which the advisor appeared closer would take a significantly longer time to reach completion than those in which the advisor appeared far away.

Table 7.1 shows the median length of time for each condition to reach completion. There were two dialogues which took much longer than any of the rest, both of which were in the life-size far condition. These dialogues skewed the data for that condition somewhat and thus median times are reported.

Condition	Median time (seconds)
Monitor Far	595
Monitor Close	560
Life-size Far	627
Life-size Close	683

Table 7.1 Median length of interaction (seconds)

A 2-way ANOVA for independent samples indicated no significant main effects of zoom or size and no significant interaction (zoom x size), although there was a trend towards the task taking longer in the conditions when the life-sized image was used [$F(1, 54) = 3.23$, $p < 0.1$].

This trend is in the predicted direction in the sense that interactions tended to be longer in the larger conditions, perhaps an early indicator suggesting that the size was slightly more important than the camera zoom in creating a feeling of closeness. As with all non-significant results however, trends must be treated with caution.

7.3.1.2 Turns

The number of turns was measured for each individual in each dialogue. By the nature of turns, and due to the fact that this was a two person interaction, the number of turns will always be very similar between both individuals. Analyses of the total number of turns in the dialogue and the number of customer turns were carried out and showed no significant differences between conditions. Table 7.2 shows the mean number of turns for each condition.

	Customer		Total	
	Close	Far	Close	Far
Monitor	53.40	50.87	94.93	92.40
Life-size	54.80	48.69	96.87	86.46

Table 7.2 Mean number of turns per interaction

A 2x2 ANOVA for independent samples indicated there to be no significant main effects or trends between conditions for either the total number of turns or the number of customer turns. There was also no significant interaction between size and camera zoom.

It had been hypothesised that there would be more turns in the conditions that the financial advisor appeared closer. The results have not shown this to be the case, although the mean was highest in the Life-size Close condition.

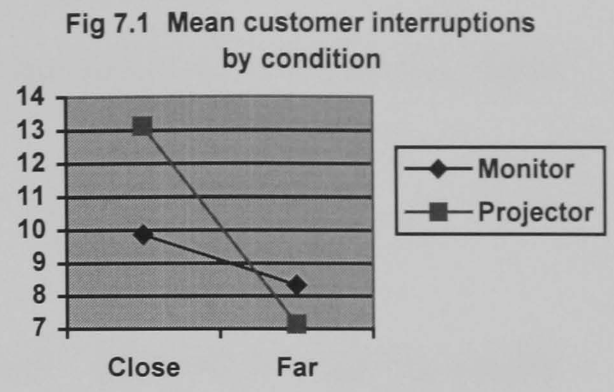
7.3.1.3 Overlaps

Instances of overlapping speech were measured and also recorded by role. Customer overlaps were taken to be instances where the customer spoke before the financial advisor had stopped speaking. Instances when the customer made an utterance merely to provide feedback and without interrupting the flow of the financial advisor were considered backchannels and not included as overlaps. A 2-way (2x2) ANOVA for independent samples was computed for both the total number of interruptions in a dialogue and the total number of customer overlaps in a dialogue. Table 7.3 shows the mean number of overlaps per condition.

Results of the ANOVA showed that there were no significant effects of size or zoom when comparing the total number of overlaps in a dialogue. Analysis of customer interruptions indicated a main effect of zoom which fell just short of significance [$F(1,54) = 3.98$; $p = 0.051$]. This indicates that there was a trend towards the customer interrupting more in the close conditions. There was no significant main effect of system and no significant interaction of zoom and system. The mean number of interruptions for each condition are plotted in Fig 7.1.

	Customer		Total	
	Close	Far	Close	Far
Monitor	9.87	8.33	13.93	13.67
Life-size	13.13	7.15	17.33	9.69

Table 7.3 Median number of interruptions per interaction



It has been hypothesised that there would be more instances of overlapping speech in the conditions in which the financial advisor appeared closer. The results in Fig. 7.1 show a definite trend towards the interruptions being higher in the Life-size Close condition and lowest in the Monitor Far condition. Although falling just shot of statistical significance this result does show a trend in the expected direction.

It is also useful to interpret overlaps in terms of the rate of overlaps occurring. This involves describing the number of overlaps as a function of the number of turns. To investigate this the number of customer overlaps were divided by the number of customer turns and multiplied by 100. This gives the percentage of customer turns which were overlaps. Mean interruption rates are presented in Table 7.4.

	Customer		**Total**	
	Close	**Far**	**Close**	**Far**
Monitor	17.95	14.85	13.67	13.23
Life-size	21.48	13.90	16.43	10.53

Table 7.4 Mean percentage of turns that are interruptions per interaction

A 2-way ANOVA indicated that there was a significant main effect of zoom [$F(1, 54) = 6.46$, $p < 0.05$] for the customer but no significant main effect of size, or any significant interaction. Analysis of the total data showed no significant main effects or interactions.

This result shows that there were more frequent instances of overlapping speech in the zoomed in conditions. This concurs with the earlier finding that customers interrupted more in the conditions when the camera was zoomed in but that there was no difference in the total number of turns.

7.3.1.4 Backchannels

The number of backchannels spoken by the customer was recorded and the results were analysed by using a 2-way ANOVA for independent samples. The results of this ANOVA showed no significant main effects and no significant interaction. None of the effects even registered as trends. Table 7.5 shows the mean number of backchannels spoken in each condition.

	Customer	
	Close	**Far**
Monitor	5.93	6.80
Life-size	4.67	7.23

Table 7.5 Mean number of customer backchannels per condition.

The financial advisor very rarely gave any backchannel responses, a result both of the fact that there were few long periods of customer speech and the style of the financial advisor was to be quiet when the customer spoke. This was a different style to the financial advisor in Study 2, the possible impact of this difference in style shall be discussed in detail later.

7.3.1.5 Words

The mean number of words spoken in each dialogue are reported in Table 7.6. It should be noted that the financial advisor spoke considerably more than the customer in all dialogues.

	Customer		Total	
	Close	**Far**	**Close**	**Far**
Monitor	302	385	1440	1540
Life-size	340	264	1676	1459

Table 7.6 Mean number of words spoken in a dialogue for the customer and in total.

A 2-way ANOVA comparing the means in each condition revealed that changing either the camera zoom or the size of image used, did not have a significant effect on the number of words spoken in a dialogue either by the customer or in total. There were no significant main effects or any significant interaction for the mean number of words spoken by either the customer only or in total.

Another way of interpreting the amount an individual speaks is in terms of an individual's share of the dialogue. Mean dialogue share for each condition is presented in Table 7.7.

	Customer	
	Close	**Far**
Monitor	18.72	23.06
Life-size	19.82	18.07

Table 7.7 Mean percentage of words spoken in a dialogue by the customer.

It had been expected that the customer would say more in the closer conditions than in the far conditions. An analysis of dialogue share by 2-way ANOVA for independent samples however, showed that there were no significant effects of condition or any significant interaction between zoom and image size. As with the analyses on turns and words, the results showed that there was no effect caused by condition in the present study.

7.3.1.6 Words per Turn

In study 3 it was found that in the close condition, customer turns tended to be shorter than in the far condition. It was therefore hypothesised that customer turns would on average be shorter in those conditions in which the financial advisor appeared closer.

The average number of words per turn is calculated simply by dividing the number of words spoken by an individual by the number of turns they made. This analysis includes backchannels which are generally only a single word in length. This should not affect the analysis however, as there was no significant difference in the number of backchannels in any condition.

	Customer		Financial Advisor	
	Close	Far	Close	Far
Monitor	5.29	7.21	16.67	20.22
Life-size	6.08	5.24	18.94	18.37

Table 7.8 Mean turn length by role and system.

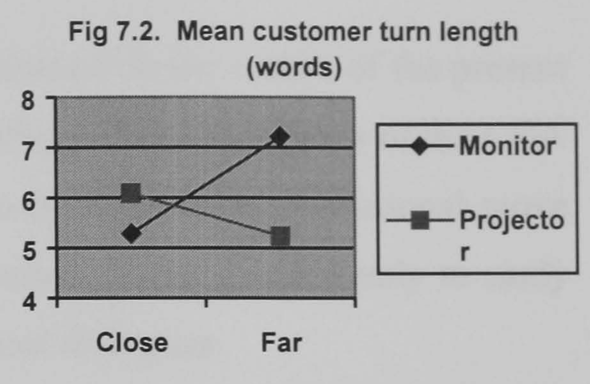


Table 7.8 shows the mean turn lengths for both the customer and the financial advisor. A 2-way ANOVA for independent samples indicated no main effects of image size or zoom but a significant interaction between the two variables [$F(1,54) = 6.22$; $p < .05$] for the customer. The mean customer turn length is graphed in Fig 7.2.

Post hoc tests of simple main effects showed a significant main effect of system in the far conditions [$F(1, 54) = 6.58$, $p < 0.05$], but not for the close conditions. There was also a significant main effect of zoom when using the monitor [$F(1, 54) = 6.23$, $p < 0.05$] but not for the projector conditions.

These results indicate that when the participants were using the monitor, the mean turn length was longer when the advisor appeared far away – a direct replication of Study 2, but that this same finding does not occur for those participants who viewed the life-sized image. The analysis investigating the mean turn length for the financial advisor showed no significant effects.

These results are partially as hypothesised. Mean turn length was significantly longer in the Monitor Far condition than the Monitor Close condition, or the Life-size Far condition. The Monitor Far condition is the one which should have given the greatest impression of distance. There does not seem to have been a corresponding decrease in turn length for the condition in which the advisor appeared closest (Life-size Close) however.

7.3.1.7 *Acknowledgements*

Full Conversational Games Analysis was not conducted on the results of the present study due to the time consuming nature of the coding. Since the results of the CGA on Study 2 showed the level of acknowledgements to be the conversational move most affected by the differences in perceived distance, it was decided only to study the prevalence of acknowledgements in the customer dialogues.

A measure was made of the number of acknowledgements made by the customer in each dialogue. Only verbal acknowledgements were recorded, although it is noted that non-verbal communication methods can be used to acknowledge, previous studies have indicated very few non-verbal methods are used when communicating across video (Whittaker, 1995).

Although full Conversational Games Analysis coding was not to be done, it is possible to search just for particular types of move in the same way. Acknowledge moves are vocal acknowledgements of having heard and understood, and announce a readiness to hear the next move. They are characterised by often being short affirmative responses but which are not in response to a question. The following extract shows an example of a typical acknowledgement:

FA: The PEPs have an initial charge of 4% of the original investment

C: yeah (ACKNOWLEDGEMENT)

FA: and there is also a yearly charge after that

The number of acknowledgements made was recorded in the same way as for full Conversational Games Analysis with the exception being that only acknowledge moves were recorded.

The mean number of acknowledgements made by the customer in each condition are summarised in Table 7.9.

	Acknowledgements		Percentage of turns that are acknowledgements	
	Close	Far	Close	Far
Monitor	40.00	32.33	70.75	60.01
Life-size	35.67	32.00	61.02	64.81

Table 7.9 Mean number of acknowledgements and the mean percentage of turns which are acknowledgements (customer only).

Results of a 2-way ANOVA for independent samples indicated that there were no significant differences between conditions on either of the acknowledgement measures.

It had been hypothesised that there would be a greater percentage of acknowledgements in the conditions that the financial advisor appeared closer. The results show a trend towards a replication of the results of Study 2 which showed more acknowledgements in the monitor close condition than the monitor far condition. The same pattern is not evident in the results of the participants in the life-size conditions.

7.3.1.8 Summary of Communication Analysis Results

The results of the communication analysis for the present Study showed fewer significant differences in behaviour than were found in Study 2. Table 7.10 compares the results of the dialogue analysis in the present analysis with those of Study 2.

Measure	Results of Study 2	Results of Present Study
Time	Trend towards taking longer in close condition	Trend towards taking longer in the Life-size conditions
Turns	More turns in close condition	No significant difference between conditions
Backchannels	No significant difference between conditions	No significant difference between conditions
Overlaps	More overlaps in Close condition	Trend towards more Overlaps in Life-size Close condition. Trend to more Overlaps in Close zoom conditions compared to Far zoom conditions.
Rate of Overlaps	Higher rate of overlaps in close condition	Higher rate of overlaps in close conditions than far conditions
Words	More words in close condition	No significant difference between conditions
Share of Dialogue	Participants hold greater share of the dialogue in the Close condition	No significant difference between conditions
Words per Turn	Fewer words per turn in close condition	Fewer words per turn in Monitor Close and Life-size Far conditions than in Monitor Far condition
Acknowledgements	More acknowledgements in the Close condition	Trend towards more acknowledgements in the Monitor Close condition than the Monitor Far condition

Table 7.10 Summary and comparison of results of communication analysis between Study 2 and Study 4.

The results in Table 7.10 show that there are no wholly contradictory results between the two studies although there are some results which have failed to be replicated, most notably differences in the number of turns and words spoken.

7.3.2 The Questionnaire

The questionnaire which participants completed immediately following the task was in three parts - previous experience, opinions of the technology and the task, and social factors. All data were analysed in terms of the two independent variables - size (life-size or monitor) and zoom (close or far). All 61 participant questionnaires were analysed, no participants were considered outliers, as their answers would not skew the data as in the dialogue analysis. A copy of the questionnaire is contained in Appendix C.

7.3.2.1 Previous Experience

The questions about previous technical experience were to ensure that all conditions were balanced for experience and to provide a general demographic of the typical participant. Results showed that all participants reported to have experience with computers, only 6% of whom would still consider themselves inexperienced, none of the participants had ever banked by video conferencing before and only 15% had ever used video conferencing equipment in the past.

7.3.2.2 The Technology

Questions in this section asked participants about their impressions of the technology during the task.

7.3.2.2.1 Eye contact

In Study 2, participants were not asked about eye contact but despite this, 27% mentioned it as a problem in an “any other comments” question. In the present study, participants were asked to rate how good the eye contact was on a 5 point scale ranging from "Major problem" through to "very good". Results are summarised in Table 7.11.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Major problem	2	0	0	0
Minor problem	1	2	7	3
No problem	10	4	4	7
Good	2	4	3	4
Very good	0	5	1	1

Table 7.11 Frequencies of participant answers to the question "How good did you consider the eye contact to be?"

These results show that eye contact appears to be considered least good in the monitor far condition and best in the monitor close condition. In order to use non-parametric statistics on the data, it was necessary to merge some of the cells to ensure expected cell frequencies were at least 5. Data were merged so as to compare in terms of "problem", "no problem", and "good". A Kruskal Wallis test for significance indicated that this fell just short of significance [Chi Square (3) = 7.61, $p = 0.055$] but a Mann Whitney U test comparing only the two monitor conditions indicated that there was a significant difference between the two monitor conditions [$U = 56.0$, $p < 0.05$], indicating that more participants in the Monitor Close condition considered eye contact to be good than in the Monitor Far condition. The Mann Whitney U test also indicated that there was a significant difference between the monitor far and the life-size far conditions [$U = 68.5$, $p < 0.05$], indicating that more participants in the Life-size Far condition considered eye contact to be a problem than those in the Monitor Far conditions. There were no other significant differences between conditions in terms of eye contact.

These results show that participants were more likely to rate eye contact as being a problem in the Life-size Far condition than in the Monitor Far condition. Participants were also more likely to rate the eye contact as "good" in the Monitor Close condition than the Monitor Far condition. This shows that eye contact was considered to be more of a problem when participants were using the monitor to view the advisor appearing far away than when using either the life-sized image or zooming the image in so the agent appeared close. Using both the life-sized and the zoomed in image however, did not show a significant improvement in terms of eye contact to the monitor far condition.

7.3.2.2.2 Perceived proximity of advisor

Participants were asked about the image of the advisor that they saw and to say whether they considered that the advisor appeared too close or far away. Examination of the results indicated no participant considered the advisor appeared too close and 13 considered the advisor too far away. Of these thirteen, 7 used the monitor far condition, 5 used the life-size far condition and one the life-size close condition. These results are displayed in Table 7.12.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Too Far	7	0	5	1
Too Close	0	0	0	0

Table 7.12 Number of participants in each condition who considered the advisor to appear at an inappropriate distance.

It is obviously doubtful whether the individual really meant to answer that they appeared too far in the life-size close condition, but regardless of this, when the camera is zoomed out, approximately 33% of participants feel that the advisor appears too far away, a number which rises to over 45% when the size of the image is small. These results do suggest that zooming the camera had a greater effect on the impression of distance than image size (as the image in the Monitor Close condition was smaller than in the Life-size Far condition).

7.3.2.2.3 Size of the advisor image

In a related question to the previous one, participants were also asked whether they considered the image of the financial advisor to be either too large or too small. Only 2 considered the image too large - both of whom were in the life-size close condition (where the image was largest).

The frequency of those who considered that the advisor appeared too small can be seen in Table 7.13.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Too small	7	1	5	2

Table 7.13 Frequency of participants who responded "yes" to the question "did the advisor appear too close?"

These results show that, as with the previous question, camera zoom appeared more important than the equipment used to view the other individual. Even though the advisor is of equivalent size in the life-size far and monitor close conditions, they are considered too small more often in the life-size far condition.

7.3.2.2.4 Picture quality

Participants were asked various questions about the quality of the image that they saw. For each question they were merely to answer yes or no. Results of these questions are summarised in Table 7.14. In any one cell, there were 15 participants asked.

"Was the image?"	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Unclear	0	0	4	3
Slow/Jerky	0	1	0	1
Well Positioned	6	13	7	11
Badly Positioned	1	1	1	1

Table 7.14 Frequencies of participants answering "yes" to questions about the picture quality.

These results indicate that the picture was regarded as unclear by almost 25% of participants using the life-size conditions but none of those using the monitor. This can be considered a reflection of the lower resolution of the projector system.

The video was of "broadcast quality" at approximately 25 frames per second, but it is still interesting that still a small proportion of individuals considered it less than ideal - slow and jerky. Most participants appeared to consider the video window well positioned, although the number of participants rating the video window as well positioned was slightly higher in the close conditions.

7.3.2.2.5 Any technological difficulties?

Participants were given the chance to give details of any technological difficulties in completing the task. 13 participants answered this question, spread evenly across all conditions. Most complained that the volume was slightly low for their liking. Other comments were all expanding upon answers already given in previous questions. The audio volume was the same for all participants.

7.3.2.3 The Task

In this section of the questionnaire, participants were asked about their opinions of how difficult and realistic they found the task. They were also asked how willing they would be to use this type of technology to do their banking, and how it compared in terms of suitability with other technologies.

7.3.2.3.1 Realism

Participants were asked how realistic they considered the task to be. Since they were being asked to role play it was important that individuals considered the task realistic. The frequencies of responses in each condition are reported in Table 7.15.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Very realistic	1	4	3	2
Realistic	11	9	7	12
Neither	1	1	3	1
Unrealistic	2	1	2	0
Very unrealistic	0	0	0	0

Table 7.15 Frequencies of participant answers to the question "How realistic did you consider the task to be?"

These results show that individuals did consider the set-up realistic, only 8% reporting that they considered it unrealistic, and over 80% considering it realistic or very realistic. There were no differences between conditions.

7.3.2.3.2 Difficulty

Participants were asked how difficult they found the task to complete, participants answered on a 5 point Likert type scale from "very difficult" to "very easy". The frequencies of each answer are reported in Table 7.16.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Very difficult	0	0	0	1
Difficult	4	5	8	6
Neither	7	5	3	4
Easy	4	5	4	4
Very easy	0	0	0	0

Table 7.16 Frequencies of participant answers to the question "How difficult did you consider the task to be?"

There were no significant differences between conditions for the perceived level of difficulty. The results show that no participant considered the task very easy and only one found it very difficult. Marginally more participants found the task on the difficult side, but there was no difference between conditions. These results suggest that the task was probably challenging enough to keep individuals' attention without being particularly troubling, and that the different conditions did not cause participants any difficulty in themselves.

7.3.2.3.3 Willingness to use video-mediated technology in banking situations?

Participants were asked to state whether they would be willing to use this type of technology to perform various banking tasks. Table 7.17 shows the percentage of individuals in each condition indicating that they would be willing to use the video system to perform a particular task.

"Would you be willing to....?"	Monitor Far	Monitor Close	Life-size Far	Life-size Close	Total
Accept financial advice	80	100	80	67	82
Close an account	67	60	73	73	68
Open an account	73	87	93	87	85
Complain	47	27	40	53	42
Apply for a loan	67	60	60	53	60
Apply for a mortgage	60	47	40	40	47
Apply for an overdraft	73	73	60	67	68

Table 7.17 Percentage of participants answering "yes" to the question "Would you be willing to use the technology to...?"

These results indicate that opening an account and accepting financial advice are perceived as the most appropriate actions for using the technology. Complaining and applying for a mortgage are the least suitable for the technology. There appears to be no difference between the conditions on any of the questions.

7.3.2.3.4 Comparing with other banking methods

There were two questions which attempted to get an idea of how participants felt about the video-mediated technology in comparison to other methods of banking. The first question asked how far individuals would travel to get face-to-face advice as opposed to using the video-mediated kiosk. Results of this question are summarised in Table 7.18.

"How far would you travel to get face-to-face advice?"	Monitor Far	Monitor Close	Life-size Far	Life-size Close	Total (%)
Zero	1	0	2	2	8
0-15 minutes	5	3	5	7	33
15-30 minutes	4	9	4	3	33
30-60 minutes	5	3	4	3	25
Over an hour	0	0	0	0	0

Table 7.18 Frequency of participants answers to the question "How far would you be willing to travel to receive face-to-face advice....?"

The results in Table 7.18 indicate that most individuals would prefer to get face-to-face advice than use a video-mediated system, as only 8% say they would be unwilling to travel any distance to get face-to-face advice. One quarter of the respondents would be willing to travel for more than half an hour. There was again no significant difference between any of the conditions.

These results were confirmed by another question which asked participants to rank in order of preference 5 alternative methods of gaining financial advice. The alternatives offered were video, face-to-face, telephone, postal, and leaflets (no human contact). Results of this question showed that face-to-face was by far the most popular with 79% rating it as their first choice, next highest was video with 11% (and 54% rating it second). Video was consistently more popular than audio only conferencing (telephone).

7.3.2.4 Social Factors

The final part of the questionnaire investigated the effect of the different conditions on how participants felt about the advisor and the advice they were given. In order to investigate this 10-point Likert-type scales were used for the majority of questions. These were in place of the 5 point scales used in Study 2.

7.3.2.4.1 Persuasion

During the interaction, the advisor was attempting to subtly persuade the customer to select the "bonus account" option. There were 4 different options but some individuals chose to divide their money between two of the options so a fifth possibility was "combination". Table 7.19 summarises the number of individuals in each condition choosing each option.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close	Total
Bonus Account	60%	66.7%	62.5%	46.7%	59%
Shares	0%	0%	6.3%	6.7%	3.3%
PEP	26.7%	26.7%	12.5%	26.7%	23%
TESSA	13.3%	0%	6.3%	13.3%	8.2%
Combination	0%	6.7%	12.5%	6.7%	6.6%

Table 7.19 Percentages of participants choosing each option

Overall a chi-square test indicated that significantly more people chose the Bonus Account option than any other option [chi-square = 64.98, $p < .01$]. When asked why they had made their decision, 4 participants explicitly stated that it was on the recommendation of the financial advisor - all of these individuals had chosen the bonus account, corresponding to 11% of the total number of individuals who chose the bonus account. This result suggests that there was a significant effect of persuasion overall, however as there was no control group in which the advisor attempted no persuasion, this assumption is not proven.

The aim of the experiment however was to investigate whether or not participants differed in their response to persuasion under different conditions. To investigate whether there was a significant difference between conditions concerning the success of persuasion, a one-sample chi-square test was used. The results of this indicated that the distribution of those individuals choosing the bonus account was not significantly different between conditions [chi-square (3) = 1.99, $p > .05$] although there was a slight trend towards it being lower in the life-size close condition. The results therefore indicate that neither the level of camera zoom, nor the difference in the image size used had any effect on the persuasibility of the participants.

There were also questions which specifically asked individuals about their impression of the amount of persuasion used. These questions were both answered on a 10 point scale and were "Did you feel that each option was presented equally well by the advisor?" and "How much influence did you feel the financial advisor had on your final decision?". Results of each question are summarised in Table 7.20.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close	Total
Was each option presented equally well?	7.40	8.33	7.31	6.27	7.33
How much influence did the advisor have?	5.20	6.13	5.94	5.93	5.80

Table 7.20 Mean score on a rating scale of 1 to 10 a score of 10 corresponds to "strongly agree" or "a lot"

A Kruskal Wallis test for significant differences between conditions indicated that the question asking about equal presentation of options approached significance [Chi square (3) = 7.61, p =.055] and a Mann Whitney U test indicated that there was a significant difference between the two close conditions [U = 55.5, p<.05]. This shows that when the advisor appeared zoomed in close then if the monitor was used, the options appeared to be presented equally but if the life-sized image was used they appeared to be less equally presented. A similar analysis showed no effect of condition for the second question, indicating that neither image size nor zoom had any effect on the amount of influence the advisor was perceived as having.

These results suggest that the financial advisor was perceived as attempting less persuasion in the Monitor Close condition than the life-size close condition. This did not seem to lead to any difference in the amount of influence the financial advisor had on the customer's final decision in those conditions.

7.3.2.4.2 The Advisor

There were several questions which asked specifically about how individuals perceived the financial advisor. These questions were also answered on a 10-point scale. The results of these questions are shown in Table 7.21.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close	Total
1. How Useful?	8.27	8.47	7.31	7.47	7.87
2. Opportunity to ask questions?	8.07	7.87	7.31	7.53	7.69
3. How well answered questions?	8.50	8.40	8.19	7.27	8.08
4. Trust of financial advisor	7.40	8.00	7.00	8.27	7.67

Table 7.21 Mean score on a rating scale of 1 to 10 a score of 10 corresponds to "strongly agree" or "a lot"

1. "How useful do you feel the advisor was in helping you to make your choice?" (1=very useless, 10= very useful)

2. "Do you feel you had the opportunity to ask all the questions you would have liked to have asked?" (1= strongly disagree, 10 = strongly agree)

3. "How well do you think the advisor answered your questions?" (1 = not at all, 10 = very well)

4. "How much did you trust the financial advisor?" (1=distrust a lot, 10=trust a lot)

Kruskal Wallis tests of significance were carried out on the results comparing individual conditions and also comparing across conditions but between individual variables. No question approached significance. It must therefore be concluded that the different conditions had no effect on participants' opinions of the financial advisor.

7.3.2.4.3 The advice given

Table 7.22 summarises the results from the questions which specifically asked about the advice as opposed to the advisor.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close	Total
1. Understand info?	7.73	7.47	6.88	6.40	7.11
2. Trust advice?	7.40	8.60	8.56	7.20	7.95
3. Same choice without advice?	7.13	6.07	6.44	6.20	6.46

Table 7.22 Mean score on a rating scale of 1 to 10

1. "How well did you understand the information given to you by the financial advisor?" (1=not at all, 10= very well)

2. "How much did you trust the advice given to you by the financial advisor?" (1 = distrust a lot, 10 = trust a lot)

3. "How likely do you think it is that you would have made the same choice without the advice from the financial advisor?" (1= very unlikely, 10 = very likely)

As with the questions about the advisor, Kruskal Wallis tests of significance were carried out on the results comparing individual conditions and also comparing across conditions but between individual variables. No question approached significance. Again it must be concluded that the differences in conditions led to no effect on the participants opinions of the financial advice that they received.

7.3.2.4.4 Comparison with other mediums

Participants were finally asked how likely it was that they would have come to the same decision were they to have completed the task face-to-face or by telephone. This was again answered on a ten point scale. Table 7.23 summarises the results of these questions.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close	Total
Face-to-face	7.87	8.33	8.63	7.80	8.16
Telephone	6.87	6.40	7.00	6.07	6.58

Table 7.23 Mean score on a rating scale of 1 to 10 a score of 10 corresponds to "strongly agree" or "a lot"

The results show that individuals considered that they were less likely to choose the same option had advice been dispensed across the telephone as opposed to face-to-face. There were no differences between individual conditions.

7.3.2.5 Summary of Questionnaire Results

As in Study 2, the questionnaire has failed to find many statistically significant differences between the different conditions. The main findings of the questionnaire were that participants in the Monitor Close condition seemed to be the most impressed with the technical set-up, while the Life-size Far condition got the worst reviews. Participants in the Monitor Close condition also considered the financial advisor to be using less persuasion than in the Life-size Far condition.

Other than these results the questionnaire may be considered more useful as a tool to show participants' general impression of the task and technology. To this end, participants seemed reasonably impressed by the technology, found the task to be realistic and liked and trusted the financial advisor and the advice given.

7.3.3 The Interview

After having completed the questionnaire, participants were given a short surprise interview. This was designed to evaluate individuals' understanding of the options on offer during the task, without them specifically having been trying to remember them. The interview took the form of 23 questions about information which was available in the text of the whiteboard information screens and would generally be discussed with the advisor. A scoring system was used by giving 2 points for a correct answer, 1 point for a partially correct answer and 0 points for a totally incorrect answer. A copy of the interview can be found in appendix F.

7.3.3.1 Overall

Table 7.24 shows the mean scores for participants in each condition. The outliers which skewed the dialogue analysis were kept in the interview analysis because they will not have a skewing effect. The outlier due to age was removed from this analysis on the basis of the experiment seeking to keep a homogenous population.

	Close	Far
Monitor	25.13	28.13
Life-size	24.00	24.38

Table 7.24 Mean interview score (out of 46).

A two way (2x2) ANOVA for independent samples, using image size and camera zoom as independent variables indicated that there was no significant main effects and no significant interaction of image size and zoom on the overall score. This suggests that no particular condition would seem to lead to better understanding per se.

7.3.3.2 Individual Questions

As well as testing the overall effect of condition on understanding, results for each individual question were analysed using a 2-way (2x2) ANOVA for independent samples. This was to investigate if there were particular questions which were recalled better in any particular condition.

7.3.3.2.1 Camera zoom effects

One question showed an effect of camera zoom. The question "How do you make money with shares?" showed a significant main effect of camera zoom [$F(1, 57) = 6.58$, $p < 0.05$]. Examination of the means showed that the question was answered significantly better in the far conditions (mean = 2.0) than the close conditions (mean = 1.67).

It should be noted, however, that the probability of a Type 1 error when using a significance level of 0.05 (1 in 20) for 23 questions is quite high. It should also be noted that there was a ceiling effect in this question, with no-one getting it wrong in the far (zoomed out) conditions. For these reasons, there is little that can be read into this apparently significant effect.

7.3.3.2.2 Image size effects

The 2-way ANOVAs indicated that 5 questions out of the 23 led to significant main effects of system. For four of these questions, understanding was better when using the monitor and for one question understanding was better using the life-sized image.

The four questions where understanding was better with the monitor were:

"Which options are tax free?" [$F(1, 57) = 4.73$, $p < 0.05$]

"Are immediate withdrawals allowed with the bonus account?" [$F(1, 57) = 6.23$, $p < 0.05$]

"What is the maximum initial investment allowed with a TESSA?" [$F(1, 57) = 4.87$, $p < 0.05$]

"What is the minimum amount of money to keep in a PEP?" [$F(1, 57) = 4.78$, $p < 0.05$]

The question for which understanding was better when using the life-sized image was:

"Which options have charges?" [$F(1, 57) = 5.75$, $p < 0.05$]

As was noted with the camera zoom effects, the probability of a Type 1 error, rejecting the null hypothesis when it is true is high, and probably the most likely explanation for the single result indicating understanding improved in the life-size condition.

The four questions indicating understanding was better in the monitor conditions than the life-size conditions does suggest a slight advantage for the monitor conditions at certain times. The most likely explanation for that would be that the image of the whiteboard changed between systems, with a more blurred image in the life-size condition which may have either made certain parts less readable. The larger image on the projector screen may also have meant a decrease in the amount of information available in peripheral vision.

7.3.3.2.3 Interactions of image size and zoom

The ANOVAs indicated no significant interactions of image size and zoom for any individual questions.

7.3.3.3 Interactivity

The main reason for expecting an effect of understanding across this experiment was due to the theory of Clark and colleagues that interactivity aids understanding by virtue of grounding. To test this theory in this experiment, the data were divided by identifying the twenty most interactive dialogues and the twenty least interactive dialogues. Twenty was chosen so as to divide the data into three groups of equal size (very interactive, slightly interactive and non-interactive). Because the categorisation of participants into interactivity groups was fairly arbitrary, statistical analysis of the interactivity results only compared the interactive group with the non-interactive group.

The constitution of the interactive and non-interactive groups is shown in Table 7.25.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Very interactive	4	4	5	7
Non-interactive	4	7	5	4

Table 7.25 Number of individuals making up the different interactivity groups.

It can be seen that there are 50% more participants from the life-size conditions who produced highly interactive dialogues than the monitor conditions.

Interactivity is not a commonly used measure, and indeed could be described in different ways depending upon the nature of the dialogue. Some measures include time, total number of turns, number of customer turns, number of questions asked, share of the dialogue etc.

When listening to the interactions in the present experiment, it was clear that there were two main different types of strategy which participants could use to complete the task. The customer could either listen and allow the financial advisor to speak, generally only speaking when the financial advisor specifically asked them to (lecture style) or the customer continually gave verbal feedback and was verbally involved in the interaction (interactive style). Due to these differences in style, for the purposes of this analysis the best measure of interactivity was considered to be the number of customer turns in the dialogue - that is how often the customer spoke. Time was considered inappropriate due to different reading speeds of individuals, interactions may take a long time due to the amount being said or due to the customer doing a lot of reading. The range of the data of the number of customer turns spoken by participants in each interactivity condition is presented in Table 7.26.

	Minimum	Maximum	Mean
Very interactive	67	136	85.9
Non-interactive	10	39	27.6

Table 7.26 The range of the number of customer turns spoken in each interactivity condition.

The data presented in Table 7.26 shows that the participants in the very interactive group made over 3 times as many turns as those in the non-interactive group.

7.3.3.3.1 Overall

A t-test for independent samples compared the mean recall scores for each group of participants (interactive vs. non-interactive). This showed that scores were significantly higher in the interactive group (mean = 29.0) than in the non-interactive group (mean = 22.6) [$t(38) = 3.40$, $p < .01$]. This result shows that being interactive during the task led to greater understanding of the options that were discussed.

7.3.3.3.2 Individual questions

T-tests were also computed for each individual question, comparing the mean score of the interactive and non-interactive groups. Results of this analysis revealed 7 of the 23 questions were affected by interactivity, all of which had higher mean scores in the interactive group than the non-interactive group.

The seven questions which showed a significant difference were:

"Which options are risk free?" [$t(38) = 2.46$, $p < .05$]

"What were the options?" [$t(38) = 2.18$, $p < .05$]

"Who chooses where to invest your money in a PEP?" [$t(38) = 2.78$, $p < .01$]

"What are the charges on the PEP?" [$t(38) = 2.28$, $p < .05$]

"Which options involve the stockmarket?" [t(38) = 2.15, p < .05]

"What is the interest rate on the bonus account?" [t(38) = 3.42, p < .01]

"What does PEP stand for?" [t(38) = 2.43, p < .05]

The reason that all the questions did not show an effect of interactivity was mostly due to ceiling or floor effects. A question such as "how do shares work?" for instance was answered correctly by over 80% of participants. These results show that interactivity does lead to better understanding in general, and at no point did interactivity lead to a decrease in understanding.

7.3.3.4 Predictors of Interactivity

In order to find out what the major predictors of interactivity were, data from all areas of results were entered into a stepwise multiple regression which indicated that there were five major variables which predicted the number of customer turns (which was the chosen definer of interactivity), this regression gave an adjusted r squared value of 0.98. Those predictors are presented in Table 7.27.

Variable	t	p
Acknowledgements	11.18	p < .01
Customer words	10.62	p < .01
Customer words per turn	8.52	p < .01
Realism	3.15	p < .01
Interview score	2.25	p < .05

Table 7.27 - regression table for predictors of interactivity

The results of this regression analysis show that interactivity is highly correlated to the number of acknowledgements the customer made and the number of words spoken by the customer. There was also a negative correlation between the number of words per turn and the level of interactivity. The other factors which were related to the level of interactivity of the dialogues were realism - participants considered the task more realistic after they had been interactive – and interview score, interactive participants scored more highly on the interview recall test than non-interactive participants.

Customers who make more turns (speak more often) make more acknowledgements, say more words, say fewer words per turn, consider the task more realistic and show greater understanding of the content of the task.

7.3.4 The Video Tapes

During the interaction, there was a video camera set at 90° to the participant, which recorded each individual's body position throughout the trial. This recorded any movement from participants from when they were seated in front of the screen until when they completed the task.

When the tapes were analysed it became clear that individuals' all followed a similar pattern for their seating arrangement throughout the course of the experiment. Individuals were led into the room and sat on a pre-positioned chair. They then settled into a comfortable position. As soon as the experiment started and the advisor spoke they then would either remain in their position, move forward or move back (sometimes merely leaning forward or back, sometimes physically moving the chair) and settle into a new position. Once they had settled into this new position, they rarely moved during the course of the interaction. Because of this pattern of behaviour, individuals' were categorised as to whether they moved forward once they had initially sat down, stayed the same or moved back. Table 7.28 summarises the frequency of individuals behaviour in each category.

	Monitor Far	Monitor Close	Life-size Far	Life-size Close
Moved forward	7	7	7	4
Remained neutral	8	7	8	7
Moved back	0	0	0	2

Table 7.28 Frequencies of individuals body position during task.

The results show that of 57 participants analysed (3 were rejected due to technical problems). 30 remained neutral, 25 moved forward when they began interacting and 2 leaned back when interacting. The breakdown by condition showed only the Life-size Close condition to appear to differ at all from the others, a Kruskal-Wallis test however indicated that this did not reach significance – presumably due to the small numbers.

The body position data was then compared with the answers participants' gave in the questionnaires to investigate whether there was any link between participants' behaviour in altering their body position and the answers that they gave to questions asking how close or far the advisor appeared.

Of the two participants who moved away from the screen (both in the life-size close condition), neither reported that they felt the advisor appeared too close, although one did say they thought the advisor appeared too large. Of the 25 who moved closer to interact with the advisor, 7 reported that the image was too small but these were spread across all conditions (MC-1; MF-2; LC-1; LF-3).

In the questionnaire, 13 participants reported that they considered the advisor too far away. Of these, 4 had physically moved closer while 9 remained neutral during the trial.

7.4 Discussion

The primary aim of this experiment was to investigate how the size of image and amount of camera zoom may interact to create different impressions of distance, which in turn affect individuals' behaviour during a video-mediated interaction. It is known from previous studies in chapters 2 and 5 that both image size and camera zoom are positively correlated with feelings of closeness and that their effects on the impression of distance are cumulative. It was therefore expected that the largest differences in behaviour would be seen between the Life-size Close condition and the Monitor Far condition. It was also expected that by comparing the results of the Life-size Far condition and the Monitor Close condition, would show whether image size or camera zoom led to a greater feeling of perceived proximity.

7.4.1 Dialogue analysis

Previous experiments have shown that decreasing the perceived distance leads to longer interactions, with more but shorter turns (Capella, 1994; Short et al. 1979). Previously the financial advice task led to results showing that when the image of the financial advisor was zoomed in, the customer spoke more often, said more, and made more instances of overlapping speech than when the image appeared far away. These results have only been partially replicated in the present experiment: there were a greater rate of overlaps in the two conditions where the advisor appeared zoomed in, but there was no effect of condition on the number of turns or words spoken.

7.4.1.1 Length of dialogue

Length of dialogue can be considered in terms either of the actual time taken to complete the interaction or the total number of turns. The present experiment found no significant differences in the length of dialogue along either of these scales, although there was a trend towards interactions lasting longer in seconds in the life-size conditions than those using the monitor.

This trend is in the direction that would be expected, with the larger image leading to a feeling of greater proximity. The fact that there was no similar trend as far as the level of zoom on the camera is concerned suggests that on this variable, size of image may be more important than camera zoom. It is also possible however, that the differences in screen size may have led to an increase in the length of time taken to complete the task due to physical factors such as the length of time it takes to read a page of text on the projector screen, as opposed to the monitor screen.

The analysis showed no trend of an increase in the number of turns in any condition. For this result, differences in the image size cannot be implicated, due to the failure to replicate between the two monitor conditions. The only difference between the original financial advice study (Study 2) and the present study was the different financial advisor. Every confederate who plays the part of the financial advisor will have their own distinct style, regardless of how much instruction they are given, in the same way that participants differ in the amount that they speak. Whilst listening to the interactions, it became clear that the advisor in the present study compared to the previous one was more formal and less interactive. Although still allowing the customer to speak as much as they wished, the present advisor elicited less information from customers than the previous one. This is illustrated by a decrease in the overall mean number of turns from 120 in the earlier study to 93 in the present experiment, despite no accompanying decrease in the median length of time for an interaction. It is certainly possible that the more formal the expert is, the less the customers will behave in a familiar manner, regardless of visual cues.

7.4.1.2 Instances of Overlapping Speech

Results showed that for both technical systems, when the camera was zoomed in the customer overlapped more often. Overlapping speech was also most common in the life-size close condition. This replicates the results of the previous study which showed that zooming the camera in close led to the customer interrupting more. It also suggests that the monitor close condition led to a greater impression of perceived closeness than the life-size far condition.

Overlapping speech could be caused by either of two reasons. They may be a result of an individual failing to correctly judge when another was going to pause and beginning a speaking turn before the other person had finished. Alternatively, they may be caused by an individual deliberately wanting to stop the other speaker, possibly to query a point, or put their own point of view across. The former of these reasons can be viewed as an error, whilst the latter may indicate a lack of formality.

As speaking turns are monitored by very subtle communication cues, involving both verbal and non-verbal communication (Sacks, Schegloff, & Jefferson, 1974) it is reasonable to assume that differences in image quality will lead to differences in the ability of individuals to perceive the appropriate cues and consequently to accurately judge the precise timing of turn taking. Indeed, audio lag has been shown to have major effects on the turn taking processes (e.g. Tang & Isaacs, 1993).

In the different conditions in this experiment, although there was no audio lag, it is reasonable to suggest that the different sizes and resolutions of the advisor images may have lead to differences in the ability of an individual to judge when the advisor will pause. Close analysis of the results shows that in fact the overlaps were most likely not to have been caused by poorer image quality. The zoomed in conditions led to a greater level of overlaps – had it been a resolution issue then effects of image size would have been predicted. This suggests that the overlaps were caused by the customers being less formal and jumping in more, possibly suggesting a greater feeling of familiarity.

7.4.1.3 *Customer interactivity*

As explained earlier, interactivity is a broad term for a range of behavioural nuances. There can be many measures of interactivity in a dialogue, but on the whole. research has shown that when individuals appear closer then they are more interactive. This has been shown by increases in the length of dialogue, the number of interruptions and the amount any individual says during an interaction.

Due to the nature of the present experiment, interactivity was measured by the number of turns the customer made during the dialogue. Because there was information to read on the whiteboard, time may have been affected more by the speed and amount of reading an individual engaged in than a measure of how much people were talking. The overall number of turns in a dialogue was split fairly evenly between the financial advisor and the customer due to the task being two-person. However, it was possible for customers to make more turns than the financial advisor by backchannels, and interruptions which failed to break the flow of the financial advisor. Since the experiment was interested in customer interactivity, a purely customer based method of measuring interactivity was considered more appropriate. The number of words a customer spoke was also not considered to be the best measure of interactivity because sometimes an individual may say nothing for long periods and then have spoken for one turn using a large number of words, if they were giving a lot of information for example. Finally, the average number of words per turn was a possibility but again was rejected due to the fact that long customer turns either signified giving all the information at once (low interactivity) or asking a lot of questions (high interactivity).

Using customer turns was felt to be an appropriate method because it indicated how often someone spoke. Interactive dialogues tended to involve the customer speaking a lot of short turns - often simply an acknowledgement, they would ask questions and interrupt the financial advisor. Non-interactive dialogues were more "lecture-style" where the financial advisor would be allowed to talk for a long time and then the customer would talk before the financial advisor would talk for a long time again. This is shown by the fact that the dialogues ranged from only 10 customer turns in the least interactive to 136 in the most interactive.

Although it wasn't considered the best method of judging customer interactivity, a significant effect of customer turn length was found for the monitor conditions. This effect showed that when using the monitor, participant turn lengths were longer when the advisor appeared farther away. This does suggest that participants were less interactive in the monitor far condition than the monitor close condition.

7.4.1.3.1 Customer monitoring of the dialogue

The results showed that the interactivity of a dialogue in terms of customer turns was linked very strongly to the number of acknowledgements the customer made. There were far more acknowledgements in the interactive dialogues than in the non-interactive ones. This result shows that the customers in the more interactive dialogues were giving verbal feedback to the financial advisor and presumably indicating that they were monitoring the dialogue closely.

7.4.1.4 Non-verbal communication information

It was hypothesised that the approach-avoidance equilibrium proposed by Argyle & Dean (1968) may be relevant when considering participants reactions to perceived proximity. This model predicted that when individuals felt that their personal space was being invaded they would alter their body position in order to preserve a comfortable distance. Similarly, if the remote individual was considered to be too far away, participants may approach them to cut don the interpersonal distance by leaning or moving forwards.

The results showed that nearly half of the participants (47%) did indeed alter their body position once the advisor started speaking. It was particularly noticeable, however, that only 2 moved away, creating more space, both in the life-size close condition, predicted to give the closest view of the advisor. In all of the other three conditions, participants were split evenly between those that moved closer and those that remained neutral. This suggests that the feelings of distance were more powerful than impressions of being too close. This view is corroborated by the questionnaires, which indicated that no-one felt that the advisor appeared too close but several felt that the advisor appeared too far away.

One problem in interpreting these results is that the effect that the behaviour has on the feelings of the individual. It is reasonable to suppose that if an individual feels the advisor is too far away then the act of moving closer to the screen may alter that feeling. When asking that individual to subsequently report whether they considered the advisor to be too close or too far away it is unclear whether they will be reporting about their feelings before or after altering their body position.

The results of the present experiment allow some tentative conclusions to be reached. Of 25 participants who moved forward to talk to the advisor, only 7 (28%) considered the advisor too far away. Alternatively, 6 of 13 (46%) who considered the advisor too far away did not change their body positioning. This suggests that most people when they move to decrease the distance, report that the distance was roughly correct. When considered along with the finding that participants very quickly altered their position if they were going to, one is led to the conclusion that this behaviour is fairly automatic and occurs to increase comfort, which it does successfully. There are limits of course, and those 28% who moved closer and still considered the individual too far away, may simply have felt unable to get close enough.

The task also has a built in constraint on body position due to the existence of the whiteboard. In order to be able to read text, participants must be within certain limits of distance and must be facing straight ahead, changing bodily orientation may cause reading difficulties. This could well have constrained participants' movements and consequently moving to an appropriate distance may not have been an option for them. If this is indeed the case, then one would expect an equivalent study being completed without any textual information would have more movements of body position and fewer impressions of people feeling at the wrong distance, although this would not drop to zero.

7.4.2 Questionnaires

The questionnaires revealed few significant differences between conditions on questions investigating individuals' subjective impressions of the task, technology or the advisor. This was despite the questionnaire being extensively remodelled since Study 2 in an attempt to make it more sensitive. Some of the questions which notably failed to reveal any differences were the questions which asked participants how they felt about the advisor or the advice given.

It could be argued that the lack of significant differences on analysis of the questionnaire was due to the lack of homogeneity of the type of dialogue in terms of interactivity. This is refuted however, because the regression analysis revealed that there was no link between interactivity and the level of perceived persuasion or any other scale of the social factors questionnaire. Interactivity was only related to realism, with those participants who had been more interactive, rating the task as being more realistic.

One area where the questionnaire did show a significant result was the issue of eye contact. Participants in the Monitor Close condition tended to rate eye contact as good whereas those in the Life-size far condition rated it as a problem. Considering that the size of image in these conditions was comparable little difference had been expected. The fact that the monitor was preferred is most likely to be a reflection of the poorer resolution of the projector screen, and also showed that zooming in the camera led to a greater feeling of good eye contact.

7.4.2.1 Customer opinions

When asked about the image of the advisor, 42% of participants in the far condition considered the advisor appeared too far away, none in the close condition considered the advisor to appear too close. This is in contrast to the previous experiment, in which the advisor was more commonly considered too close than too far away. Why this should be the case is unclear, considering that for at least the monitor conditions the only difference was the actual advisor. It is possible that somehow the advisor appeared farther away in general perhaps by some difference in their demeanour, speculation that would also explain the lower number of turns in general compared to the previous proxemics experiment.

Overall the questionnaire suggested that individuals' opinions of the technology used and how much they liked it was not particularly affected by either the perceived proximity of the advisor or the technological system used. Most people considered the technology used to be of a high quality and well laid out, with no differences between conditions on any questions investigating their likelihood to use the equipment in real life.

As in Study 2, the questionnaire in the present study has proven to be most useful for gaining an insight into participants' general impressions of the task. As a tool for investigating psychological effects caused by fairly subtle differences in a technical set-up, the questionnaire has again proven to be rather ineffective.

7.4.2.2 Understanding

Although the experimental manipulations did not lead to the same effects on participant interactivity as in the previous study, the interview recall task still provided useful results. The results of comparing the scores of the interview with the level of interactivity displayed in the dialogues clearly demonstrated a link between interactivity and understanding. Specifically, the results show that an increase in the level of active participation in a dialogue leads to an increase in the ability of an individual to absorb facts and understand general concepts being explained by another individual.

This finding takes further the work of Schober & Clark (1989) who showed that understanding was greatly improved when an individual took part in a conversation rather than simply overheard it. It is most likely that the cause of the extra interactivity was from closer monitoring of the dialogue by the customer. It was hoped to show a clear link between perceived closeness leading to greater understanding by virtue of greater interactivity. The failure of the present study to replicate the increase in interactivity caused by perceived proximity in the original study meant that this could not be done. Although concrete evidence of a link between perceived closeness and greater understanding has not been provided, it does appear to be fair to say that if understanding is to be affected by perceived distance, it is likely to be in the direction that closeness leads to better understanding.

There are clear implications from this research for any situation which involves attempting to transfer information from one individual to another such as education and helplines. It is important to note that although there is likely to be a correlation between time and interactivity in many situations, the present research has demonstrated that time was not the key factor in determining the level of understanding. Time was not a significant predictor of interactivity and thus these results cannot be accounted for by a direct relationship between time and understanding.

7.4.3 Confederate Variables

Part of this experiment involved a direct replication of the previous financial advice experiment. In doing so, there was only one difference – the individual who played the role of the financial advisor. As has already been alluded to, it is possible that differences in results may be due specifically to this change.

Just as there are subtle differences which affect impressions of perceived distance, so there are very subtle behaviours which regulate communication, and at what distance individuals should communicate. One behaviour which has been shown to affect communication in a similar way to changes in perceived distance is gaze, Hughes & Goldman (1978) showed that people are less likely to stand close to a confederate who gazes steadily at them. It is possible that if the advisor gazes at the camera a lot more, then there is an equivalent effect to the face-to-face situation and individuals will communicate as if they were farther apart.

The only information recorded on how the two advisors differed between experiments was that of dialogue. As was stated in the results section, the advisor in the present study was a lot more formal in her approach to dialogue than the advisor in the previous study. Because of the sensitive nature of dialogue this is almost certain to have had some effect on the manner in which the participants spoke, the results suggests that participants were also more formal in the present experiment than the original study.

There was no information on gaze recorded but there is also reason to predict that there were differences in gaze between the two confederates. In the original study, the confederate was a psychology graduate with experience of experimental situations but not specifically of video conferencing (although she was given some training beforehand). In the present study, the confederate was a professional actress who is trained to look into a camera. It is highly probable that the amount she looked into the camera differed and hence the impression of gaze differed. As the data for this behaviour is unavailable then this must remain only speculation, no conclusions can be drawn as to what may or may not have caused behavioural differences in the customers. It has been shown however, that changing the advisor does affect the behaviour of the customer.

7.4.4 So Which Variables Impact Upon Perceived Distance?

The main aim of this experiment was to investigate whether image size or camera zoom had a greater influence upon perceived distance and consequently influenced behaviour. The results suggest that in general, camera zoom has a greater potential to affect behaviour than image size but that both together can enhance any effect.

This conclusion is based on the results which showed that people considered the advisor appeared too far away in the life-size far condition but not the monitor close condition, when there was an equivalent image size but a difference in camera zoom. Level of camera zoom was also the critical factor in determining the number and rate of customer interruptions.

In no area of the task, with the exception of customer impressions of the technology, was image size the critical factor. Had it been, it would have been expected that there would be large differences between the life-size close and the monitor close conditions and similarly with the life-size far and monitor far conditions.

Some results did indicate an interaction between the variables. The mean customer turn length was longer in the monitor far condition than the monitor close condition but had no effect between the life-size conditions. Similarly for the percentage of turns that were acknowledgements, there were more acknowledgements in the monitor close condition than the monitor far condition. These results indicate that changing the level of camera zoom is more important than the image size but that image size can alter the effectiveness of changing the level of camera zoom.

A similar example is that of individuals' impressions of the quality of the eye contact. Eye contact was rated as poor in the monitor far condition but could be improved by either changing the zoom or the image size. Changing both however, offered no improvement, suggesting that there may be limits to the benefits, causing the advisor to appear too close leads to eye contact becoming a problem again.

7.5 Conclusions

Both image size and camera zoom contribute to an impression of perceived distance, which in turn affects the way in which individuals interact with a stranger. When the stranger appears close participants make more interruptions and more acknowledgements than when they appear far away. Increasing the image size to life size appears to have little effect on the way individuals interact or complete a task. It has also been shown that understanding is aided by a more interactive dialogue and that this may tend to occur more frequently in conditions when the advisor appears closer.

8 Chapter 8 . The Functional Effect of Familiarity and Eye Contact in Task Oriented Communication

8.1 Introduction

Although it has been stated that proxemic information is one of the most fundamental pieces of non-verbal communication, it must also be noted that proximity itself can be influenced by other areas of non-verbal communication such as culture and personal relationships. The present research is an investigation of the latter of these, and seeks to understand the effect of familiarity between speakers on conversational behaviour in a task-oriented situation.

One of the hypotheses raised by the proximity research previously reported in Chapter 4 is that a forced close proximity may lead to a feeling of perceived familiarity. Effectively the hypothesis is that while familiarity may affect proximity, so proximity may unconsciously affect impressions of familiarity. Some of the behaviours that are thus performed between familiar individuals may also be seen when individuals are interacting at a close proximity (or close virtual proximity).

Familiarity is known to aid interpretation of an individual's communication. Studies have shown that pairs of friends are better than strangers at inferring each others thoughts and feelings during interactions (Stinson & Ickes, 1992) and are also able to purposely communicate hidden messages to each other, indiscernible to strangers (Fleming, Darley, Hilton, & Kojetin, 1990). Common sense would suggest that this improved communication and increased level of understanding between friends would be a result of more contact between friends, and thus an increased experience of any individual's particular nuances.

An alternative to this "experiential model" is proposed by Ansfield, DePaulo & Bell (1995) who point to literature which show that pairs of friends, siblings or partners tended to be more similar in their perceptions of stranger's expressions than pairs of strangers (Brauer & DePaulo, 1980; Blanck, Zuckerman, DePaulo & Rosenthal, 1980; Zuckerman & Przewuzman, 1979). Ansfield et al. completed a study which tested individuals' skill at judging the expressions of themselves, their friends and strangers. The results showed no difference between the ability of individuals to judge their friend's expressions over their own or strangers, but that there was a striking similarity between female friends' interpretation of expressions of both familiar and unfamiliar faces. These results suggest that friends (especially female friends) are more similar in their understanding of non-verbal information.

Millar & Millar (1995) investigated the ability of individuals to detect deception being practised by friends or strangers when manipulating the availability of sensory cues. Participants described a photograph which was either pleasant or unpleasant. In half of the trials they had to lie and in half they had to tell the truth. All trials were completed face-to-face, but the listener either wore a blindfold (audio only) or could see as well (audio-visual). Results showed that when listeners had full visual and aural information, they were better at detecting the deceptions of strangers than friends but when only aural information was present, participants were more successful at detecting the deceptions of their friends than strangers. There was also a bias towards considering friends as more truthful than strangers, a finding that has been found several times in the past (e.g. McCornack & Levine, 1990; McCornack & Parks, 1986).

It is still too early to discriminate between the two models of why familiar individuals interpret friends' communication better than strangers, although it appears that feelings of trust appear to interfere somewhat with the ability to intercept hidden messages. One of the aims of the present study is to investigate how familiar pairs of individuals communicate and why this may lead to better task performance. This investigation may help to add to the experiential model debate.

So far all these studies have investigated the effects of familiarity in a face-to-face situation, none has been done investigating the effects of familiarity in a video-mediated situation. The finding that deception is easier to detect between familiar people in an audio–only environment, compared with audio-visual being better for strangers, has clear implications for certain types of tele-sales – depending upon whether or not deception or persuasion may be required.

The earlier studies reported in this thesis show structural similarities between dialogues at close proximity as shown in study 2 and those of familiar pairs as reported by Boyle et al. (1994) in a task oriented situation. These similarities are that familiarity and close proximity both lead to increases in dialogue length in terms of the number of conversational turns, the number of words spoken, and led to shorter turns of speech.

The reason that there is a possibility of a link between the effects of proximity and familiarity is that it is known that the relationship between two individuals affects the distance at which they interact. It is possible that there is a two-way psychological effect on the individuals involved. If strangers are forced to interact at a distance they would normally reserve for a friend, then they may behave in a manner similar to that normally reserved for a friend. To investigate this hypothesis further, it is necessary to go beyond the structural aspects of the dialogue and compare the effects on dialogue content of close proximity and familiarity.

8.1.1 Conversational Games Analysis

In Chapter 5 , Conversational Games Analysis was applied to the dialogues of the original Financial Advice task experiment. CGA has in the past been used to code the HCRC Map task corpus (Anderson et al., 1991). A detailed description of Conversational Games Analysis and examples of types of move found in both the Map task and the Financial Advice task are presented in Chapter 5 . As a reminder however, the moves which are coded in CGA are as follows:

There are 12 different moves, which are divided into two categories - initiating moves and response moves. Initiating moves begin a game and response moves occur during a game. A brief description of the communicative function of each move is given below.

8.1.1.1 Initiating moves

INSTRUCT. A direct or indirect instruction to be done immediately or shortly.

CHECK. Checks self-understanding of a previous message or instruction by requesting confirmation, either directly or indirectly.

QUERY-YN. Yes-No question asking for new or unknown detail about task (not clarifying instructions)

QUERY-W. As Query-YN but an open ended question.

EXPLAIN. Describes status quo or position in task, with respect to goal. Freely offered, not elicited.

ALIGN. Check the other persons understanding or accomplishment of a goal; checks alignment of both participants plans or position in task with respect to goal; checks attention, agreement or readiness.

8.1.1.2 Response Moves

CLARIFY. Clarifies or rephrases what has previously been said; usually repeats given or known information.

REPLY-Y / REPLY-N. Affirmative (REPLY-Y) or negative (REPLY-N) response to a Query-YN, CHECK or ALIGN move. Elicited by other person.

REPLY-W. An elicited response that is not categorisable as either positive or negative.

ACKNOWLEDGE. Vocal acknowledgement of having heard and understood. Announces readiness to hear next move.

READY. Indicates intention to begin new game and focuses attention on oneself, in preparation for the new move.

8.1.2 Familiarity, Eye Contact and the HCRC Map Task Corpus

An experiment by Boyle, Anderson and Newlands (1994) examined the effects of familiarity and visual contact on conversation in a task-oriented setting. By analysing the structure of dialogues as well as task performance on the Map Task (Brown et al. 1994), they showed clear differences in conversational structure between familiar and unfamiliar pairs of individuals and between pairs with or without eye contact. The dialogues used in this experiment were drawn from the HCRC Map Task Corpus (Anderson et al, 1991).

Boyle et al (1994) showed that those participants who could not see each other, indulged in longer dialogues in terms of the number of turns and words spoken.

Although these turns were shorter in terms of the number of words per turn spoken. Participants deprived of eye contact also interrupted each other more, more often and had more incidences of overlapping speech. There was also a greater rate of backchannels, indicating more verbal feedback in the absence of visual communication cues.

A similar analysis showed that familiar pairs of participants tended to perform the task slightly better, although they used more conversational turns and more words to do so. They also found that familiar pairs interrupted each other less and had fewer instances of overlapping speech. When controlling for the length of dialogue in turns they found that there was no difference in the percentage of turns which contained an interruption between familiar and unfamiliar pairs of speakers. This is a surprising finding given that the results show that familiar pairs make more turns per dialogue, and also make fewer interruptions per dialogue than unfamiliar pairs, yet the number of turns which have interruptions (rate of interrupting) shows no difference. One possible explanation for the pattern of the published results is that when an individual attempts to interrupt an unfamiliar interlocutor, they are either less successful or more persistent at attempting to gain the floor within a single turn, and so interrupt more often. Consequently, unfamiliar individuals may be less willing to yield the floor to interruptions, instead finishing what they say.

The logic of this argument is as follows, in Boyle et al's data familiarity leads to more turns, but fewer interruptions. Yet the number of turns which contain an interruption shows no difference. Table 8.1 shows the interruptions data as reported by Boyle et al along with the estimated rate of interruptions derived from the reported data. The estimated rate of interrupting is taken by assuming no speaking turns could contain more than one interruption (although that is clearly not the case). This also means that the estimated rate of interruptions is the theoretical maximum rate of interruptions for that condition - based upon the numbers of interruptions, thus it is impossible for the actual rate of interruptions to be higher than the estimated rate of interruptions, it must be lower or equal.

	Familiar	Unfamiliar	Eye Contact	No Eye Contact
Mean turns	179.7	145.7	142.5	182.9
Mean Interruptions	16.38	21.32	12.6	24.8
Reported rate of interruptions	not reported	not reported	8.7%	12.07%
Difference	not reported (but not significant)		3.3% (significant, p<.01)	
Estimated rate of interruptions [(interruptions/turns)*100]	9.12%	14.63%	8.84%	13.56%
Estimated Difference	5.51%		4.72%	

Table 8.1 Rate of interruptions in dialogues (results reported by Boyle, Anderson & Newlands, 1994).

The data show that the estimated rate of interruptions for eye contact came out slightly higher than the reported rate, as would be expected due to some turns containing more than one interruption. However, the difference of 3.3% was significant. The estimated difference between interruption rates for the familiarity conditions is higher than the equivalent rate for eye contact, so it would be expected that the actual difference was also higher than the equivalent rate of eye contact. The results however showed that this is not the case. Therefore we might conclude that there is little difference between the familiar and unfamiliar interruption rates. As the estimated familiar interruption rate is the theoretical maximum interruption rate (given the total number of interruptions recorded), then the actual rate of interruptions in the unfamiliar condition is much lower than the estimated interruption rate. This means that there must be many more instances of several interruptions taking place during a single turn in the unfamiliar condition.

Boyle et al. (1994) make no reference to this finding being odd, but conclude that familiar pairs are better at co-ordinating their turn taking than those who do not. If in fact the reason that more interruptions are occurring is that interruptions are being less successful then the co-ordination problems which unfamiliar people have in dialogue are based not on smooth turn-taking but on taking the floor at any time other than when their partner wishes them to. When an individual is speaking during the task, it is as likely for them to be interrupted by an unfamiliar person as a familiar person, when that interruption happens, however, the unfamiliar person is more likely to make more interruptions during the same speaker turn.

The present analysis seeks to take the results of Boyle et al. (1994) further by using CGA to investigate the differences in content between familiar and unfamiliar dialogues. Doherty-Sneddon et al. (1997) have already analysed the effect of eye contact in these dialogues but it was included in this analysis in order to remain true to the original experimental conditions and investigate any interactions between familiarity and eye contact. By doing these analyses it is hoped to gain a greater understanding of how familiar speakers communicate, and to detail how pairs of participants may benefit from their longer dialogues by producing better task outcomes. A detailed examination of coded dialogues should go some way to illuminating the following questions: What are familiar speakers attempting to achieve in their extra talk? What kinds of communicative functions occur more frequently in dialogues between familiar speakers and how might this produce more effective task outcomes?

Familiarity has also been shown to affect visual communication between individuals. In general it has been shown that familiar persons gaze more at one another (Rutter et al., 1978). Rutter & Stephenson (1979) however, showed that in a task based situation strangers gazed at each other more, with gaze presumably performing a communicative function. Boyle et al. (1994) showed that in completing the Map task, information followers gazed at familiar partners more than unfamiliar partners, and that this was most prevalent in times of communicative difficulty.

In the present investigation 32 of the 64 dialogues with eye contact possible were analysed for gaze distribution. The aim of the gaze analysis was to investigate the effect of familiarity on gaze patterns in the Map task. Half of the dialogues were between familiar pairs and half involved unfamiliar pairs.

The present study seeks to investigate the relationship between familiarity and visual contact. When visual cues are missing, there will almost certainly be a difference in the pattern of conversational content. This study should indicate whether or not there is a difference in the way that friends compensate for a lack of visual information compared with the way that strangers do. It is hypothesised that the deprivation of visual contact will affect the conversational patterns of friends in a different way to those of strangers.

This study will also examine the inherent differences in conversational content between the two roles in the Map Task. It is expected that the two roles each have a very different pattern of conversational moves as each role has very different goals. It should be remembered that the information giver is attempting to transfer knowledge to the information follower, while the information follower is attempting to receive knowledge from the information giver. The analysis will allow a distinction to be made between differences in conversational style caused purely by the task, and those caused by the independent variables (and how these role differences relate to familiarity).

The analyses of the effects of real familiarity on task oriented dialogues will also be compared with the effects of apparent proximity reported in study 1. It is known that in terms of dialogue length, familiarity and perceived closeness have similar effects, and it has been suggested that this may be due to apparent closeness leading to an effect of an interlocutor being treated as if they were familiar. Comparing the results of the CGA across the two studies should show whether familiarity and perceived closeness both lead to similar changes in dialogue content. It should be noted however, that differences in task inevitably lead to differences in the patterns of dialogue function. To study this, a comparison will be made between the roles of the Map task and the role of the customer in the financial advice task. It is hypothesised that there will be a similar conversational distribution between the customer in the close proximity financial advice task and the information follower in the Map task.

8.2 Method

The experiment involved analysing 128 pairs of dialogues from the HCRC Map Task Corpus (Anderson, Bader, Bard, Boyle, Doherty, Garrod, Isard, Kowtko, McAllister, Miller, Sotillo, Thompson & Weinert, 1991). The dialogues were taken from a total of 64 participants (32 male, 32 female; mean age = 20), each of whom participated four times. Dialogues were transcribed and coded using conversational games analysis and gaze coding by researchers at the HCRC. The present study involved analysing the relative frequency of each type of conversational move.

8.2.1 The Task

The Map Task (Brown, Anderson, Shillcock and Yule, 1984) is a cooperative task involving two participants. The two speakers sit opposite one another and each has a map which the other cannot see. One speaker (the Instruction Giver) has a route marked on her map, the other speaker (the Instruction Follower) has no route. The speakers are told that their goal is to reproduce the Instruction Giver's route on the Instruction Follower's map. The maps are not identical and the speakers are told this explicitly at the beginning of the session.

8.2.2 Conditions

The dialogues studied differed according to two independent variables, familiarity and visibility. Participants in the familiar condition were pairs of friends, those in the unfamiliar condition were strangers. Eye contact was manipulated by the presence or absence of an opaque screen between the participants, when the screen was present, participants could still hear but could not see their partner. The other independent variable for each participant was whether they were the information giver or the information follower.

Every participant attempted the task 4 times, twice with a friend (once as information giver and once as information follower) and twice with a stranger (again, once each role). Each participant participated either in the eye contact or the no eye contact condition. Order was counterbalanced across all conditions.

8.2.3 The Corpus

The HCRC Map Task corpus (Anderson et al. 1991) allows online searching of 128 experimental dialogues. These dialogues are fully transcribed and include full data on disfluencies, overlaps, time stamping and conversational games analysis. For the purposes of this study, dialogues were analysed by conversational move as it was differences in dialogue content that were of interest.

8.3 Results

The number of moves in each dialogue was recorded and compared for each of the three independent variables – familiarity, visibility and role. Analyses were conducted on the total number of each type of move and the total number of moves. An analysis was also done comparing each move as a percentage of the total moves in the dialogue, which has the effect of standardising each dialogue for length. This was done by taking the total number of any particular type of move as spoken by one or both interlocutors (depending upon whether the analysis was across or between roles), dividing by the total number of moves spoken in each dialogue (again by one or both interlocutors) and multiplying by 100 to get the percentage. The advantage of completing both types of analysis is that it can indicate the difference between those moves affected by the variable and those which are merely influenced by the length of the dialogue.

Each analysis involved a 3-way (2x2x2) ANOVA for independent samples. The independent variables were familiarity (familiar or unfamiliar), visibility (eye contact or no eye contact) and role (information giver or information follower). The dependent variable was the number of moves of a particular type spoken by an individual in a particular dialogue. There were 128 dialogues analysed, corresponding to 256 individual contributions across the three independent variables.

The effects of visibility were analysed by Doherty-Sneddon et al. (1997). Although in the present study the effect of visibility (eye contact) was again analysed, the results were the same and consequently shall not be reported in this study. The purpose of analysing the results again was to investigate any interactions between familiarity and visibility. The results found by Doherty-Sneddon et al. showed that in conditions where visibility was present there was an overall increase in the total number of moves made which was accounted for by increases in the number of REPLY-Y, ACKNOWLEDGE and READY moves made. For a full discussion of the effects of visibility, interested readers are referred to the original paper.

8.3.1 Total moves

Analysis of the total number of moves showed main effects of familiarity [$F(1, 62) = 23.61$, $p < 0.01$]; and role [$F(1, 62) = 18.32$, $p < 0.01$]. There were no significant interactions. The mean number of moves made in each category is presented in Table 8.2.

Condition	Info giver	Info follower	Mean
Familiar	130.39	104.02	117.21
Unfamiliar	105.94	83.03	94.48
Mean	118.16	93.52	

Table 8.2 Total mean number of moves by an individual in each condition.

These results indicate that individuals make more conversational moves in general when talking to someone they know, regardless of whether they are giving or receiving information. The results also show that in the task the information giver makes more conversational moves than the information follower.

Having shown that dialogues contain more conversational moves in some conditions, each type of move was analysed individually to investigate what types of move the extra conversation was in each condition. Because of the differences in the overall number of moves in the dialogue it is useful to study the frequency of each type of move as well as the number of moves per se. Each type of move was analysed separately revealing the following results.

8.3.2 Instruct

A 3-way (2x2x2) ANOVA showed significant main effects of familiarity [$F(1, 62) = 7.14$, $p < .05$] and role [$F(1, 62) = 729.66$, $p < .01$]. There was also a significant interaction of familiarity and role [$F(1, 62) = 5.56$, $p < .05$]. The results of the interaction are presented in Table 8.3.

	Information Giver	Information Follower
Familiar	34.83	0.34
Unfamiliar	30.89	0.17

Table 8.3 Mean No. of INSTRUCT moves for familiarity and role.

Table 8.3 shows that when pairs of individuals are familiar, the Information Giver makes more INSTRUCT moves than when unfamiliar. This effect does not occur for the information follower because the number of INSTRUCT moves spoken by the information follower is so low as to constitute a floor effect.

Table 8.4 summarises the significant main effects as shown by the ANOVA.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	17.16	7.14	< .05
	Unfamiliar	15.53		
Role	Info giver	32.86	729.66	< .01
	Info follower	0.26		

Table 8.4 Mean number of INSTRUCT moves for an individual in each condition.

The results show that the information giver makes more INSTRUCT moves than the information follower. Familiarity also leads to an increase in the number of INSTRUCT moves made.

When the data for INSTRUCT moves were standardised for the total number of moves spoken in a dialogue, a 3-way ANOVA revealed significant main effects of familiarity [$F(1, 62) = 4.79$, $p < .05$], role [$F(1, 62) = 1170.23$, $p < .01$] and a significant 2-way interaction of role and familiarity [$F(1, 62) = 5.33$, $p < .05$]. No other comparisons reached significance. Table 8.5 shows the mean percentage of INSTRUCT moves between familiarity and role conditions, across eye contact.

	Information Giver	Information Follower
Familiar	28.64	0.32
Unfamiliar	31.98	0.19

Table 8.5 Mean % of INSTRUCT moves for familiarity and role.

A post hoc Tukey test showed that there was a significant difference between the amount of INSTRUCT moves said by the information giver in the unfamiliar condition and the familiar condition. The information giver gave significantly more instructions when they were talking to a stranger than a familiar person, this accounted almost entirely for the significant main effect of familiarity. It is also clear that the information giver made many more INSTRUCT moves than the information follower.

Taking both analyses together it can be seen that the information follower never makes more than an occasional INSTRUCT move, and that this is unaffected by familiarity. The information giver on the other hand gives a lot of instructions moves and although they make more in the longer familiar dialogues, this actually corresponds to an overall decrease in the rate of instructing in the familiar condition. This suggests that little of the extra speech in the familiar dialogues can be attributed to INSTRUCT moves.

8.3.3 Check

A 3-way (2x2x2) ANOVA showed significant main effects of familiarity [F(1, 62) = 8.77, p < .01], and role [F(1, 62) = 131.70, p < .01]. There were no significant interactions. Table 8.6 shows the mean number of CHECK moves in each condition and summarises the results of the ANOVA.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	9.89	8.77	< .01
	Unfamiliar	7.25		
Role	Info giver	2.24	131.70	< .01
	Info follower	14.91		

Table 8.6 Mean number of CHECK moves for an individual in each condition.

The results show that the information follower makes more CHECK moves than the information giver. Familiarity also leads to more CHECK moves being made.

When the data was standardised for the total number of moves spoken in a dialogue, a 3-way ANOVA revealed a significant main effect of role [F(1, 62) = 448.31, p < .01], and a significant main effect of familiarity [F(1, 248) = 6.02, p < .05]. There were no significant interactions. The results are summarised in Table 8.7.

Variable	Condition	Mean % of Moves	F	p
Familiarity	Familiar	8.96	6.02	< .05
	Unfamiliar	7.79		
Role	Info giver	1.61	448.31	< .01
	Info follower	15.14		

Table 8.7 Mean percentage of CHECK moves for an individual in each condition.

The results indicate that the information follower more frequently checks their understanding of a previous utterance than the information giver. When speakers are familiar, this checking occurs more frequently.

8.3.4 Query-YN

Analysis of the mean number of QUERY-YN moves spoken in a dialogue showed a significant main effect of role [$F(1, 62) = 76.37$, $p < .01$]. Analysis of the means shows that the information giver made more QUERY-YN moves (mean = 9.05) than the information follower (mean = 4.60). There was also a significant interaction between familiarity and visibility [$F(1, 62) = 76.37$, $p < .01$]. Table 8 summarises the means for the interaction between visibility and familiarity.

	Visiblility	No visibility
Familiar	7.61	6.72
Unfamiliar	5.84	7.14

Table 8.8 Mean No. of QUERY-YN moves for familiarity and visibility.

A post hoc Tukey test showed that there were significantly fewer QUERY-YN moves spoken by unfamiliar pairs than familiar pairs when they could see each other. This result suggests that visibility leads to a decrease in an individual's propensity to ask questions of the other, familiarity however, overcomes this. People may well feel less comfortable to ask a stranger a question if they cannot see them.

When the data was standardised for the total number of moves spoken in a dialogue, a 3-way ANOVA again revealed a significant main effect of role [$F(1, 62) = 47.57$, p

< .01]. This showed that a greater percentage of the information givers dialogues consisted of QUERY-YN moves (mean = 8.28%) than the information follower (mean = 4.70%).

The results indicate that the information giver makes more QUERY-YN moves than the information follower. Individuals may also feel less inclined to ask a stranger a question if they cannot see them, than if they can or if their partner is a friend.

8.3.5 Query-W

A 3-way (2x2x2) ANOVA showed significant main effects of familiarity [$F(1, 62) = 25.18$, $p < .01$], and role [$F(1, 62) = 46.08$, $p < .01$]. There were no significant interactions. Table 8.9 shows the mean number of QUERY-W moves in each condition and summarises the results of the ANOVA.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	3.87	25.18	< .01
	Unfamiliar	2.16		
Role	Info giver	1.70	46.08	< .01
	Info follower	4.34		

Table 8.9 Mean number of QUERY-W moves for an individual in each condition.

The results show that the information follower makes more QUERY-W moves than the information giver. Familiar individuals make more QUERY-W moves than unfamiliar individuals.

Upon standardisation of the data, a 3-way ANOVA revealed significant main effects of role [$F(1, 62) = 105.11$, $p < .01$], and of familiarity [$F(1, 62) = 9.99$, $p < .01$]. There was no significant interactions. The mean percentages of QUERY-W moves for each role and familiarity condition are tabulated in Table 8.10.

Variable	Condition	Mean % of Moves	F	p
Familiarity	Familiar	3.42	9.99	< .01
	Unfamiliar	2.40		
Role	Info giver	1.27	105.11	< .01
	Info follower	4.57		

Table 8.10 Mean percentage of QUERY-W moves for an individual in each condition.

These results show that the information follower made more QUERY-W moves than the information giver. Participants also asked their listener more questions when talking to a familiar partner than an unfamiliar one. Both of these findings applied both to the percentage of QUERY-W moves in a dialogue and the number of QUERY-W moves spoken per se.

8.3.6 Explain

A 3-way (2x2x2) ANOVA showed significant main effects of familiarity [$F(1, 62) = 10.65$, $p < .01$] and role [$F(1, 62) = 12.29$, $p < .01$]. There were no significant interactions. The mean number of EXPLAIN moves for familiarity and role are presented in Table 8.11.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	10.02	10.65	< .01
	Unfamiliar	7.17		
Role	Info giver	6.90	12.29	< .01
	Info follower	10.29		

Table 8.11 Mean number of EXPLAIN moves for an individual in each condition.

The results in Table 8.11 show that the information follower makes more EXPLAIN moves than the information giver. Familiar individuals volunteer more explanations than unfamiliar individuals.

When the data was standardised, a 3-way ANOVA revealed a significant main effect of role [$F(1, 62) = 75.32$, $p < .01$]. There was no significant main effect of familiarity or any significant interactions. The mean percentages of EXPLAIN moves for each role are shown in Table 8.12.

Variable	Condition	Mean % of Moves	F	p
Role	Info giver	5.74	75.32	< .01
	Info follower	9.90		

Table 8.12 Mean percentage of EXPLAIN moves for an individual in each condition.

These results show that the information follower made a greater percentage of EXPLAIN moves than the information giver.

Overall, the findings show that more EXPLAIN moves are made by the information follower than the information giver. Although more EXPLAIN moves were made by participants who were familiar, there was no difference in the percentage of EXPLAIN moves across the dialogue.

8.3.7 Align

A 3-way (2x2x2) ANOVA showed significant main effects of familiarity [$F(1, 62) = 7.06$, $p < .05$], and role [$F(1, 62) = 82.31$, $p < .01$]. There were no significant interactions. Table 8.13 shows the mean number of ALIGN moves in each condition.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	7.89	7.06	< .05
	Unfamiliar	5.49		
Role	Info giver	12.30	82.31	< .01
	Info follower	1.08		

Table 8.13 Mean number of ALIGN moves for an individual in each condition.

The results show that the information giver makes more ALIGN moves than the information follower. More ALIGN moves are also made by familiar individuals than unfamiliar.

When the data is standardised, a 3-way ANOVA reveals significant main effects of role [$F(1, 62) = 159.97$, $p < .01$], and familiarity [$F(1, 62) = 9.05$, $p < .01$]. There was also a significant interaction of familiarity and role [$F(1, 62) = 5.00$, $p < .05$]. Table 8.14 shows the mean percentages of ALIGN moves for the interaction between familiarity and role

	Information Giver	Information Follower
Familiar	**10.42**	**1.04**
Unfamiliar	**8.08**	**0.70**

Table 8.14 Mean % of ALIGN moves for familiarity and role.

Posthoc tukey tests showed that there was a significantly higher percentage of align moves spoken by the information giver to a familiar partner. As with the INSTRUCT moves, there appears to be a floor effect for the Information Follower, as they very rarely make any ALIGN moves.

These results also explain the significant main effects of role and familiarity. The information givers talk contained a significantly higher percentage of ALIGN moves than those of the information followers. The same is true for familiar and unfamiliar participants, although as the interaction indicates, this is only relevant for the information giver.

Taken together the results show that the information giver uses significantly more ALIGN moves than the information follower who rarely makes any ALIGN moves. For the information giver familiarity leads to more and a greater percentage of ALIGN moves being made in the dialogue. Familiarity causes speakers to check a familiar listener's understanding more frequently than that of an unfamiliar listener. This indicates that ALIGN moves are one of the moves which are increased by familiarity.

8.3.8 Clarify

Results of the 3-way (2x2x2) ANOVA studying CLARIFY moves showed significant main effects of familiarity [F(1, 62) = 11.31, p < .01], and role [F(1, 62) = 170.29, p < .01]. There was also a significant interaction of familiarity and role [F(1, 62) = 5.66, p < .05]. Table 8.15 shows the mean number of CLARIFY moves in each condition.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	5.80	11.31	p < .01
	Unfamiliar	3.61		
Role	Info giver	8.86	170.29	p < .01
	Info follower	0.55		

Table 8.15 Mean number of CLARIFY moves for an individual in each condition.

A post hoc Tukey test on the interaction between familiarity and role showed that there were significant differences between the number of CLARIFY moves made by the information giver when they were familiar rather than unfamiliar but that there was no effect of familiarity on the information follower, although given the low number of CLARIFY moves by the Information Follower, a floor effect is clearly possible. Table 16 shows the mean number of CLARIFY moves in each condition.

	Information Giver	Information Follower
Familiar	10.73	0.88
Unfamiliar	6.98	0.23

Table 8.16 Mean No. of CLARIFY moves for familiarity and role.

This result shows that familiarity affects the number of CLARIFY moves said by the information giver but has no effect on the information follower.

When the data was standardised, a 3-way ANOVA revealed significant main effects of role [F(1, 62) = 264.80, p < .01], and familiarity [F(1, 62) = 4.98, p < .05]. There were no significant interactions. The mean percentages of CLARIFY moves for each role and familiarity condition are tabulated in Table 8.17.

Variable	Condition	Mean % of Moves	F	p
Familiarity	Familiar	4.35	4.98	< .05
	Unfamiliar	3.44		
Role	Info giver	7.33	264.80	< .01
	Info follower	0.46		

Table 8.17 Mean percentage of CLARIFY moves for an individual in each condition.

Taken together the results show that the information giver makes significantly more CLARIFY moves than the information follower. When a participant is talking to a familiar person, they also tend to use more CLARIFY moves.

8.3.9 Reply-Y

Results of the 3-way (2x2x2) ANOVA studying REPLY-Y moves showed a significant main effect of familiarity [$F(1, 62) = 5.89$, $p < .05$] and a significant main effect of role [$F(1, 62) = 4.86$, $p < .05$]. There were no significant interactions. Table 8.18 summarises the means for familiarity and role.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	14.14	5.89	p < .05
	Unfamiliar	11.14		
Role	Info giver	11.56	4.86	p < .05
	Info follower	13.73		

Table 8.18 Mean number of REPLY-Y moves for an individual in each condition.

Table 8.18 indicates that there were more REPLY-Y moves made in the familiar condition than the unfamiliar. There were also more REPLY-Y moves made by the information follower than the information giver.

When the standardised data was analysed, a 3-way ANOVA revealed a significant main effect of role [$F(1, 248) = 67.66$, $p < .01$]. There was no significant main effect of familiarity or any significant interactions. The mean percentages of REPLY-Y moves for each role are tabulated in Table 8.19.

Variable	Condition	Mean % of Moves	F	p
Role	Info giver	9.18	67.66	< .01
	Info follower	14.56		

Table 8.19 Mean percentage of REPLY-Y moves for an individual in each condition.

When considered together the results show that a significantly greater percentage of the information followers dialogues are REPLY-Y moves than the information giver. This also corresponds to an overall increase in REPLY-Y moves per dialogue per se. Familiar individuals also make more REPLY-Y moves per se, but this does not correspond to an increase in rate of REPLY-Y moves being spoken.

8.3.10 Reply-N

Results of the 3-way (2x2x2) ANOVA studying REPLY-N moves showed a significant main effect of familiarity [F(1, 62) = 8.93, p < .05]. There was no effect of role or any significant interactions. Table 8.20 summarises the effect of familiarity.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	3.92	8.93	p < .05
	Unfamiliar	2.98		

Table 8.20 Mean number of REPLY-N moves for an individual in each condition.

Analysis of the standardised data indicated a significant main effect of role [F(1, 62) = 38.92, p < .01]. There was no significant main effect of familiarity or any significant interactions. The mean percentages of REPLY-N moves for each role are tabulated in Table 8.21.

Variable	Condition	Mean % of Moves	F	p
Role	Info giver	2.60	38.92	< .01
	Info follower	3.95		

Table 8.21 Mean percentage of REPLY-N moves for an individual in each condition.

When considered together the results show that the information follower makes significantly more REPLY-N moves than the information giver as a percentage of the dialogue, although this does not correspond to any difference in amount per dialogue per se. Familiar individuals on the other hand, make more reply-n moves per se, but this does not correspond to an increase in rate of REPLY-N moves being spoken.

8.3.11 Reply-W

Results of the 3-way (2x2x2) ANOVA studying REPLY-W moves showed a significant main effect of familiarity [$F(1, 62) = 9.15$, $p < .01$], and of role [$F(1, 62) = 4.37$, $p < .05$]. There were no significant interactions. Table 8.22 summarises the results of familiarity and role.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	4.43	9.15	p < .01
	Unfamiliar	2.95		
Role	Info giver	4.20	4.37	p < .05
	Info follower	3.17		

Table 8.22 Mean number of REPLY-W moves for an individual in each condition.

Analysis of the standardised data indicated no significant main effect of role or familiarity or any significant interactions.

When considered together the results show that although familiar pairs of individuals make significantly more REPLY-W moves than unfamiliar, this increase does not correspond to a higher percentage of the total moves in a dialogue. A similar effect is found for the information giver. This shows that REPLY-W moves do account for some of the extra talk in the longer, familiar interactions.

8.3.12 Acknowledge

Results of the 3-way (2x2x2) ANOVA studying ACKNOWLEDGE moves showed a significant main effect of role [$F(1, 62) = 332.75$, $p < .01$]. There was no effect of familiarity or any significant interactions. Table 8.23 summarises these results.

Variable	Condition	Mean Number of Moves	F	p
Role	Info giver	10.76	332.75	p < .01
	Info follower	32.77		

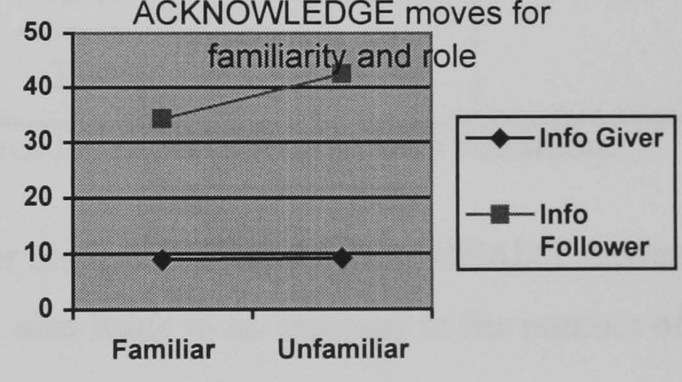
Table 8.23 Mean number of ACKNOWLEDGE moves for an individual in each condition.

When the standardised data was analysed, the 3-way ANOVA indicated significant main effects of familiarity [$F(1, 62) = 17.93$, $p < .01$] and role [$F(1, 62) = 636.30$, $p < .01$], and a significant interaction of familiarity and role [$F(1, 62) = 17.38$, $p < .01$]. Mean percentages of ACKNOWLEDGE moves for the interaction of familiarity and role are presented in Table 8.24 and Fig 8.1.

	Information Giver	Information Follower
Familiar	**8.99**	**34.28**
Unfamiliar	**9.17**	**42.47**

Table 8.24 Mean percentage of ACKNOWLEDGE moves for familiarity and role.

Fig 8.1. Mean % of ACKNOWLEDGE moves for familiarity and role



A post hoc Tukey test on the interaction of familiarity and role showed that a significantly higher percentage of the dialogue was ACKNOWLEDGE moves for the information follower when they were unfamiliar, however familiarity had no effect on the percentage of ACKNOWLEDGE moves spoken by the information giver.

Analysis of the main effects shows that overall, the information follower made a significantly higher percentage of ACKNOWLEDGE moves per dialogue than the information giver. Similarly, unfamiliar pairs made more acknowledgements than familiar pairs.

When considered together the results show that the information follower makes more acknowledgements than the information giver. If the information follower is familiar with the information giver however, they acknowledge less than if they were strangers.

8.3.13 Ready

Results of the 3-way (2x2x2) ANOVA studying READY moves showed significant main effects of familiarity [$F(1, 62) = 7.63$, $p < .01$] and role [$F(1, 62) = 90.53$, $p < .01$]. There was no effect of familiarity or any significant interactions. Table 8.25 summarises the results for role.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	8.85	7.63	p < .01
	Unfamiliar	7.24		
Role	Info giver	12.78	90.53	p < .01
	Info follower	3.31		

Table 8.25 Mean number of READY moves for an individual in each condition.

The results show that the information giver made significantly more READY moves than the information follower. Familiarity also leads to an increase in the number of READY moves an individual makes.

When the data was standardised, a 3-way ANOVA revealed a significant main effect of role [$F(1, 62) = 124.10$, $p < .01$]. This indicated that the information giver made a higher percentage of READY moves (10.55%) than the information follower (3.08%). There was no significant main effect of familiarity or any significant interactions.

Overall, the findings show that the information giver makes more READY moves than the information follower. Familiarity has no effect on the number of READY moves spoken.

8.3.14 Total Initiating Moves

To study the number of conversational games that were initiated, a 3-way (2x2x2) ANOVA studying the total number of initiating moves (INSTRUCT, CHECK, QUERY-YN, QUERY-W, EXPLAIN, ALIGN) was completed. The results of this showed significant main effects of familiarity [$F(1, 62) = 23.11$, $p < .01$], and role [$F(1, 62) = 89.37$, $p < .01$]. There were no significant interactions. The results show that the information giver makes more initiating moves, and consequently more conversational games on average in a dialogue, than the information follower, a finding which is also true for friends rather than strangers. These results are summarised in Table 8.26.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	56.43	23.11	< .01
	Unfamiliar	44.10		
Role	Info giver	65.06	89.37	< .01
	Info follower	35.47		

Table 8.26 Mean number of total initiating moves for an individual in each condition.

When the data is standardised, a 3-way ANOVA reveals a significant main effect of role [$F(1, 62) = 322.42$, $p < .01$], as well as a significant interaction of familiarity and role [$F(1, 62) = 8.44$, $p < .05$]. There was no main effect of familiarity. The mean percentages of initiating moves for each role are presented in Table 8.27.

Variable	Condition	Mean % of Moves	F	p
Role	Info giver	56.53	322.42	< .01
	Info follower	35.44		

Table 8.27 Mean percentage of total initiating moves for an individual in each condition.

Means for the interaction between role and familiarity are presented in Table 8.28.

	Information Giver	Information Follower
Familiar	56.16	37.76
Unfamiliar	56.89	33.11

Table 8.28 Mean % of moves that are initiating moves for familiarity and role.

A post hoc Tukey test on the interaction of familiarity and role showed a significant effect of familiarity for the information follower but not the information giver. This shows that when pairs are familiar, a higher percentage of the information followers' dialogue consists of initiating moves. This indicates that games are generally shorter for the information follower than the information giver.

Taken together, the results show that initiating moves are made more and more often by the information giver than the information follower. Familiar pairs make an increased number of total initiating moves on average in a dialogue but the rate of making initiating moves is increased for the information follower but not the information giver.

8.3.15 Total Queries

Results of a 3-way (2x2x2) ANOVA studying the total number of queries (QUERY-YN, QUERY-W) showed significant main effects of familiarity [$F(1, 62) = 15.37$, $p < .01$], and role [$F(1, 62) = 6.07$, $p < .05$]. There was also a significant interaction of familiarity and visibility [$F(1, 62) = 4.39$, $p < .05$]. Table 8.29 shows the mean number of total queries for an individual in each condition and summarises the results of the ANOVA.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	11.04	15.37	< .01
	Unfamiliar	8.66		
Role	Info giver	10.76	6.07	< .05
	Info follower	8.94		

Table 8.29 Mean number of total queries for an individual in each condition.

Table 8.30 shows the means for the interaction between familiarity and visibility.

	Visible	Not visible
Familiar	**11.16**	**10.92**
Unfamiliar	**7.50**	**9.32**

Table 8.30 Mean No. of QUERIES for familiarity and visibility.

These results show that, as with the QUERY-YN results, when strangers cannot see each other, they are less likely to ask questions than when individuals can see each other or are familiar.

When the standardised data was analysed, a 3-way ANOVA showed no significant main effects and no significant interactions. The results show that the number of queries an individuals makes across a dialogue is directly related to the overall length of a dialogue, rather than the independent variables of condition, and role. Familiar participants rather than unfamiliar and the information giver as opposed to the information follower, tend to make more moves overall and consequently tend to make more queries overall.

8.3.16 Total Replies

Results of the 3-way (2x2x2) ANOVA studying the total number of replies (REPLY-Y, REPLY-N, REPLY-W) showed a significant main effect of familiarity [$F(1, 62) = 10.71$, $p < .01$]. There was no significant main effect of role and no significant interactions. Table 8.31 shows the mean number of total queries for an individual in each familiarity condition.

Variable	Condition	Mean Number of Moves	F	p
Familiarity	Familiar	22.49	10.71	< .01
	Unfamiliar	17.06		

Table 8.31 Mean number of total replies for an individual in each familiarity condition.

When the standardised data was analysed, a 3-way ANOVA showed a significant main effect of role [$F(1, 248) = 81.43$, $p < .01$]. There was no significant main effect of familiarity and no significant interactions. Analysis of the means showed that the information follower makes a greater percentage of replies (mean = 21.78%) than the information giver (mean = 15.29%).

The results show that familiar individuals make more replies than strangers, but this is directly related to the overall length of a dialogue. The nature of the task leads to the information giver making more replies than the information follower.

8.3.17 Summary of Results

8.3.17.1 Distribution Pattern of Conversational Moves

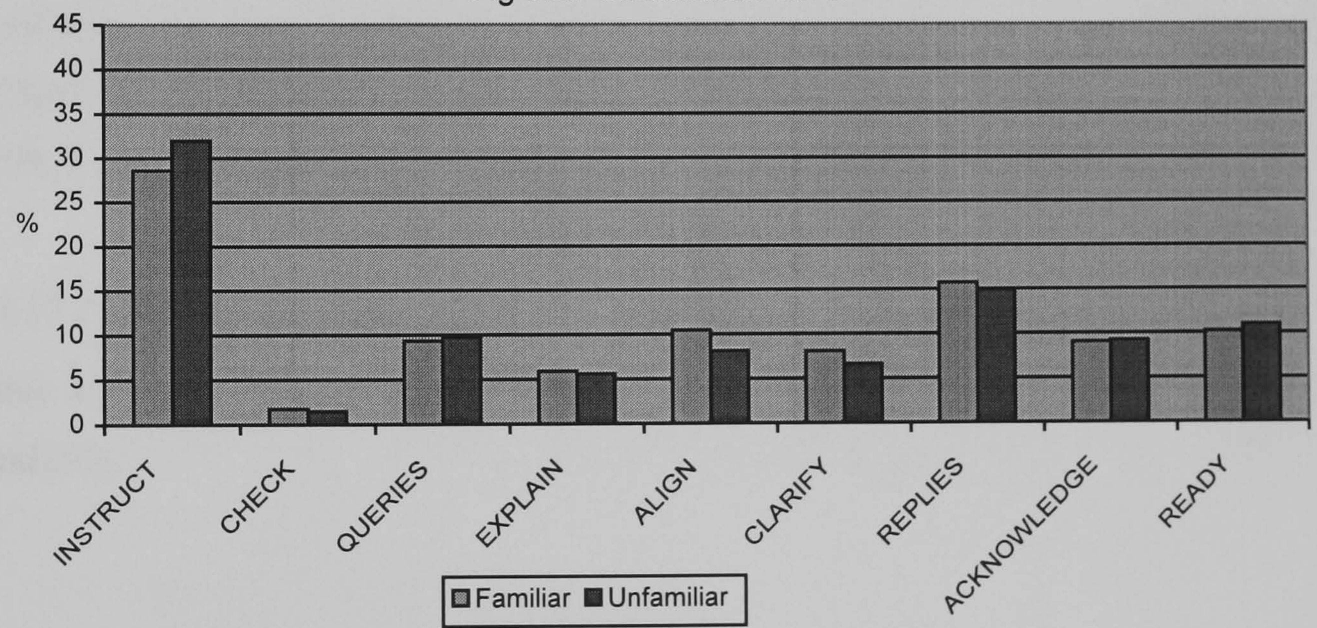
In order to consider the effects of role and familiarity in terms of their impact on the content of dialogue it is helpful to look at the relative distribution of conversational moves in each of the conditions. Table 8.32 shows the mean percentage of each type of conversational move for each role and familiarity condition and this is graphically presented in figures 8.2 and 8.3.

	Information Giver		Information Follower	
	Familiar	Unfamiliar	Familiar	Unfamiliar
INSTRUCT	28.64*	31.98*	0.32	0.19
CHECK	1.74	1.47	16.17*	14.11*
QUERIES	9.33	9.77	9.63	8.91
EXPLAIN	5.89	5.59	10.59	9.21
ALIGN	10.42*	8.08*	1.04*	0.70*
CLARIFY	8.06*	6.60*	0.64*	0.28*
REPLIES	15.69	14.90	22.85	20.72
ACKNOWLEDGE	8.99	9.17	34.28*	42.47*
READY	10.17	10.93	3.56	2.61

Table 8.32. Mean percentage of each type of move by role and familiarity.

* significantly different at $p < 0.05$ (for familiarity)

Fig 8.2. Information Giver



It is clear from Fig 8.2 that the major changes in the amount of each type of conversational move for the information giver are that familiarity leads to a decrease in the percentage of INSTRUCT moves spoken and a corresponding increase in ALIGN and CLARIFY moves.

Fig 8.3. Information Follower

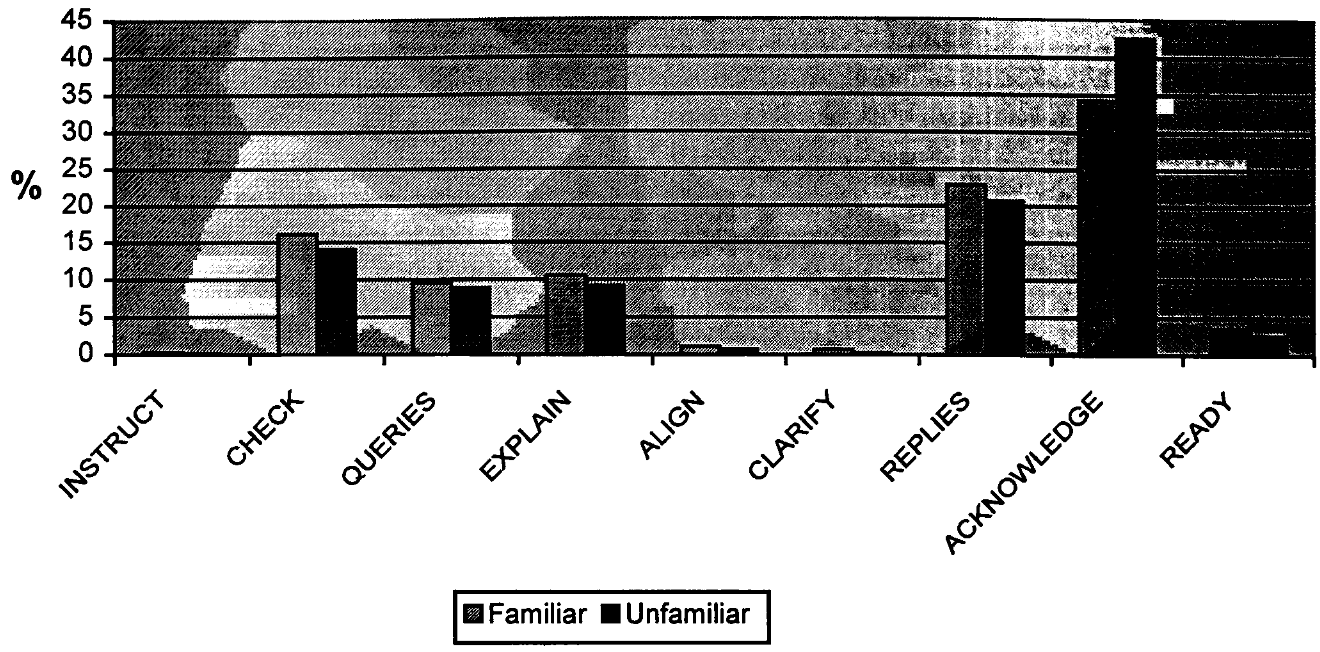


Fig 8.3 shows a strikingly different pattern of conversational moves being used for the information follower than the information giver, with hardly any INSTRUCT, ALIGN or CLARIFY moves being used. It also clearly shows the effect of familiarity on the information follower is a decrease in the proportion of ACKNOWLEDGE moves, compensated by increases in the level of CHECK, and REPLY moves as well as marginal increases in QUERIES and EXPLAIN moves.

8.3.17.2 Mean Number of Conversational Moves

Table 8.33 provides a summary of the mean number of each type of move in each condition.

	Information Giver			Information Follower		
	Unfamiliar	Familiar	Net diff.	Unfamiliar	Familiar	Net diff.
TOTAL MOVES	105.94	130.39	+24.45	83.03	104.02	+20.99
INSTRUCT	30.89	34.83	+3.94	0.17	0.34	+0.17
CHECK	1.91	2.58	+0.67	12.59	17.22	+4.63
QUERY-YN	8.80	9.31	+0.51	4.19	5.02	+0.83
QUERY-W	1.02	2.39	+1.37	3.31	5.36	+2.05
EXPLAIN	5.97	7.83	+1.84	8.38	12.20	+3.82
ALIGN	10.28	14.33	+4.05	0.70	1.45	+0.75
CLARIFY	6.98	10.73	+3.75	0.23	0.88	+0.65
REPLY-Y	10.13	12.98	+2.85	12.16	15.30	+3.14
REPLY-N	2.63	4.00	+1.37	3.33	3.84	+0.51
REPLY-W	3.77	4.64	+0.93	2.13	4.22	+1.11
ACKNOWLEDGE	9.86	11.66	+1.80	32.47	33.06	+0.59
READY	11.95	13.61	+1.66	2.53	4.09	+1.56

Table 8.33 Mean number of each type of conversational move for familiarity by role.

The figures in Table 8.33 summarise the effect familiarity has upon participants depending upon which role they have. Regardless of role, there is an increase in the total number of moves spoken, the net difference columns show which type of moves provide the extra content. For the information giver, extra content is provided by increases in the number of ALIGN, INSTRUCT, CLARIFY and REPLY-Y moves. For the information follower, the major moves affected are the CHECK, EXPLAIN, REPLY-Y and QUERY-W moves.

It should be noted however, that the only moves not to significantly differ overall were QUERY-YN, ACKNOWLEDGE, READY and CLARIFY (for the information follower alone). This indicates that even those moves which had low increases (e.g. REPLY-W) were still consistently affected across the experiment.

The point of this summary is to indicate which of the moves contribute most to the extra dialogue – it would be wrong to assume that just because a move shows a statistically significant difference across the course of the experiment that it contributes largely to the increase in dialogue overall. Thus, although REPLY-W moves are significantly affected by familiarity, REPLY-W on average only accounts for 1 extra move per participant, which in real terms accounts for less of the extra dialogue than some other moves. In contrast, the main causes of the extra talk between familiar speakers are the number of affirmative replies, clarifications, explanations and the amount of checking of their understanding.

8.4 Gaze

Instances of gaze were analysed for 32 of the dialogues where eye contact was possible. 16 of these involved familiar pairs and 16 involved unfamiliar pairs of participants. Gaze was recorded for whenever either individual looked up or when both looked up (mutual gaze). The data were analysed by using independent samples t-tests with familiarity as the independent variable.

Table 8.34 shows the mean number of gazes per dialogue by role, as well as the results of an independent samples t-test comparing familiar with unfamiliar dialogues.

Variable	Condition	Mean Number of Gazes	t (30)	p
Information Giver	Familiar	58.00	1.82	< .10
	Unfamiliar	41.88		
Information Follower	Familiar	42.56	1.82	< .10
	Unfamiliar	25.13		
Mutual Gaze	Familiar	25.69	2.51	< .05
	Unfamiliar	11.63		
Total	Familiar	126.25	2.34	< .05
	Unfamiliar	78.63		

Table 8.34 Mean number of gazes per dialogue divided by role and familiarity.

These results show that in this task based situation, familiar individuals gazed at each other significantly more often than unfamiliar individuals. Although this was the trend for each particular role, neither reached statistical significance on their own. The results also show that there were significantly more instances of mutual gaze by familiar individuals than unfamiliar individuals.

Considering that familiarity has been shown to lead to longer dialogues in terms of conversational moves, it could be argued that it is therefore likely for familiarity to lead to more gaze purely as a function of the length of the dialogue. To overcome this problem, the gaze data were standardised for the length of the dialogue by dividing the number of gazes by the total number of moves and multiplying by 100 to give the rate of gazing per 100 moves. Analysis of this standardised data showed the same pattern as for the non-standardised results with significantly more instances of mutual gaze between familiar pairs than unfamiliar pairs [$t(30) = 2.26$, $p < .05$]. The rate of one-way gazing did not significantly differ according to familiarity.

These results show that when dialogues are standardised for length there are still more instances of mutual gaze by familiar participants than unfamiliar pairs. This suggests that mutual gaze is not related to the length of the dialogue, whereas the amount of gaze from any given individual may simply be a function of the length of the dialogue.

8.5 Discussion

8.5.1 Role Differences

When considering the results of the Map task it is always important to consider the differences between the two roles. In such a task-oriented scenario, the two participants have very different individual goals to reach the joint goal of producing a well drawn map route. The Information Giver is aiming to dispense knowledge that they have (of the map route) to the Information Follower who in turn is attempting to receive information and confirm that their understanding matches the information that is being dispensed.

Conversational Games Analysis has shown very clear differences in the pragmatic functions of the conversational moves by participants. The moves which are said more by the information giver both as a share of the dialogue and per se are the READY, INSTRUCT, QUERY-YN and ALIGN moves. This is explained by the fact that the information giver is in control of the dialogue, by virtue of the fact that they have more of the relevant knowledge. They therefore say when they are ready to start a new game (READY) and also do virtually all the instructing, telling the Information Follower what to do with the pen (INSTRUCT). As a further consequence of passing on information, it is the information giver who spends more time seeking confirmation of the other participants understanding (ALIGN). QUERY-YN is used primarily to gain understanding of the Information Follower's position so that instructions can be suitably tailored e.g. “”do you have a ruined monastery?” or “can you see a well?”. It is soon shown to be inefficient to simply say “go around the well” when it is known that the landmarks on the two maps do not match.

The information follower on the other hand makes substantially more CHECK, QUERY-W, EXPLAIN and ACKNOWLEDGE moves than the Information Giver (despite the fact that the Information Giver makes more moves overall). As a consequence of the Information Giver doing most of the instructing, the Information Follower has more need to check their own self-understanding of a message or instruction (CHECK) and also offer feedback to show they have heard and understood (ACKNOWLEDGE). Since they have less information to begin with, it should also come as no surprise that they are required to ask more open-ended style questions (QUERY-W) as opposed to the more direct ones asked by the Information Giver who knows exactly the information they require. Finally, as information is exchanged between the two roles it is clear that almost all the information being given by the Information Follower is contained in EXPLAIN moves, most of the Information Givers information can be given in the form of an instruction.

Replies appear less affected by role as do CLARIFY moves. This is not surprising when considering the task. The most usual form of dialogue during the task may appear something like this:

IG: Do you see a bridge? (QUERY-YN)

IF: Yes, (REPLY-Y) it is in the middle of the page (EXPLAIN)

IG: Go past the bridge on the left (INSTRUCT)

IF: The left? (CHECK)

IG: Yes (REPLY-Y)

IF: Ok (ACKNOWLEDGE)

This extract indicates how an instruction would be passed on from the information giver to the information follower, and typically shows how REPLY-Y moves may be equally used, while INSTRUCT and QUERY-YN moves are information giver based and CHECK, EXPLAIN and ACKNOWLEDGE moves may be more information follower based.

8.5.2 Familiarity

The results of Boyle et al (1994), which used the same dialogues to investigate the effects of familiarity on other aspects of the task, showed that familiar pairs of participants tended to use more conversational turns and more words and consequently performed the task slightly better. They also found that familiar pairs interrupted each other less and had fewer instances of overlapping speech, but when controlled for the length of dialogue in turns they found that there was in fact no increase in the rate of interruptions. The present results help to explain why the increase in dialogue leads to improvements in performance.

The present study has shown that familiarity leads to an increase in the total number of moves, irrespective of role. Analysis of which types of move were affected shows that all moves showed an increase when participants were familiar (across role and visibility conditions) with the exception of QUERY-YN, ACKNOWLEDGE, READY and CLARIFY (info follower not affected). The moves which that showed the largest increases however were REPLY-Y, EXPLAIN, CHECK, QUERY-W, and ALIGN.

With so many of the individual moves showing increases it may appear that familiarity is leading simply to a general increase in the amount of talk being used across the board. Close analysis of the pattern of those moves that are affected however shows a definite pattern. All the moves which show the greatest increases are interactive moves which elicit information or offer information. In contrast to that, those moves which decrease or are unaffected by familiarity are those moves which are the least interactive or least rich in information. It appears that when friends are offering feedback there is a greater percentage of moves which are richer in information such as EXPLAIN instead of ACKNOWLEDGE. They also appear to be more willing to check understanding, CHECK and ALIGN are similar types of move - the former checks self understanding of a message while the latter checks the other participants understanding of a message or goal. It is interesting to note that these are not role dependent, both participants show an increase in the percentage of the dialogue that is taken by these moves when familiar. It is therefore safe to conclude that familiarity leads to an increase in speakers willingness to interact with their partners and to offer and seek extra information – notably about how well they understand one another. This is reflected in an increase in checking, aligning and general queries (as opposed to specific questions).

So it appears that familiarity leads to a less task oriented discussion, with more explaining and more open-ended questions, feedback tends to be richer and consequently there is less basic feedback in the form of ACKNOWLEDGE moves.

It is fair to say that individuals probably feel more comfortable talking to a friend than a stranger and consequently the increased amount of conversation may be attributable to the increased feelings of comfort. It may feel easier to ask questions and to check understanding for instance with a friend than with a stranger. One of the major questions raised by the research of Boyle et al. however, is why do familiar pairs of individuals perform the task better than unfamiliar pairs?

There are a number of possible hypotheses as to why performance may be improved between friends. Familiar pairs of individuals may simply be offering/eliciting more information, or they may have a better understanding of what is said by their partner (be "on the same wavelength" more). Finally it is possible that the way in which familiar people complete the task is different - they may provide more precise instruction than the unfamiliar pairs.

The improvement in performance must be the result of the information follower understanding the information given to them by the information giver better. Consequently either more relevant knowledge is disseminated by the information giver, or the knowledge disseminated by the information giver is understood better by the information follower, or both. Since the results show that there was little extra instruction given by the information giver – only a mean of 3.5 extra instructions per dialogue - it is unlikely that the information giver was more extensive in their instructions, considering they used a similar number of increments.

The increase in the number of QUERY-W and EXPLAIN moves does suggest an increase in the amount of knowledge being disseminated, as opposed to an improved overall understanding of information offered by the information giver. There are also increases in the number of CHECK and ALIGN moves which should elicit more feedback. Since the extra moves do appear to be task-oriented then it is fair to assume that they contribute to the information followers understanding more, rather than simply "being on the same wavelength". It is likely then that familiar pairs perform better due to spending more time explaining and offering richer feedback than unfamiliar pairs.

The increase in the amount of mutual gaze in which familiar pairs indulge may provide extra communicative information which is unaccounted for in the CGA. This, however cannot be the major reason for improved performance in the familiar condition because it only applies to those dialogues where eye contact was present. As there was no interaction between eye contact and familiarity in terms of performance then the mutual gaze can not have been the overall cause of improved task performance.

Once a performance improvement is discounted as a possible reason for the increase in mutual gaze, attention must turn to another hypothesis for the purpose of mutual gaze in the familiar condition. Published research suggests that mutual gaze may have an affective communicative function (e.g. Exline & Winters, 1966; Rubin, 1970; Rutter et al. 1978) or be regulatory (e.g. Kendon, 1967). The present research, showing that mutual gaze is not linked to any impact on task performance, suggests that the function of mutual gaze in the map task may indeed be mainly affective as opposed to regulatory. This would account for the increase in mutual gaze in the familiar condition.

It appears that the key to improved performance between familiar pairs of individuals in the Map task may be the increase in conversational content caused by the closer monitoring of the dialogue. This is in accordance with Ansfield et al. (1995) who suggested that it may not be greater experience of communicating with an individual which led to familiar pairs showing improved communicative performance. Instead it may be a result of more information actually being spoken. The increases in the number of CLARIFY and CHECK moves for example, show that there is more effort made to ensure understanding has been reached before moving on to the next piece of information. The present results show that improvements are due to a different style of communication which leads to more actual communication rather than what is being said being more intuitively understood.

It is important to note that the increase in performance appears to be caused by an increase in the overall amount of communication being used, rather than specifically extra verbal communication. This distinction is shown by the fact that Boyle et al. (1995) showed that a lack of visual cues leads to an increase in the number of moves and consequently the amount of verbal information being used. This did not lead to an increase in performance however, presumably because the extra moves were to make up for the non-verbal communication lost by removing the visual channel. Familiarity has also led to an increase in the amount of verbal communication, but has also led to an increase in performance. If one takes into account the visual cues of communication in addition to the verbal cues, it appears that performance may be improved by an increase in the *overall* amount of communication information used.

8.5.3 Familiarity and the Lack of Visual Cues

The present research found interactions between eye contact and familiarity only for QUERY-YN and QUERIES overall. It appears that when strangers are deprived of visual cues they ask fewer questions than if individuals can either see each other or are familiar. With the exception of this finding, the results suggest that the effects of familiarity on dialogue occurred across eye contact conditions. Similarly the results indicate that the effects of depriving eye contact occur regardless of whether or not individuals are familiar with each other.

The knowledge of the independence of these variables is important because it suggests that the effects of familiarity will occur across different communicative contexts. It seems likely that familiar individuals will talk more and with richer feedback and greater interactivity regardless of how well they can see each other, which potentially is important for VMC.

8.5.4 Comparison with Proximity findings

In Chapter 5 , where CGA was used to investigate the dialogue differences between individuals talking to a stranger across a video link who appeared either close or far away, several differences in the process of communication were found. When an individual appeared zoomed in close there was an overall increase in the total number of moves. When considering which moves were more frequent in the close condition, it was found that there was an increase in the number of QUERY-YN, REPLY-W and ACKNOWLEDGE moves and that ACKNOWLEDGE moves were the most heavily affected. In the HCRC corpus however, familiar pairs performing the map task showed no increase in the number of acknowledgements or Query-YN moves used, although they did show an increase in the number of REPLY-W moves.

It must be remembered that the proximity research was completed using a different task (the Financial advice task). Despite this comparisons between the effect of familiarity and proximity may still be useful. Both tasks had relatively similar numbers of moves being spoken by participants. In the financial advice task, dialogues consisted of a number of 105.9 moves in the close condition and 71.9 in the far condition. The map task resulted in each participants mean contribution to a dialogue being 94.5 when talking to an unfamiliar interlocutor and 117.2 when familiar. The number of turns per total dialogue was also larger in the map task (179.7 turns for familiar; 145.7 turns unfamiliar) than the financial advice task (140.1 - close; 100.6 - far). The comparable increase in numbers of turns and moves suggest a structural resemblance between the familiar and the close dialogues.

When comparing the tasks it is clear that the role of customer in the financial advice task is much closer to the role of Information Follower in the Map task. Results showed that the Information Follower used more CHECK, QUERY-W, REPLY-Y, and EXPLAIN moves. It has already been explained that these moves show an increase in the amount of information being elicited by the Information Follower. This was also characteristic of the close condition in the Financial Advice task. In the Financial Advice task the major difference between the two conditions was in the style of the dialogue. The far condition led to very lecture style interactions compared to the much more interactive style of the close conditions. The major difference was the increased number of ACKNOWLEDGE moves, which provided extra feedback and essentially provide the message "yes I am listening and I understand, keep talking" in comparison to the lack of feedback present in the far condition.

To compare this with the Map task, in the Map task familiarity appears to elevate participant feedback from acknowledging and obeying, to checking, explaining and questioning. In the Financial Advice task close proximity appears to elevate participant dialogue from generally listening silently to acknowledging and giving feedback. Even although the two findings show differences in the distribution of the specific moves which are affected, it is clear that both results could represent different stages on a continuum of interaction and feedback. In their own way, both familiarity and close proximity appear to raise the amount of information elicited by participants.

There are of course some key task differences. In the financial advice task, participants have very little information to begin with and have to learn and give opinions on new information. In the Map task on the other hand, both participants have a relatively large amount of information when they begin and are required to exchange this information. Secondly in the financial advice task, although the conversational moves of both the advisor and the customer were recorded and analysed, only the customer was subject to the manipulation of condition. While these differences have implications for individual moves in any analysis (one would not expect differences in the amount of ALIGN moves by the customer in the financial advice task for instance due to the lack of information being offered by the customer to the financial advisor which would require aligning) these differences should not affect the net changes in general style of an interaction as caused by an experimental manipulation.

8.6 Conclusions

The results have indicated that in a task-oriented conversation, familiarity may lead to discussion which is less didactic, with more explaining, more open-ended questions and richer feedback than when unfamiliar individuals are interacting. Familiar pairs of individuals also gaze at each other more than unfamiliar pairs of individuals. When individuals are unfamiliar they may also be more persistent or less willing to yield the floor to an interruption by the other person than when they are familiar.

Depriving participants of eye contact when completing the task, clearly removes the opportunity for non-verbal feedback, and consequently an increase in short verbal feedback moves is recorded. The moves which are increased the most are acknowledgements, positive replies and statements of readiness - all of which can be conveyed using facial expression.

Familiarity and eye contact also appear to be independent of each other. Although both familiarity and a lack of eye contact are known to increase the length of dialogues in terms of the number of turns spoken to complete a task, it was previously unknown if there was any relationship between the two. By virtue of the fact that there were no interactions between eye contact and familiarity for any moves, the present research has shown that all differences caused by familiarity occur regardless of whether eye contact is present or not and vice versa.

So could familiarity be equal to perceived proximity? For this to be claimed the effects of familiarity on communication must roughly equate to the effects of perceived proximity on communication. The results showed that participants in the Map Task who were familiar tended to offer and elicit richer information from each other than if they were unfamiliar. In the Financial Advice task, participants in the close condition also tended to be richer in the information they gave than participants in the far condition. Although the type of conversational moves affected differed between tasks, it can be seen that the effects on communication elicited by perceived proximity and familiarity were broadly similar. So the answer to this question is broadly a yes, familiarity and perceived proximity do appear to have similar effects on communication.

This research has illustrated the benefit of analysing the content of dialogues to illuminate the nature of structural differences of dialogue. Analysis of content can illustrate differences in style of conversation which may in turn help to explain more precisely the psychological bases underlying communication.

9 Chapter 9. Conclusions

The aim of this thesis was to investigate the role of proximity in video-mediated communication. This question required an analysis of whether visual impressions of proximity could exist in typical video-mediated communication situations, and if so, how they were governed. Consideration was also given to the potential effects on communication such cues may have, and the consequences such effects may have in an interpersonal interaction. All of this empirical work was framed within a paradigm that emphasised the applied nature of the research. Many of the findings should be of direct relevance to those who use, design and implement video-mediated communication technologies.

The results of the research have indicated that impressions of proximity can be manipulated within the context of video-mediated communication technology. Manipulations in the perceived proximity have been shown to lead to differences in the manner in which individuals communicate. This work has also offered insights concerning a possible connection between the perception of proximity and familiarity, and also proved to be enlightening for gauging individuals' impressions of technology.

9.1 Isolating Proximity in a Video-Mediated Interaction

The first requirement of investigating the impact of proximity in video-mediated communication was to determine whether indeed it existed in the first place. This in turn required an attempt to isolate how proximity may be perceived in video-mediated interactions. To investigate this, the concept of proximity in a face-to-face interaction was taken as the starting point.

In a face-to-face interaction, proximity can be considered as either a physical or a perceptual dimension. The physical basis of proximity is concerned with the actual distance between two individuals. In a video-mediated interaction this can be expressed both in terms of the physical remoteness of the two individuals and the physical distance from each individual to their screen and camera. On a perceptual level on the other hand, proximity can be considered in terms of how far away an individual appears to be - as if this was a face-to-face interaction. The perception of proximity is not directly quantifiable, as it relies on how close an individual feels.

Some research has suggested that individuals may communicate differently depending upon how far away they believe the other person to be. Moon (1998) showed that when individuals were being interviewed by a computer which they believed was located far away, they answered differently to when they believed that they were being interviewed by a computer they thought was closer. Participants completed a survey on a computer and were told that the results of the survey were being collected either on the computer which they were to use, a computer based in a city a few miles away, or in a computer based several thousand miles away.

The survey that they completed was actually the Balanced Inventory of Desirable Responding (BIDR) which measures response distortion along two scales – impression management and self-deceptive enhancement. Impression management is the deliberate tendency to over-report desirable behaviours and under-report undesirable ones. Self-deceptive Enhancement, on the other hand is designed to measure overconfidence in one's judgements. Moon found that there was a significant effect of distance on a scale of impression management but no effect of self-deceptive enhancement. When participants believed that the results of the survey they were completing were to be collected by a computer based long way away, they showed a tendency to distort their responses to make themselves appear more favourable.

A second experiment showed that under similar conditions, the perceived distance of a remote computer sending messages of advice on how to complete a task were more persuasive than identical messages that participants believed originated from only a few miles away.

The present thesis however, was only concerned with the perceptual element of proximity. The research did not manipulate actual proximity in a video-mediated interaction.

It is not claimed that these two dimensions of proximity are mutually exclusive. In both a face-to-face or video-mediated interaction the perceptual experience of proximity is of course partially determined by the physical reality of proximity. This does not, however, preclude analysis of the two separately. It is only when analysed separately that an understanding of the precise cause of any effect on communication can be assessed. This thesis has consistently sought to analyse in detail the role of perceived proximity in the communication process without the interference of manipulating the physical distance between the interlocutors.

9.2 The concept of perceived proximity

For the purposes of this thesis, perceived proximity has been defined as the experience of the apparent distance of another individual. This can be considered in terms of how close someone feels as well as how close they appear.

Just as there are two ways of considering proximity, so there are also factors within perceived proximity. Cues available in video-mediated communication can be both aural and visual. For the purposes of this thesis, only the visual factors were manipulated. Audio may be a crucial factor for increasing the experience of proximity in a remote communication situation, indeed it may even be a stronger cue to perceived proximity than visual cues, but to study both was beyond the scope of this thesis. Attempting to alter both during the tasks would have led to confusion as to which factor was causing the effects of proximity. As it was, merely a visual manipulation was shown to produce differences in the perception of proximity. In choosing which aspect of perceived proximity to study, there were a number of reasons for selecting only the visual basis of proximity.

The primary reason for studying visual rather than audio cues was that the visual channel is the sense that human beings use above all others. They are also more diverse, consequently leading to a wider scope to the study. Visual cues to proximity in a video-mediated communication system include the size of the image, the amount of background visible in an image, the presence of other objects and so on. Audio cues to distance may be only volume and frequency. The availability of visual information is also the addition that sets video-mediated communication apart from the older technology of audio-only communication.

There is another reason for studying the visual rather than audio cues to proximity. In a video-mediated interaction the nature of the visual information being transmitted is primarily controlled by the sender (who can adjust their camera). Audio information, on the other hand, is usually controlled by the receiver (by way of a volume control). As such, individuals using video-mediated communication equipment tend to have control of the visual cues to proximity that are being sent (as in a face-to-face interaction). The equivalent aural cues to proximity however, are likely to be controlled by the receiver. The value in this distinction lies for anyone who seeks to use perceived proximity as a factor in driving the nature of an interaction in a particular way.

9.3 Demonstrating Perceived Proximity

The first challenge of the thesis was to investigate whether or not perceived proximity could exist in video-mediated communication in any systematic way. Study 1 sought to investigate this by asking participants to create a feeling of a particular interpersonal distance simply by changing the zoom on a video camera. Participants attempted to equate the feeling of distance of an image of a remote individual with that of a co-present individual.

Despite the fairly abstract nature of the task, participants were able to grasp the concept of equating the feeling of distance of a remote individual presented on a monitor with that of an individual sitting in the same room. The results showed that all participants stated that increasing the focal length on the camera led to an impression of the remote individual appearing closer.

The questionnaire results also showed that participants felt that the general social rules of proximity as identified by Hall (1966) did apply to individuals represented by video-mediated technology. The impressions of presence of the remotely presented individuals were less pronounced but followed the same pattern as those of co-present individuals.

The fact that participants could reliably equate changes in camera zoom with changes in the impressions of distance shows that perceived proximity can exist in a video-mediated interaction. The next question to be answered was what factors in an image govern impressions of perceived proximity.

9.3.1 The nature of perceived proximity

Study 1 compared two factors in an image that could potentially have affected the perceived distance of an individual in an image. Participants completed the task with the remote individual presented either using a large or a small window. This allowed an investigation comparing whether the principal visual cue to perceived distance was image size or the image to background ratio.

By comparing the images that individuals selected to represent a particular distance using either a large or a small video window, it was found that the factor that seemed to govern impressions of distance was image size. With the exception of the closest distance (2.5 feet), there was no significant difference in the mean size of image chosen to represent a particular distance. This effectively meant that when using the small window to select an image representing a particular distance, the camera was zoomed in more than when viewed using the large window.

This result applied to video-mediated images. Study 3 sought to investigate whether it would apply equally to still images. In several ways, Study 3 was effectively Study 1 in reverse. Rather than ask participants to change an image until they appeared to be a certain distance away, the procedure for this study involved participants judging how far away individuals actually were in photographic images presented to them.

In this study the zoom of the camera never changed, but the actual distance of the individuals was manipulated. The other variables were image size and the presence or absence of objects in the background.

The results of Study 3 indicated again that image size was important in the perception of distance. A large image tended to lead to the impression that the individual in the photograph appeared closer, while the opposite was true of small photographs. However image size was not the only variable that was shown to have an affect on the impressions of proximity. The presence of objects in the background of an image also led to an impression of closeness of the individual.

These results serve to underline that perceived proximity is dependent upon the combination of a number of factors. When considered alongside the results of Study 1 it appears that when individuals are asked to make judgements about impressions of the distance of individuals in photographs or a video image, the size of the object image (i.e. that of the individual) is the best predictor of perceived distance, although there are other factors that are likely to have an effect.

9.3.2 Perceiving Absolute Distance

The results of studies one and three have investigated factors which affect the perception of absolute distance (also termed egocentric distance) in either a video image or a photograph. It is known that under natural, unrestricted viewing conditions, the perception of distance is remarkably consistent (Baum & Jonides, 1979). Indoor observations are usually slightly underestimated, while outdoor observations are slightly overestimated (Da Silva & Fukusima, 1986).

When distance perception is attempted under conditions which remove the surrounding context (for instance low illumination) there is evidence that observers tend to underestimate the distance of far objects and overestimate the distance of near objects. This phenomenon is known as the "Specific Distance Tendency" (Owens & Leibowitz, 1980; Gogel, 1982; cited in Durgin, Profitt, Olson & Reinke, 1995). The research presented in Study 3 demonstrated that the introduction of particular types of contextual information can be powerfully influential concerning participants ratings of absolute distance. The presence of objects in the background or a large image size both gave the impression of closeness. Although of course in pictures the context may be false, and lead to fewer correct answers, these results are consistent with the Specific Distance Tendency – the absence of cues to distance leads to a default distance being selected.

These results have also gone some way to explaining the nature of the effect of camera zoom. Persson (1998) concluded that close-up images did create impressions of proximity, but offered no objective study demonstrating this. The present research has not only shown that close-up images do indeed create feelings of closeness, but has also gone on to investigate why. By comparing the effects of image size with the relationship between the image and the background, we have shown that the major factor in creating impressions of proximity in mediated materials may be the size of the object image. Furthermore, we have also demonstrated that this is not the only factor, but that changing the background in an image may also affect the apparent proximity of the object image.

The results of studies 1 and 3 showed that perceived distance could be manipulated in a video image. As proximity is known to affect conversational interactions in a face-to-face situation, the next aim of the thesis was to investigate whether manipulations in the appearance of a video image could lead to differences in the way in which individuals interacted using video-mediated communication technology.

9.4 The effects of proximity in a video-mediated interaction

Study 2, as reported in Chapter 4, endeavoured to investigate whether or not differences in the perceived proximity of a remote individual could lead to differences in the manner in which individuals would interact with them. The study showed that there were significant differences between the styles of the dialogues of those individuals who participated in the close condition and those that took part in the far condition. The style of the dialogues in the close condition was generally more interactive than those in the far condition.

The results showed that when the financial advisor appeared zoomed in, participants said more, spoke more often and interrupted more than when the financial advisor appeared zoomed out. This also led in general to interactions taking longer in the close condition.

There were two striking aspects of the results of this study. The first is that the questionnaire data showed no significant differences between participants in the close condition and the far condition on any of the scales of subjective opinion regarding opinions of either the task or the technology. Since it is known from the dialogue analysis that there were behavioural differences between participants in the two groups, this result suggests that the dialogue analysis is a more sensitive tool than the questionnaire.

The other interesting aspect of the results is that all of the behavioural effects that were discovered applied to the customer, not the financial advisor (with the exception of turns, which by definition must almost certainly apply to both interlocutors in a dyadic interaction). This is strong evidence showing that the experimental manipulation was having a direct effect on the behaviour of the participants. This is because the financial advisor's view of the customer didn't differ between conditions.

The results of Study 2 showed that it was possible for manipulations in perceived distance to affect the behaviour of individuals in a video-mediated interaction. These results did not in themselves however, indicate the nature of the differences between the interactions occurring in the two conditions. Chapter 6 detailed the application of Conversational Games Analysis to the results of Study 2.

The results of the Conversational Games Analysis being applied to Study 2 showed that the key difference between typical dialogues could be found in the number of ACKNOWLEDGE moves made, indicating an increase in the amount of verbal feedback offered. Participants in the close condition tended to make significantly more ACKNOWLEDGE moves than participants in the far condition. There were other differences but these accounted for far less of the difference in the number of turns (participants asked on average 1-2 more questions, but gave 17-20 more acknowledgements per dialogue).

These results show that when the remote interlocutor appeared in a close-up view, participants gave more verbal feedback during the conversation than if the remote interlocutor appeared to be farther away. No measurements were made of the amount of non-verbal communication information that participants used during the course of their interactions. No comment can therefore be made as to whether this information, in the form of nods, gestures, facial expressions and so on, differed between conditions of proximity.

Studies investigating the difference between face-to-face and video-mediated interactions have suggested that video-mediated interactions do not support these types of non-verbal communication. Cukor, Baer, Willis, Leahy, O'Laughlan, Murphy, Wither & Martin (1998) report that because these types of information exchange are not supported in a video-mediated interaction, the benefits for using video in telemedicine are limited to an increase in social presence. They go on to suggest that low cost videophones may be as suitable for telemedicine as expensive

ISDN based systems. Although the present experiment used high quality video conferencing, there is reason to suggest that the effect of perceived proximity would also work using systems with a low frame rate. The results of study three showed that even in still images, differing appearances of proximity could exist. For perceived proximity, image size is likely to be more crucial than frame rate – something that may not be true for other forms of non-verbal communication information.

The results of Study 2 showed that manipulations in the camera view did lead to differences in the way in which individuals communicated. Study 1 had already shown that changing the level of camera zoom used on the image of a video-mediated individual led to differences in the impressions of perceived proximity. In light of this, Study 4 aimed to combine the findings of studies 1 and 2 by manipulating both camera zoom and image size in a video-mediated interaction.

The results of Study 4 were disappointing in many ways. The main result of Study 2 was not replicated, as no significant difference was found between the participants in the Monitor Close condition and the Monitor Far condition. Indeed no significant between groups effects were found regarding the dialogue analysis in Study 4. The reasons why there were no effects are unclear, however the most likely explanation is the difference in the financial advisor. In Study 2 the financial advisor took a fairly informal style in comparison to the confederate in Study 4.

Because of the lack of success of replication in Study 4, it is arguably unsurprising that the results of the study failed to produce any significant effects regarding the two new conditions involving the life size video image.

Despite this however, Study 4 did produce interesting effects when comparing those participants that were very interactive against those that were very non-interactive. Those participants that were interactive had significantly higher recall scores than those participants that took a less interactive approach to the task. The implications of this result are that anything that makes individuals more interactive is liable in turn to lead them to greater understanding. Although in Study 4, there was no between groups effect of interactivity it is likely that the participants in Study 2 would have shown an effect of condition on understanding. Care must be taken in interpreting this result but it seems fair to suggest that an individual appearing close in a video-mediated interaction is more likely to lead to greater understanding than an individual appearing far away for the individual who is interacting with the image.

9.4.1 Video-Mediated Communication

The research presented in Studies 2 and 4 have demonstrated that perceived proximity is a factor which can impact upon the manner in which individuals communicate using a video link. Previous research in the field of video-mediated communication has shown that different variables affect the structure of participants' dialogue in an interaction. An example is the study of O'Connail, Whittaker & Wilbur (1993) who showed that the quality of a videoconference link impacted upon turn-taking in a video-mediated interaction. The present research has shown that turn taking can also be affected by the video image which is transmitted. A close image could lead to more instances of overlapping speech and more turns overall than a far image.

O'Malley et al. (1996) showed that increases in turn taking were found in video-mediated interactions compared to audio-only groups. Tang & Isaacs (1993) found that more turns were made in face-to-face than video-mediated interactions. When considering the present research alongside these previous studies, it seems that increases in turn taking may signify a greater informality within conversation, rather than as a sign of inefficiency in completing a task.

The analysis of the content of the dialogues in study 2 has also indicated that as a variable within a videoconferencing system, perceived proximity can affect the content of individuals' conversations when using them. Doherty-Sneddon et al. (1997) demonstrated that the effect of direct eye contact being available in a video link was to increase the amount that individuals checked one anothers understanding during a conversation. The results of the present research showed that when talking to a remote interlocutor who appeared very close, individuals made more verbal feedback in the manner of acknowledgements than if the remote interlocutor appeared far away.

The results of the dialogue analysis of the Financial Advice task has outlined the effects of perceived distance as a variable within video-mediated communication systems. Studies 2 and 4 have shown that differences in perceived proximity can lead to differences in the manner in which individuals interact in a video-mediated interaction. The most logical next concern is to understand why such effects may occur.

9.5 Why does proximity affect communication?

To have a proper understanding of what effects anything is likely to have, it is reasonable to ask the question of why such effects will occur. Proximity appears to be a fundamental form of non-verbal communication information. In a face-to-face interaction, the function of proximity is to communicate certain information. The information that it can be used to convey may include that of interpersonal relationships, engagement in a conversation, as well as information about the content of a conversation.

In face-to-face communication, proximity is often used in conjunction with both body position and gaze. In a dyadic interaction for example, if a stranger positions themselves closer than would normally be expected, individuals tend to alter their body position so that they are less face-on to the stranger.

In Study 4, body position was analysed between groups. It had been hoped that the results would show that when the financial advisor appeared close, that participants would tend to sit further back from the screen or turn their bodies side on to the screen, indicating a desire to increase the interpersonal distance between themselves and the financial advisor. The data comparing body position between those in the close and far conditions however yielded no significant differences. Indeed, as many participants sat forward as sat back.

The reasons for this can only be speculated upon. Firstly the task contained a reading element, and in reading the information on the screen, there was always going to be a distance from the screen that was most comfortable. Turning side on to the screen was also not a realistic body position to read the text from. No information however, was recorded regarding how often each individual gazed at the video window portion of the screen.

Another reason for the lack of results is that participants may have felt constrained by the experimental setting. They were led into the experimental room and asked to sit on a pre-positioned seat. There was a camera facing them, which was clearly on view. It is possible that these factors led to some participants feeling less natural with their body position.

It is unclear exactly what effect the movement of body position may have had in the experiment. When body position was moved, is this a sign of the lack of comfort in the situation at first, and consequently does moving provide a neutralising effect?

9.5.1 Familiarity

Another area of interpersonal communication that has close links with proximity is familiarity. As was discussed in depth in Chapter 8 , interpersonal familiarity appears to share many of the same effects on communication as proximity, and also affects proximity directly. In Chapter 4 it was noted that there was a great similarity between the effects of strangers standing closer than normal in a face-to-face situation and the differences between conversations between familiar and unfamiliar pairs in experimental tasks. The only difference between the effects of familiarity and the effects of proximity appeared to be that familiarity led to a decrease in the number of interruptions made, while close proximity led to an increase in the number of interruptions made. This led to the hypothesis that close proximity led to an impression of "perceived familiarity".

In Chapter 8 , this hypothesis was tested, by investigating the relationship between familiarity and proximity in terms of dialogue content. The results indicated that the effect of familiarity on dialogues between familiar individuals completing the Map Task was broadly similar to the effect of proximity in the financial advice task. The dialogues differed in terms of richer feedback although how this was expressed differed between the tasks. Participants who interacted with an interlocutor who appeared close in the Financial Advice task were more likely to make acknowledgements than those who interacted with an interlocutor appearing far away. Participants who completed the Map task with a familiar partner made fewer acknowledgements but volunteered more information than those who completed the task with a stranger.

The results were not clear enough to constitute conclusive proof that perceived proximity leads to an impression of perceived familiarity but they at least suggest a link between the two.

9.6 The Use of an Experimental Approach

Throughout this thesis, an experimental approach has been adopted. As was alluded to in the opening chapter, this was not the only available option, a more ethnographic approach could also have been used. It is worth considering whether with hindsight, the experimental approach was a success.

The overall aim of the thesis was to explore whether or not proximity, as a form of non-verbal communication, could exist in video-mediated communication, and if so, how is it governed and what effects may it have on communication? The experimental approach appears to have been reasonably successful as far as these aims are concerned. Tightly controlled laboratory experiments have been used to demonstrate that differences in camera zoom can lead to differences in the perception of proximity of a remote individual; have shown that these differences can cause effects on how individuals communicate with one another; have informed as to the most likely determinants of virtual proximity and finally have demonstrated similarities between the effects of proximity and those of other contextual variables (familiarity). All of these results show the success of the experimental approach. Without the tight controls that could be placed on the task and the configuration of the technology it would not have been possible to systematically show these effects.

However, there is one area in which the experimental approach has failed to be informative. Laboratory based experiments do not provide information as to the actual effects of differences in proximity in real life applications of video-mediated communication. Attempts were made to keep the context of experiments as realistic as possible but still, there is no escaping the fact that they were laboratory experiments. Consequently, no firm conclusions can be drawn in this thesis with regard to the actual effects that changing the perceived proximity of individuals in a video-mediated interaction. Field studies and naturalistic observations may have been better able to inform the research in this manner.

There is of course a trade-off between the number of field studies and experimental studies that could have been done. Time and resources are finite, and to conduct field studies would have meant that some of the experimental studies would have had to have been sacrificed. The main advantage of the purely experimental approach taken is that it has broadly shown the potential for proximity to be important as a variable in remote communication. It seems appropriate that some future research into the subject would attempt to understand the extent of this importance for real world applications.

9.7 Advantages of a multi-faceted approach

Although the general methodology throughout the thesis was experimental, within this paradigm many different dependent variables were measured. When attempting to understand something as complex as "the way people interact" it is important to attempt to use as wide a range of methodologies as possible. The work outlined in this thesis has used a wide range of research methods, borrowing tools from many different areas of psychology to gain a multi-faceted approach. Consequently in this thesis, open ended questionnaire items are reported next to detailed dialogue coding. Scaled questionnaires and scored results are reported, but so are non-quantifiable impressions of an interaction.

This application of a diversity in approaches leads to a more detailed understanding of how people interact with new technology. Without such a range of approaches it is likely that different conclusions would have been reached.

When considering the best way of finding out what an individual thinks, it is traditional simply to ask them. However, this is not always the most reliable measure. The present research has shown that comparing what people do with what they report may lead to contradictory findings. Consider the results found in Study 2 which showed that while questionnaire data indicated no differences in the way individuals perceived any aspect of the task, the behavioural analysis showed that there were differences between groups as to how individuals behaved during the task.

The results of Study 2 showed that behavioural measures can be more sensitive than questionnaires. However this sensitivity does not come without a cost. The behavioural measures described in this thesis are for the most part extremely labour intensive. Transcription takes a minimum of 8 hours for every single hour of two-party dialogue, coding the speech requires a similar amount of time, and recording body movements is also extremely time consuming. When these methods are applied to groups of individuals rather than dyadic interactions the amount of time each takes is liable to increase exponentially.

As well as conducting research into conversational behaviour thoroughly, the present research seeks also to inform researchers on a practical level into ways of shortcutting the process of behavioural analysis of communication. After having completed the full conversational games analysis in Study 2, a less time consuming solution was sought to analyse the dialogues from Study 4. Study 2 had shown that the key to the extra interactivity in dialogues lay in the number of acknowledgements that the participants made during conversation. This permitted a more focussed approach to the coding of the dialogues in Study 4. The analysis in the latter study sought only examples of acknowledgements, rather than coding the entire dialogues, and consequently saved a lot of coding time.

This method of directed coding was only made possible by having a specific hypothesis for the nature of the expected differences between dialogues from different conditions. It would not be suitable for purely explorative research where the aim was to investigate the general pattern of dialogues under a specific condition. Another advantage of using a coding scheme that is generally applicable is that this type of directed analysis can be employed.

Some of the questionnaire items asked participants about how likely they were to use this equipment were it available. Although these questions may appear to have added little to the empirical findings of the research they were useful in two ways. Firstly they provided information which was useful to the industrial sponsors of this research when considering video-mediated technology in terms of a product.

Secondly, they gave an extra flavour of how the participants felt about the task and the technology.

This multi-faceted approach is particularly relevant in the field of HCI, which is a multidisciplinary field. Some researchers have gone so far as to suggest that this interdisciplinary approach may lead to major scientific advances for example:

> "The combination of psychology, engineering, sociology, ergonomics and so on may prove to be as fruitful as was the combination of physics, chemistry, and biology Watson and Crick employed to untangle the structure of DNA." (Atkinson, 1999).

9.8 The Financial Advice Task

It is important to reflect on some of the methodological choices that were made during the course of the research. Though this research did employ a wide variety of approaches, central to the research was the Financial Advice task. When designing the Financial Advice task in particular, there were many decisions which could have been made differently. The following section will attempt to address some of the issues involved in the decision process and evaluate the success of these decisions with the benefit of hindsight.

9.8.1 The Use of Confederates

Central to the design of the Financial Advice task was the use a confederate who was the same in all of the trials. It was felt that using a confederate would be useful in trying to standardise the experience of the customer. Had a confederate not been used then all participants would have had different financial advisors. This would have consequently led to them receiving different amounts of advice, levels of persuasion, quality of advice, etc.

Although throughout this thesis, the financial advisor has been referred to as a confederate it is worth pointing out that there were differences to the way a confederate may be used in certain social psychology tasks. In a social psychology task, and indeed in the experiments into proximity which used confederates (e.g. Argyle & Dean, 1965), the confederate is an individual who is aware of the experimental conditions. In a deception task for instance, a confederate may be asked to attempt to deceive a participant in half the trials and not to in the other half. In the present thesis, the financial advisor was unaware of which experimental condition they were taking part in and even what the experimental manipulation was.

The main problem of using a confederate in a task that investigates interpersonal communication is that dialogue is a co-operative activity. Any differences in dialogue could therefore result from the direct experimental manipulation on either interlocutor or indirectly from the manner in which one individual reacts to a change in the behaviour of the other. This is why it was important that the confederate did not know which experimental condition they were in, so there was no risk of them altering their behaviour accordingly.

The other problem of using a confederate in a communication study is that instructing one individual to behave the same way in all trials, may inhibit the natural fluidity of the cooperation and interplay within a dialogue. This is likely to make the analysis of any differences between communication caused by the experimental manipulation more conservative and therefore null results more likely. The fact that in both study two and four significant differences between the conditions were shown, merely strengthens those findings.

A related issue was the decision to choose to change the confederate between studies one and three. The failure of the attempt to fully replicate the findings of the original financial advice task in study four indicated that the choice of financial advisor could influence the outcomes of the task. The results showed that in the later study, participants were less interactive and there was no significant difference in the number of speaking turns made between those in the different conditions.

The reason a different confederate was used between experimental trials was largely practical - the original confederate was not available for the second study - however there were also advantages to this. The lack of replication has shown that the effect of altering the appearance of a video image of a remote interlocutor is not totally independent of the remote individual. The research has shown that changes in camera zoom can lead to individuals behaving differently but dialogue is a two-party activity and these behaviour changes may not always be enough to affect the conversation. This underlines the point that many factors within a video link all combine to create the video-mediated interaction experience, including the remote interlocutor.

A more specific issue about the use of the confederate was the choice of the individual to play the confederate. The two confederates were chosen for the particular attributes that they brought to the situation. The confederate in the original study was a graduate psychologist and so understood about the running of psychology experiments. This was useful in that she understood how she should behave, and was able to provide useful information regarding her experience during the trial period before the experimental trials began. The confederate for the later study had different attributes, she was a professional actress and had also worked as a call centre agent. These experiences were advantageous in two ways. She was able to complete the task in a manner that helped her complete the task realistically, and she was extremely comfortable in front of the camera.

It is possible that these attributes did in fact lead directly to the lower levels of interactivity in Study 4. Her previous experience in a telephone call centre meant that she had the skills which audio call centre staff develop. In communication experiments, audio-only conditions have consistently been shown to be less interactive than face-to-face situations (e.g. Boyle et al., 1994). Although video-mediated conditions have been more like those in a audio-only situation than face-to-face, there have been studies showing that interactions conducted without any visual communication channel are more formal, and thus less interactive than those using videoconferencing technology (e.g. Tang & Isaacs, 1993). It is possible that during the role of the financial advisor, the confederate reverted to the role of audio-only call centre agent and used the conversational style she had learned there. The confederate in study one however, never had this skill base and so may have started from a more face-to-face style conversational set.

The acting skills of the second confederate may also have had an influence on the manner of the interaction. In each study the financial advisor was instructed to look at the camera as much as possible, to create a greater impression of presence for the customer. It is conceivable that the confederate in study 4 was more comfortable looking at the camera, and thus looked at it more than the confederate in study 1. One of the factors which has been demonstrated to be linked to proximity is eye gaze. In face-to-face situations the amount that individuals make eye contact with each other decreases as they get closer (Argyle & Dean, 1965; Paterson, 1977). Thus it is possible that the ability of the confederate in the study 4 to look at the camera more, may have created an impression of greater distance. It is therefore possible that this worked against the manipulation of camera zoom attempting to create an impression of closeness.

Because no data was recorded of the financial advisors behaviour during the task it is only possible to speculate about the probable reasons for the difference between the advisors. In future experiments, it would be useful to study the behaviour of the financial advisor as well as that of the customer.

9.8.2 Only Changing the View of the Financial Advisor

Another choice which was made was only to manipulate the participant's view of the financial advisor, while the financial advisor's view was the same in all conditions. The experiment was designed this way to ensure that any findings could be directly attributed to the difference in experience of the customer.

A consequence of this decision was that the two interlocutors had different impressions of each others interpersonal distance. This meant that the interaction was one step further removed from simulating a face-to-face style interaction, as no equivalent interaction with asynchronous interpersonal distance can occur. The alternative would have been to alter the views of both interlocutors so that both got a feeling of proximity or distance. This would have made the interaction closer to a simulation of a face-to-face encounter, and thus perhaps allowed a more direct comparison with the literature on co-present findings.

There were several reasons why changing the financial advisor's view of the customer was not considered to be ideal. On a practical level, this would have alerted the confederate to the difference between the conditions, consequently they may have unwittingly altered their behaviour in each condition as a result of knowing when they were in each and guessing a hypothesis differed according to condition. By not altering the experience for them at all, the financial advisor never knew what the different conditions were, or when they were in one condition rather than the other, so could not have altered their behaviour accordingly.

Altering the views of both interlocutors between conditions would also have been unsatisfactory on a theoretical level. Had any significant differences emerged when both views differed, then it would have been unclear what caused the difference. Changing both views could make the origin of any behavioural differences unclear.

The risk in changing only the view of one individual was that if no significant differences had been found then it would have been unclear if perceived differences in proximity would never have an effect or were prevented from being expressed by the asynchrony in proximity experience.

The decision to change only one view appears to have been vindicated by the results that were found. Of all the results that showed a difference between the dialogues that were generated in each of the two experimental conditions, the results showed effects that only affected the customer, such as the increased level of interruptions when the remote interlocutor appeared closer. Since only one the view of the customer differed, anything that only affects the customer would appear to result directly from the experimental manipulation.

It would be of interest to investigate the effects of manipulating the view of both participants. This would permit more direct comparisons between face-to-face investigations of the effects of proximity which have been conducted and video-mediated research.

9.8.3 The Suitability of the Financial Advice Task

When evaluating a task, the most important question one must ask is whether or not the task successfully met the goals for which it was designed. The goals for which the financial advice task was designed were to design a task that would allow the study of a broad range of communication behaviour in a video-mediated interaction. The task had to contain a social or negotiation element to the task and retain some realism and relevance to practical applications of video-mediated communication.

The task has been largely successful in satisfying these requirements. Communication was measured in several ways, although a measure of eye contact and gaze would have been useful. Most participants felt that the task was realistic and all seemed to conduct the task sensibly, adapting to the role play well. Although no significant effects of persuasion were identified between conditions, the results did indicate that 46% of participants in study two selected the choice that they were being persuaded towards. This figure is well above chance but well below any ceiling effect so does appear suitable for a study of persuasion.

9.9 Comparing Remote Communication with Face-to-face?

Much of the study of remote communication has focussed on the relationship between how people interact when using different technologies for communication and how these compare to a face-to-face situation. The research presented throughout this thesis is investigating the what may have been expected to be a uniquely face-to-face phenomenon – proximity, and it's role in remote communication.

This research has demonstrated that proximity as a form of non-verbal communication information can have an effect across a video link. Studies of this kind may also lead to a problem of terminology. It must be accepted that by definition, remote technology will never replicate the experience of participating in a face-to-face interaction. Consequently, whenever an attribute of a co-present interaction is demonstrated to exist in remote interactions, there is a need for an appraisal of the terminology used to support it

Some may wish to argue against the use of the term "proximity" to be applied to this research. The problem is that proximity is to a large extent a fact, and actually means the physical distance between two objects and thus should be expressed in units of length. When considering how far away an individual feels however, there is no term which adequately describes the experience. "Perceived interpersonal distance" could also be considered misleading as it suggests that individuals are deliberately considering how far away someone appears. Perhaps "pseudoproximity" would be better, implying as it does that the proximity is apparent or in a sense false.

9.10 Academic and Industrial Collaboration

In modern society at present there is an increasing requirement for science to justify itself by demonstrating a degree of "usefulness" to those beyond the academic community. Though arguments may rage on a philosophical level about the relative advantages or disadvantages of this approach, it is clear that it is here to stay. This is especially prevalent when considering new technology. As technology continues apace, it is essential that research continues to assess its impact on potential users.

The work for this thesis was funded by an ESRC CASE studentship. As a collaboration between NCR and the University of Glasgow, the work always sought to retain a commercially applicable flavour. At every stage of the thesis, the aim was to understand how a new technology might actually affect the individuals who use it and how that knowledge may be applied to in a commercial situation. It is hoped that the knowledge created in this thesis will lead to an understanding that will inform both designers and users of video-mediated communication equipment, hopefully serving to improve the technology and the experience of using the technology for everyone.

For the present research, the relationship between academia and industry proved to be mutually beneficial. The application of theoretical investigation to commercial issues led to findings that were both generalisable as well as useful. It is conceivable that were this research to have been conducted without the influences of both the commercial and academic communities that it would have been different. It is certainly difficult to imagine commercial companies indulging in deeply laborious issues such as dialogue coding. Similarly, academic interests may not have given a lot of time to the commercial setting of the technology and some of the questioning of users that sought direct impressions from people of how this technology might fit into their lives.

It is the blend of these styles which has led to a well rounded basis to this research. Rarely was there a conflict in approach to research, the process was almost entirely complementary. Both sectors had an influence and the research is all the better because of that.

9.11 Implications of this research

The research presented in this thesis has implications for all individuals connected with video-mediated communication technology at all stages of production and use. A more complete understanding of the impact of technology on communication behaviour is essential for using the technology most effectively. The knowledge of how individuals interact when using technology can aid both the design process and empower the end users to get the most from the technology.

The primary experimental finding in this research is that ostensibly minor alterations to the configuration of a video-mediated communication system can result in tangible differences in the manner in which individuals interact when using it. This type of knowledge should be able to help people optimise systems for their particular use.

The present research has shown that changing the level of zoom on a camera can lead to differences in how individuals will interact when using the video-mediated communication system and also how well they are likely to understand what they talked about.

The direct implications to come from this research are that if you are setting up a video-mediated communication system for communication between two people, you should consider the goals of the interaction. If the goals are to encourage conversation, and to increase the level of interactivity – for example to hold a brainstorming meeting – then the camera should be zoomed in so that the images of the individuals appear close up. A similar set-up should be used if the primary aim of the interaction is understanding – for example a technical support customer helpline. Alternatively, if the aim of the interaction is to disseminate information, and little interactivity is desired – as in a ticket sales line perhaps – then the images of the individuals should appear zoomed out or smaller.

Such a consideration consequently assumes the media interface to be fairly seamless. It is almost certainly true however, that the nature of the interface will affect feelings of presence. A video link will often be only one part of a multimedia interface that may also include objects, text and backgrounds for example. An interface may also be either two or three dimensional, thus automatically introducing an extra depth cue. It is known that in a face-to-face situation the layout of the room, the presence of objects and so on can all influence the nature of an interaction. So in a multimedia based interaction, might the presence of a table for instance decrease the impression of personal closeness?

For those parties interested in designing technology, or increasing its use in everyday life, this research has attempted to gain some insight into how the technology is perceived by those that use it. Participants reported that they generally liked the technology, had very few problems using it but would still be more likely to use a face-to-face option if it was available in this scenario. Bouwhuis (2000) points out that usability is generally considered to be the standard criterion for interactive products in human factors engineering. Despite this however, many products that satisfy usability criteria fade away while others which are ostensibly less usable may gain widespread acceptance. Bouwhuis' research showed that the extent to which a product may fit into the lifestyle and the personal environment of the user may be as important as usability. To increase the acceptance of the product, Bouwhuis argues that the more complex a product is, the more configurable it must be. The present research has demonstrated a need to understand the possible effects of making a product more configurable.

It is worth noting that manipulating perceived proximity in a video-mediated interaction can lead to novel types of interaction. It is possible, for instance, to have an interaction where one individual may appear closer than the other individual – something that cannot happen in a dyadic face-to-face interaction. Quite how this affects the nature of potential interactions is unknown but is another area where future research can be done.

9.12 Future directions

One of the first points made in this thesis was that a video link should not be considered as a singular entity. Instead it is better considered as an umbrella term encompassing a technology with many different variables. Within this, proximity was taken to be a single factor. It must be remembered of course that single factors do not work in isolation with each other. Rather, a number of factors may all influence each other. To this end, proximity is no different to any other. The research presented in this thesis has studied proximity in isolation and as such, the claims made are always made within the premise that, proximity may lead to an effect, providing all other things are equal.

The work provided in this thesis is only a small part of a much larger field of investigation. The research has shown that the concept of video-mediated communication is a broad one, encompassing many individual factors. For a scientist, the concept of "having a video-mediated interaction" may be equated to the concept of two principal characters "having a conversation" to a director of a play. There are many factors that must be described in order to fully understand the nature of the communication situation. In the case of the director these would include the setting, the lighting, the volume etc. For the HCI researcher or designer interested in video-mediated communication technology this will include the window size, the zoom on the camera, the quality of the image etc.

The research in this thesis has concentrated solely on the concept of perceived proximity in a video-mediated interaction. There are many other factors that will also contribute to communication using new media, which have so far yet to be studied. There is a lot of future work to be done investigating the impact on communication of differences in the configuration of video-mediated communication technology.

Within the field of proximity, it would be interesting to know how issues of physical proximity may interact with the issues raised in this thesis. The relative importance of visual and aural cues to the concept of perceived proximity remains unknown. Similarly, the extent of the relative boundaries of perceived proximity is unknown.

Video-mediated communication technology has been around for over 30 years and represents many hopes for many people. The potential for the technology is enormous, the ability to connect people globally in a way that audio-only communication cannot manage. This potential will be realised fastest however, if there is a greater understanding of the manner in which individuals interact with technology.

10 References

Acker, S. R. & Levitt, S. R. (1987). "Designing videoconference facilities for improved eye contact." *Journal of Broadcasting & Electronic Media, **31**,* 181-191.

Albert, S. & Dabbs, J. M. (1970). "Physical distance and persuasion". *Journal of Personality and Social Psychology, **15**,* 265-270.

Alexander, C. N. Jr. & Rudd, J. (1981). Situated identities and response variables. In J. T. Tedeschi (Ed.), *Impression Management Theory and Social Psychological Research* (pp. 83-103). New York: Academic Press.

Allegier, A. R., & Byrne, D. (1973). "Attraction towards the opposite sex as a determinant of physical proximity". *Journal of Social Psychology, **90***, 213-219.

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H., & Weinert, R. (1991). "The HCRC map task corpus." *Language and Speech, **34**,* 351-366.

Anderson, A. H., Newlands, A., Mullin, J., Fleming, A. M., Doherty-Sneddon, G., & Van der Velden, J. (1996). "Impact of video-mediated communication on simulated service encounter." *Interacting with Computers, **8**,* 193-206.

Anderson, A. H., Bard, E. G., Sotillo, C., Doherty-Sneddon, G. & Newlands, A. (1994). "The effects of face-to-face communication on the intelligibility of speech." Paper presented at the *Annual Conference of the Experimental Psychology Society,* University of Exeter, July 5-7. Cited in O'Malley & Langton (1994).

Anderson, A. H., Mullin, J., Newlands, A., Doherty-Sneddon, G. & Fleming, A-M. (1994). "Video-mediated communication in CSCW: effects on communication and collaboration." *Proceedings of Workshop on VMC, at CSCW 94,* Chapel Hill, NC.

Anderson, A. H., O'Malley, C., Doherty-Sneddon, G., Langton, S., Newlands, A.. Mullin J., Fleming, A. M. & Van der Velden, J. (1997). "The impact of VMC on collaborative problem solving: An analysis of task performance, communicative process and user satisfaction". In Finn, K. E., Sellen, A., Wilbur, S. (Eds.) *Video-Mediated Communication.* pp133-156. Lawrence Erlbaum Assoc. Hillsdale. New Jersey.

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 (Eds.) *Video-Mediated Communication*, pp51-74, Mahwah, NJ: Lawrence Earlbaum Associates.

Ansfield, M. E., DePaulo, B. M., & Bell, K. L. (1995). Familiarity effects in non-verbal understanding: Recognizing our own facial expressions and our friends'. *Journal of Nonverbal Behaviour, **19**,* 135-149.

Argyle, M. & Cook, M. (1976). *Gaze and Mutual Gaze.* Cambridge: Cambridge University Press.

Argyle, M. & Dean, J. (1965) "Eye-contact, distance and affiliation". *Sociometry,* ***28***, 289-304.

Argyle, M. (1988). *Bodily Communication.* 2nd Edition. Routledge. New York.

Argyle, M., Ingham, R., Alkena, F., and McCallin, M. (1973). The different functions of gaze. *Semiotica, 7,* 10-32.

Atkinson, R. (1999). personal message posted on www virtual library on knowledge management http://www.brint.com/wwwboard/messages/4797.html, Sep 1999).

Austin, D. (1996). "Remote Possibilities". *Banking Technology*, Feb. 1996, pp 26-30.

Baer, L. Cukor, P., Coyle, J. T. (1997). Telepsychiatry: Application of telemedicine. In R. L. Bashshur, J. H. Sanders, & G. W. Shannon (Eds.), *Telemedicine: Theory and practice* (pp. 265-290). Springfield, IL: Charles C Thomas.

Barber, P. J. & Laws, J. V. (1994). "Image Quality and Video Communication" In Damper, R.; Hall, W. & Richards, J. (Eds). *Proceedings of IEEE International Symposium on Multimedia Technologies & Their Future Applications*: pp 163-178. London. Pentech Press.

Baum, D. R., & Jonides, J. (1979). Cognitive maps: Analysis of comparitive judgements of distance. *Perception and Psychophysics, 2,* 161-166.

Beattie (1979, 81, 86) cited in O'Donnell & Draper (1997) but not listed in references.

Becker, F. D. & Mayo, C. (1971). Delineating personal distance and territoriality". *Environment and Behaviour, 3,* 375-381.

Blanck, P. D., Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1980). Sibling resemblances in nonverbal skill and style. *Journal of Nonverbal Behaviour, 4,* 219-226.

Blohm, W. (1996). "Video dynamic range compression of portrait images by simulated diffuse scene illumination." *Optical Engineering, 35,* 255-261.

Blokland, A. & Anderson, A. H. (1998). Effect of low frame rate on intelligibility of speech. *Speech Communication, 26,* 97-103.

Bouwhuis, D. G. (2000). Parts of life: configuring equipment to individual lifestyle. *Ergonomics, 43,* 908-919.

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, 37,* 1-20.

Brauer, D. V. & DePaulo, B. M. (1980). Similarities between friends in their understanding of nonverbal cues. *Journal of Nonverbal Behaviour, 5,* 64-68.

Brave, S., Ishii, H. & Daly, A. (1998). Tangible interfaces for remote collaboration and communication. *Proceedings of CSCW'98*, 169-178.

Bridgeman, B., Lewis, S., Heit, G., & Nagel, M. (1979). Relationship between cognitive and motor-oriented systems of visual position perception. *Journal of Experimental Psychology: Human Perception Performance, **5**,* 692-700.

Brown, G., Anderson, A. H., Yule, G. and Shillcock, R. (1984). *Teaching Talk.* Cambridge, U.K.: Cambridge University Press.

Bruce, V. & Green, P. R. (1992). *Visual Perception: Physiology, Psychology and Ecology.* 2nd Edition. Lawrence Erlbaum Associates. Hove and London (UK). Hillsdale (USA).

Bruce, V. (1996). The role of the face in communication: implication for videophone design. *Interacting with Computers, 8,* 166-176.

Buxton, W. & Moran, T. (1990). "EuroParc's integrated interactive intermedia feature (iiif): early experiences." In *Proceedings of the IFIP WG8.4 Conference on Multiuser Interfaces and Applications,* pp11-34.

Capella, J. N. (1981). "Mutual influence in expressive behaviour: Adult-adult and infant-adult dyadic interaction". *Psychological Bulletin, **89**,* 101-132.

Capella, J. N. (1983). "Conversational involvement: Approaching and avoiding others". In J. M. Wieman & R. P. Harrison (Eds.), *Nonverbal interaction* (pp. 113-152). Beverly Hills, CA: Sage.

Capella, J. N. (1986). "Reciprocal and compensatory reactions to violations of distance norms for high and low self-monitors". In M. L. McLaughlin (Ed.), *Communication yearbook 9* (pp. 359-376).

Capella, J.N. (1994). "The management of conversational interaction in adults and infants". In M. L. Knapp & G. R. Miller (Eds.) *Handbook of Interpersonal Communication* 2nd ed. (pp. 380-418).

Carletta, J. (1996). Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics, 22,* 249-254.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Newlands, A., Doherty-Sneddon, G., & Anderson, A. H. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics, **23**,* 13-31.

Chapanis, A. (1975). "Interactive Human Communication" *Scientific American, **232**,* 34-42.

Chapanis, A., Ochsman, R., Parrish, R. & Weeks, G. (1972). "Studies in interactive communication: The effects of four communication modes on the behaviour of teams during cooperative problem solving". *Human Factors, **14**,* 487-509.

Clark, H. H. (1996). *Using Language.* Cambridge, UK. Cambridge University Press.

Clark, H. H. & Brennan, S. E. (1989). "Grounding in communication". In L. B. Resnick, J. Levine and S. D. Teasley (Eds.) *Perspectives on Socially Shared Cognition.* American Psychological Association. Washington.

Clark, H. H. & Schaefer, E. F. (1987). "Collaborating on contributions to conversations." *Language and Cognitive Processes, **2**,* 19-41.

Clark, H.H. & Schaeffer, E.F. (1989). "Contributing to discourse". *Cognitive Science, 13,* 259-294.

Cohen, K. (1982). "Speaker interaction: video teleconferences versus face-to-face meetings". In *Proceedings of Teleconferencing and Electronic Communications,* pp. 189-199.

Colston, H. L. & Schiano, D. J. (1995). "Looking and lingering as conversational cues in video-mediated communication". In *Proceedings of CHI '95* (Short papers).

Cook, M. (1970). "Experiments on orientation and proxemics". *Human Relations, 23,* 61-76.

Cook, M. (1979). *Perceiving Others.* London: Methuen.

Cook, M. & Lalljee, M. G. (1972). "Verbal substitutes for visual signals in interaction". *Semiotica, **6**,* 212-221.

Cool, C., Fish, R. S., Kraut, R. E. and Lowery, C. M. (1992). "Interactive design of video communication systems". In *Proceedings of CSCW '92* (Toronto ON, Oct, 1992), ACM Press, 25-32.

Corning P.A. (1997). "Holistic Darwinism: 'synergistic selection' and the evolutionary process". *Journal of Social and Evolutionary Systems, **20**,* 363-400.

Cukor, P., Baer, L., Willis, B. S., Leahy, L., O'Laughlan, L., Murphy, M., Wither, M., & Martin, E. (1998). "Use of videophones and low-cost standard telephone lines to provide a social presence in telepsychiatry". *Telemedicine Journal*, 4, 313-321.

Cunningham, S. (1996). "A staff developer's briefing on video-conferencing" *UCoSDA Briefing Paper* no. 38, September 1996.

Da Silva, J. A. & Fukusima, S. S. (1986). Stability of individual psychophysical functions for perceived distance in natural indoor and outdoor settings. *Perceptual and Motor Skills, **63**,* 891-902.

Dakins, D. R., & Jones, E. (1996). "Cream of the crop: 10 outstanding telemedicine programs". *Telemedicine and Telehealth Networks.* Miller-Freeman. Retrieved November 3, 1997 from the World Wide Web: **http://www.infomarket.ibm.com/prbb**

De Greef, P., & Ijsselsteijn, W. (2000). "Social presence in the photoshare tele-application". *PRESENCE 2000 – 3rd International Workshop on Presence*, 27-28 March 2000. Delft, The Netherlands.

Doherty-Sneddon, G. (1995). "The role of visual signals in childrens communication". Paper presented at the British Psychology Society Conference (Developmental Section). University of Strathclyde.

Doherty-Sneddon, G., Anderson, A.H., O'Malley, C., Langton, S., Garrod, S. and Bruce, V. (1997). "Face-to-face and video mediated communication: a comparison of dialogue structure and task performance". *Journal of Experimental Psychology: Applied, **3**,* 1-21.

Efran, M. G. & Cheyne, J. A. (1974). "Affective concomitants of the invasion of shared space: Behavioural, physiological and verbal indicators". *Journal of Personality and Social Psychology, **29**,* 219-226.

Ekman, P. (1971). "Universals and cultural sdifferences in facial expression and emotion". In Cole, J.K. (Ed) *Nebraska Symposium on Motivation.* Nebraska University Press, Lincoln. Cited in O'Donell & Draper (1997).

Ekman, P. (1973). *Darwin and Facial Expression: A Century in Review.* New York: Academic Press.

Exline, R. V. & Winters, L. C. (1966). "Affective relations and mutual glances in dyads". In S.S. Tomkins and C. Izard (eds) Affect, Cognition and Personality. London: Tavistock.

Farmer, S. M. & Hyatt, C. W. (1994). "Effects of task language demand and task complexity on computer-mediated work groups". *Small Group Research, **25**,* 331-336.

Festinger, L. (1957). *A Theory of Cognitive Dissonance.* Stanford: Stanford University Press.

Field, M J. (1996). "Telemedicine: A guide to assessing telecommunications in health care". Washington D. C.: National Academy Press.

Fish, R. S., Kraut, R. E., Root, R. & Rice, R. E. (1993). "Video as a technology for informal communication". *Communication of the ACM, **36(1)**,* 48-61.

Flammia, G. & Zue, V. (1995). "Empirical results of dialogue coding". In *4th European Conference on Speech Technology and Communication, EUROSPEECH 95*, 1905-1968.

Fleming, J.H., Darley, J. M., Hilton, J. L. & Kojetin, B. A. (1990). "Multiple audience problem: A strategic communication perspective on social perception". *Journal of Personality and Social Psychology, **58**,* 593-609.

Forston, R. F. & Larson, C. V. (1968). "The dynamics of space: An experimental study in proxemic behaviour among Latin Americans and North Americans". *Journal of Communication, **18**,* 109-116.

Furnace, J., Hamilton, N. M., Helms, P. & Duguid, K. (1996). "Medical teaching at a peripheral site by videoconferencing". *Medical Education, **30**,* 215-220.

Fussell, S. R. & Benimoff, I. (1995). Social and cognitive processes in interpersonal communication: Implications for advanced telecommunications technologies. *Human Factors, **37**,* 228-250.

Gale, S. (1990). "Human aspects of interactive multimedia communication." *Interacting with Computers, **2(2)**,* 175-189.

Gornall, L., Pengelly, S., & Shearn, D. (1999) "'Videoconferencing in the Valleys' A case study of the 'ALPs' Project". *Report for the JTAP (JISC Technology Applications Programme) Information Security Forum.* URL: http://www.jtap.ac.uk/reports/htm/jtap-035.html (last accessed 07/09/2000).

Green & Williges, R. C. (1995). "Evaluation of alternative media used with a groupware editor in a simulated telecommunications environment." *Human Factors, **37**,* 283-289.

Gregory, R. L. (1998). *Eye and Brain: The Psychology of Seeing* – 5th ed. Oxford University Press. Oxford.

Grigsby, J. & Sanders, J. H. (1998). "Telemedicine: where it is and where it's going." *Annals of Internal Medicine, **129**,* 123-127. Available online at http://www.acponline.org/journals/annals/15jul98/telemed.htm (last accessed 28/9/00)

Guetzkow, H. (1950). "Unitizing and categorizing problems in coding qualitative data". *Journal of Clinical Psychology*, ***6***, 47-58.

Hall, E. T. (1959). *The Silent Language.* Garden City, New York: Doubleday.

Hall, E. T. (1966). *The Hidden Dimension.* Garden City, NY: Doubleday

Harper, R. G., Wiens, A. N. & Matarazzo, J. D. (1978). *Non-verbal Communication: The State of the Art.* Wiley. Chichester.

Harrison, B., Ishii, H. & Chignell, M. (1994). "An empirical study on orientation of shared workspaces and interpersonal spaces in video-mediated communication." [WWW document]. URL http://www.dgp.toronto.edu/OTP/papers/video.mediated.collaboration/ishii.html (visited 1998, 04 November).

Heath, C. & Luff, P. (1991). "Disembodied conduct: communication through video in a multimedia environment." In *Proceedings of CHI'91: Human Factors in Computing Systems, reaching through technology.* Robertson, S.P.; Olson, G.M.; and Olson, J.S. (Eds.) pp99-103. New York, NY: ACM Press.

Heath, C., Luff, P. & Sellen, A. (1994). "Reconfiguring Media Space". In Emmot, S. & Travis, D (Eds.), *Proceedings of the International Symposium on POTS to PANS: User issues in the multimedia revolution from Plain Old Telephony Services to Pictures And Network Services.* Hintlesham Hall, Suffolk: British Telecom.

Hewson, T. J., Allison, G. & Clarke, A. M. (1995). "Providing Future Telecommunication Services to Naive Users." *Lecture Notes in Computing Science, **998**,* 92-103.

Hollier, M. P. & Voelcker, R. (1997). "Towards a multimodal perceptual model." *BT Technological Journal, **15**,* 162-171.

Hughes & Goldman, (1978). *Perceptual and Motor Skills, **46**,* 579-584.

Hughes, J., Kristofferson, S., O'Brien, J., & Rouncefield, M. (1995). "The organisational politics of meetings and their technology - two case studies of video supported communication" Paper presented at the first IFIP WG 8.6 Conference on Adoption and Diffusion of IT, Oslo October 1995. Selected for publication in Kautz, K and Pries-Heje,J. (eds) (1996) *Diffusion and Adoption of Information Technology*. London. Chapman & Hall.

Isaacs, E. & Tang, J. (1997). "Studying Video-based collaboration in context: From small workgroups to large organisations." In *Video-mediated Communication*, Finn, K., Sellen, A., Wilbur, S. (Eds.) 1997. Mahwah, NJ. Lawrence Erlbaum.

Isaacs, E. & Tang, J. (1993). "What video can and can't do for collaboration: A case study." In *Proceedings of the ACM Multimedia 93 Conference*. Anaheim, CA.

Isaacs, E. A. & Clark, H. H. (1987). "References in conversations between experts and novices." *Journal of Experimental Psychology: General, **116**,* 26-37.

Isaacs, E. A., & Tang, J. C. (1994). "What video can and cannot do for collaboration: A case study." *Multimedia Systems, 2,* 63-73.

Jameson, D. G., Hobsley, M., O'Hanlon, P. & Buckton, S. (1996). "Real-time interactivity on the superJANET network." *Interacting with Computers, **8**,* 285-295.

Johansen, R. (1984). *Teleconferencing and Beyond.* New York, NY: McGraw-Hill.

Johnson, C. F. & Dabbs, J.M. (1976). "Self-disclosure in dyads as a function of distance and subject-experimenter relationship." *Sociometry, **39***, 257-263.

Kendon, A. (1967). "Some functions of gaze direction in social interaction." *Acta Psychologica, **32**,* 1-25.

Kendon, A. (1990). *Conducting Interaction: patterns of behavior in focused encounters - (studies in interactional sociolinguistics; 7).* Cambridge University Press.

Kerckhoff, A.C. (1974). "The social context of interpersonal attraction." In T.L. Huston (Ed.), *Foundations of Interpersonal Attraction.* New York: Academic Press.

Kerlinger, F. N. (1964). *Foundations of behavioral research.* New York: Holt.

Kleinke, C. L. (1986). "Gaze and eye contact: A research review". *Psychological Bulletin, **100**,* 78-100.

Knowles and Dillon (1996). "Distance education and the continuing professional education of architects: the role of satellite videoconferencing and learning." *Journal of Architectural Planning Research, **11**,* 1-12.

Kowtko, J. C., Isard, S. D., Doherty-Sneddon G. M., (1991). "Conversational games within dialogue". *Workshop on Discourse Coherence*, University of Edinburgh, 1991.

Kowtko, J., Isard, S., & Doherty-Sneddon, G. (1992). "Conversational games within dialogue." Research paper HCRC/RP-31, Human Communications Research Centre, University of Edinburgh.

Krauss, R. M. & Fussell, S. R. (1990). "Mutual knowledge and communicative effectiveness." In J. Galegher, R. E. Kraut & C. Egido (Eds.). *Intellectual Teamwork: Social and Technical Bases of Collaborative Work* (pp111-145). Hillsdale, NJ: Erlbaum.

Krauss, R. M. & Fussell, S. R. (1996). "Social psychology models of interpersonal communication". In E. T. Higgins and A. W. Kruglansji (Eds.) *Social Psychology, Handbook of Basic Principles* (pp. 655-701). New York: Guildford Press.

Kraut, R. E., Miller, M. D. & Siegel, J. (1996). "Collaboration in performance of physical tasks: Effects on outcomes and communication." In *Proceedings of CSCW'96*, Cambridge MA, USA. pp 57-66.

Kristoffersen, S. & Rodden, T. (1996). "Working by walking around - requirements of flexible interaction management in video-supported collaborative work." In Sasse, M. A.; Cunningham, R. J. & Winder, R. L. (Eds). *People and Computers XI: Proceedings of HCI '96,* pp 315-329.

Latane, B. (1981). "The psychology of social impact." *American Psychologist, **36**,* 343-356.

Leary, M. R. (1983). *Understanding Social Anxiety: Social, Personality and Clinical Perspectives*. Beverly Hills, CA: Sage.

Leather, D. (1978). *Nonverbal Communication Systems*. Boston: Allyn and Bacon. Inc.

Lombard, M. & Ditton, T. B. (1997). "At the heart of it all: The concept of presence". *Journal of Computer-Mediated Communication,* [On-line], *3*(2). Available: http://207.201.161.120/jcmc/vol3/issue2/lombard.html

MacDonald, A. M. (Ed.). (1972). *Chambers Twentieth Century Dictionary*. Edinburgh. W and R Chambers. New Edition.

Mantei, M., Baecker, R., Sellen, A., Buxton, W., Milligan, T. & Wellman, B. (1991). "Experiences in the use of a media space." In *Proceedings of CHI'91 Human Factors in Computing Systems,* pp203-209, New York, NY: ACM Press.

McAndrew, P., Foubister, S.P. & Mayes, T. (1996). "Videoconferencing in a language learning application." *Interacting with Computers, **8**,* 207-217.

McCornack, S. A. & Levine, T. (1990). "When lovers become leery: The relationship between suspicion and accuracy in detecting deception." *Communication Monographs, **57**,* 219-230.

McCornack, S. A., & Parks, M. R. (1986). "Deception detection and relationship development: The other side of trust." In M. L. McLaughlin (Ed.), *Communication Yearbook 9* (pp. 337-389). Beverly Hills, CA; Sage. Cited in Millar & Millar (1995).

Mehrabian, A. (1971). "Verbal and nonverbal interactions of strangers in a waiting situation." *Journal of Experimental Research in Personality, **5**,* 127-138.

Mehrabian, A. (1972). *Nonverbal Communication*. Chicago: Aldine, Atherton.

Mehrabian, A. & Williams, M. (1969). "Nonverbal concomitants of perceived and intended persuasiveness". *Journal of Personality and Social Psychology, **13**.* 37-58.

Millar, M., & Millar, K. (1995). "Detection of deception in familiar and unfamiliar persons: the effect of information restriction." *Journal of Nonverbal Behaviour, **19**,* 69-84.

Milner, A. D. & Goodale, M. A. (1995). *The visual brain in action.* Oxford University Press. Oxford.

Monk, A., McCarthy, J., Watts, L. & Daly-Jones, O. (1994). "Measures of process." In M. Macleod & D. Murray (Eds.), *Evaluation for CSCW.* Berlin: Springer Verlag.

Mon-Williams, M. & Tresilian, J. R. (1999). "The size-distance paradox is a cognitive phenomenon." *Experimental Brain Research, **126**,* 578-582.

Moon, Y. (1998). "The effects of distance in local versus remote human-computer interaction." In In Karat, C-M., Lund, A., Coutaz, J. & Karat, J. (Eds.) *Proceedings of CHI'98: Human Factors in Computing Systems.* Los Angeles, CA, USA. pp 103-109.

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.

Morley, I. E. & Stephenson, G. M. (1970). "Formality in experimental negotiations: a validation study." *British Journal of Psychology, **60**,* 543-545.

Muhlbach, L., Bocker, M & Prussog, A. (1995). "Telepresence in videocommunications - a study on stereoscopy and individual eye contact." *Human Factors, **37**,* 290-305.

Nahemow, L. & Lawton, M.P. (1975). "Similarity and propinquity in friendship formation." *Journal of Personality and Social Psychology, **32**,* 205-213.

Nakazwa, K. (1993). "Proposal of a new eye contact method for teleconferences." *IEICE Transaction on Communication, **J76**,* 1 (January 1993), 618-625.

Nardi, B., Kuchinsky, A., Whittaker, S., Leichner, R., & Schwarz, H. (1996). "Video-as-data: Technical and social aspects of a collaborative multimedia application." *CSCW, **4**,* 73-100. Reprinted in *Video-mediated Communication*, Finn, K., Sellen, A., Wilbur, S. (Eds.) 1997. Mahwah, NJ. Lawrence Erlbaum.

Nardi, B., Schwarz, 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,* 327-341. New York: ACM.

Nass, C., Kim, E-Y. & Lee, E-J. (1998). "When my face is the interface: an experimental comparison of interacting with one's own face or someone else's face." In Karat, C-M., Lund, A., Coutaz, J. & Karat, J. (Eds.) *Proceedings of CHI'98: Human Factors in Computing Systems.* Los Angeles, CA, USA. pp. 148-154.

Newcomb, T. M. (1961). *The Acquaintence Process.* New York: Holt, Rinehart, and Winston.

Nickelson, D. W. (1996). "Behavioral telehealth: Emerging practice, research, and policy opportunities." *Behavioral Sciences and the Law, **14**,* 443-457.

Noesjirwan, J. (1977). "Contrasting cultural patterns of interpersonal closeness in doctors' waiting rooms in Sydney and Jakarta." *Journal of Cross-Cultural Psychology, **8**,* 357-368.

O'Connail, B. (1994). "A review of video-mediated communication evaluations." In *Proceedings of Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implications: position papers*, *CSCW*. Chapel Hill, NC. Oct 1994.

O'Connail, 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.

O'Donnell, C. P. & Draper, S. (1997, 29 May). "A nod is as good as a wink...or is it? A critical review of social psychology involved in video conferencing" [WWW document]. URL http://www.psy.gla.ac.uk/~steve/crvid.html (visited 1998, 29 October).

O'Malley, C. & Langton, S. (1994a). "Video-mediated communication: Strategies for research" In *Proceedings of Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implications: position papers, CSCW '94*, Chapel Hill, NC. Oct 1994.

O'Malley, C. & Langton, S. (1994b). "Making the case for face-to-face interaction: Empirical studies of video-mediated communication." In S. Emmott & D. Travis (Eds.), *Proceedings of the International Symposium on POTS to PANS: User issues in the multimedia revolution from Plain Old Telephony Services to Pictures And Network Services.* Hintlesham Hall, Suffolk: British Telecom.

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**,* 177-192.

Okada, K., Maeda, F., Ichikawaa, Y. & Matsushita, Y. (1994). "Multiparty videoconferencing at virtual social distance: MAJIC Design." In *Proceedings of CSCW'94,* Chapel Hill, NC, USA, ACM, pp 385-393.

Olson, J. S., Olson, G. M. & Meader, D. K. (1994). "What mix of video and audio is useful for remote real-time work?" In *Proceedings of Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implications: position papers*, CSCW, Oct 1994.

O'Shea, R. P., Govan, D. G., & Sekuler, R. (1997). "Blur and contrast as pictorial depth cues." *Perception, **26**,* 599-612.

Patterson, M. L. (1968). "Spatial factors in human interaction". *Human Relations, 21*, 351-361.

Patterson, M. L. (1977). "Interpersonal distance, affect, and equilibrium theory." *Journal of Social Psychology, 101,* 205-214.

Patterson, M. L. (1991). "A functional approach to nonverbal exchange". In Feldman, R. S. & Rime, B. (Eds.) *Fundamentals of nonverbal behaviour* (pp 458-495). Cambridge University Press.

Pavitt, C. (under review). *Philosophy of Science and Communication Theory.* Thousand Oaks, CA: Sage. URL: http://www.udel.edu/communication/COMM356/pavitt/appendix.htm

Poole, M. S., & Folger, J. P. (1981a). "Modes of observation and the validation of interaction analysis schemes." *Small Group Behavior*, *12*, 477-493.

Poyatos, F. (1975). "Cross cultural study of paralinguistic "alternants" in face-to-face interaction." In Kendon, A., Harris, R.M. & Key, M.R. (Eds.) *Organization of Behaviour in Face-to-Face Interaction* Mouton & Co. The Hague. Paris. Pp 285-314.

Ramsey, J., Barabesi, A. & Preece, J. (1996). "Informal communication is about sharing objects and media." *Interacting with Computers, 8,* 277-283.

Randall, D., J.A. Hughes and D. Shapiro (1994). "Using Ethnography to Inform Systems Design." *Journal of Intelligent Systems* (UK) 4(1-2): 9-28.

Reeves, B. & Nass, C. I. (1996). *The Media Equation: how people treat computers, televisions, and new media like real people and places*. Stanford, CA: Center for the Study of Language and Information. Cambridge University Press.

Reeves, B. R., Lombard, M. & Melwani, G. (1992). "Faces on the screen: pictures or natural experience?" Manuscript submitted to *Human Communication Research* (November, 1992). Cited in Nass, Kim & Lee (1998).

Reisberg, D., McLean, J. & Goldfield, A. (1987). "Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli." In R.

Campbell & B. Dodds (Eds.) *Hearing by Eye: The Psychology of Lip-reading.* Hillsdale, NJ. Lawrence Erlbaum Associates.

Rimmel, A. N., Hollier, M. P. & Voelcker, R. M. (1998). "The influence of cross-modal interaction on audio-visual speech quality perception." *Presented at the 105th AES Convention, San Francisco, CA.*

Rose, D., Furner, S., Hall, A., Montgomery, K., Katsavras, E., & Clarke, P. (1999). "Videoconferencing for speech and language therapy in schools." *BT Technology Journal, **18*** (1). Martlesham, Suffolk, England: British Telecommunications, plc. Retrieved April 24, 2000 from the World Wide Web: http://www.bt.com/bttj/vol18no1/today/papers/d_rose/abstract.htm

Rosenfeld, H. M. (1966). "Approval-seeking and approval-inducing functions of verbal and non-verbal responses." *Journal of Personality and Social Psychology, **4**,* 597-605.

Rubin, Z. (1970). "Measurement of romantic love". *Journal of Personality and Social Psychology, **16**,* 265-273.

Rudman, C. & Dykstra-Erickson, E. (1994). "Adaptation over time is required for effective use of video for group work". In *Proceedings of Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implications: position papers, CSCW '94*, Chapel Hill, NC. Oct 1994.

Rudman, C., Hertz, R., Marshall, C, & Dykstra-Erickson, E. (1997). "Channel overload as a driver for adoption of desktop video for distributed group work". In K. E. Finn, A. J. Sellen, & S. B. Wilbur (Eds.) *Video-Mediated Communication*, pp199-224, Mahwah, NJ: Lawrence Earlbaum Associates.

Rutter, D. R. & Stephenson, G. M. (1977). "The role of visual communication in synchronising conversation". *European Journal of Social Psychology, 2,* 29-37.

Rutter, D. R. & Stevenson, G. M. (1979). "The functions of looking: Effects of friendship on gaze". *British Journal of Social and Clinical Psychology, **38**,* 203-205.

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.

Rutter, D. R., Stevenson, G. M., Ayling, K., & White, P. A. (1978). "The timing of looks in dyadic conversation". *British Journal of Social and Clinical Psychology, **17**,* 17-21.

Rutter, D.R. (1987) *Communicating by Telephone.* Oxford: Pergamon Press, Elmsford, NY.

Sacks, H. (1972). "An initial investigation of the usability of conversational data for doing sociology". In: D. Sudnow (Ed.) *Studies in social interaction.* New York: Free Press: 31-74.

Sacks, H., Shegloff, E. A. & Jefferson, G. (1974). "A simplest systematics for the organisation of turn taking for conversation". *Language, **50**,* 696-735.

Sakamoto, N. (1992). *Stage Craft of Meetings.* Diamond Executive, Tokyo, 1992, 76-81.

Sasse, M. A. & Fentem, A. C. (1994?). "Interacting with multi-media, multi-user systems: observations on multi-media conferencing tools". Conference Paper to be presented at *ECCE-7 European Conference on Cognitive Ergonomics.*

Sawitsky, J. C. & Watson, M. J. (1975). "Patterns of proxemic behaviour among preschool children". *Representative Research in Social Psychology, **6**,* 109-113.

Scheflen, A. E. (1970). *Territoriality and communication.* Final Report NATO Symposium on Non-Verbal Communication, 35-38. Cited in Vine (1975).

Schegloff, E. & Sacks, H. (1973). "Opening up closings". *Semiotica, 7,* 289-327.

Schegloff, E., Jefferson, G. & Sacks, H. (1977). "The preference for self-correction in the organisation of repair in conversation." *Language, 53,* 361-382.

Schiano, D. J., Colston, H. L. & Pederson, E. R. (1994). "Psychological perspectives on the 'presence paradox'". In *Proceedings of the Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implications: position papers*, CSCW, Oct 1994. Chapel Hill, NC.

Schneider, P. (1999). "Literature review of telehealth video and audio psychotherapy". Chapter 2 of *A Comparison of Outcome Variables in Psychotherapy: Distance Technology versus Face-to-Face*. Doctoral Thesis defended in June 1999. University of Illinois. Available via World Wide Web at http://www.outreach.uiuc.edu/~p-schne/research/telehealth_literature_review.html (accessed 27/9/2000).

Schober, M. F. & Clark, H. H. (1989). "Understanding by addressees and overhearers". *Cognitive Psychology, **21**,* 211-232.

Segal, M.W. (1974). "Alphabet and attraction: an unobtrusive measure of the effect of propinquity in a field setting." *Journal of Personality and Social Psychology, **30**,* 654-657.

Sellen, A. (1994). " The effects of video-mediation on human interaction: Position paper for CSCW workshop". In *Proceedings of the Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implications: position papers, CSCW,* Oct 1994. Chapel Hill, NC.

Sellen, A. J. (1995). "Remote conversations: The effects of mediating talk with technology". *Human-Computer Interaction, **10**,* 401-444.

Sellen, A.J. (1992). "Speech patterns in video mediated conversations". In *Proceedings of CHI'92*, ACM, NY., pp 49-59.

Shannon, G., Bashshur, R., & Lovett, J. (1986). "Distance and the use of mental health services". *The Milbank Memorial Quarterly, **64**,* 302-330.

Short, J., Williams, E. & Christie, B. (1976). *The Social Psychology of Telecommunications*. London: Wiley.

Short, J. (1974). "Effects of medium of communication on experimental negotiation". *Human Relations, **27***, 225-243.

Sinclair, J. H. & Coulthard, R. M. (1975). *Towards an Analysis of Discourse: The English Used by Teachers and Pupils.* Oxford: Oxford University Press.

Solow, C., Weiss, R. J., Bergen, B. J., & Sanborn, C. J. (1971). "24 hour psychiatric consultation via TV". *American Journal of Pychiatry, **127***, 120-23.

Sommer, R. (1959). "Studies in personal space". *Sociometry, **22***, 337-348.

Sproull, L., Subramani, M., Kiesler, S., Walker, J.H., & Waters, K. (1996). "When the interface is a face". *Human-Computer Interaction, **11***, 97-124.

Steuer, J. (1992). "Defining virtual reality: Dimensions determining telepresence". *Journal of Communication, **42***, 73-93.

Stinson, L. & Ickes, W. (1992). "Empathic accuracy in the interactions of male friends versus male strangers". *Journal of Personality and Social Psychology, **62***, 787-797.

Storck, J. & Sproull, L. (1995). "Through a glass darkly – what do people learn in videoconferences". *Human Communication Research, 22*, 197-219.

Summerfield, Q. (1992). "Lip-reading and audio-visual speech perception". *Philosophical Transactions of the Royal Society of London, **B335***, 71-78.

Sundstrom, E. & Altman, I. (1976). "Interpersonal relationships and personal space: research review and theoretical model". *Human Ecology, **4***, 47-67.

Sussman, N. M. & Rosenfeld, H. M. (1982). "Influence of culture, language, and sex on conversational distance". *Journal of Personality and Social Psychology, **42***, 66-74.

Tang, J. C. & Isaacs, E. (1993). "Why do users like video? Studies of multimedia-supported collaboration". *Computer Supported Cooperative Work (CSCW), **1***, 163-196.

Tang, J. C., Isaacs, E. A. & Rua, M. (1994). "Supporting distributed groups with a montage of lightweight interactions". In *Proceedings of CSCW'94*, Chapel Hill, NC, USA, ACM, pp23-34.

Tapsell, L. (2000) "Using applied conversation analysis to teach novice dieticians history taking skills". *Human Studies 23*: 281-307

Taylor, T. & Cameron, D. (1987). *Analysing Conversation.* Oxford: Pergamon.

ten Have (1986) METHODOLOGICAL ISSUES IN CONVERSATION ANALYSIS[(1)] This is an electronic version of a paper published originally in the *Bulletin de Méthodologie Sociologique*, Nr. 27 (June): 23-51. Some references have been updated, but apart from that, the text is the one published. A first version was read in a session on 'Issues in Qualitative Data Interpretation', Research Committee 33, Logic and Methodology in Sociology, International Sociological Association, XIth World Congress of Sociology, New Delhi, August 1986. URL: http://www.pscw.uva.nl/emca/mica.htm#N_1_

Traum, D. R. & Hinkleman, E. H. (1992). "Conversation acts in task-oriented spoken dialogue." *Computational Intelligence, 8,* 575-599.

Vaksman, E. & Ellyson, S. L. (1979). "Visual and spatial behavior among US and foreign males: A cross-cultural test of equilibrium theory". *Paper presented at the Eastern Psychological Association Convention,* Philadelphia. Cited in Patterson (1991).

Veinott, E. S., Olson, J. S., Olson, G. M., & Fu, X. (1999). "Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other." In *Proceedings of Computer and Human Interaction (CHI) '99,* (Pittsburgh, PA, May 1999), ACM Press.

Veinott, E.S., Olson, J.S. & Olson, G.M. (1997). "Video matters! When communication ability is stressed, video helps." Short paper presented at CHI '97, Atlanta, Georgia, 22-27 March, 1997. [WWW document]. URL http://www.acm.org/turing/sigs/sigchi/chi97/proceedings/short-talk/esv.htm (visited 1998, 6 November).

Vine I. (1973). "Social spacing in animals and man." *Social Science Information, **12**,* 7-50.

Vine, I. (1975). "Territoriality and the spatial regulation of interaction." In Kendon, A., Harris, R. M. & Key, M. R. (Eds.) *Organization of Behavior in Face-to-Face Interaction* (pp. 357-388). Mouton Publishers. The Hague. Paris.

Vitkovich, M. & Barber, P. (1994). "Effect of video frame rate on subjects ability to shadow one of 2 competing verbal passages." *Journal of Speech and Hearing Research, **37**,* 1204-1210.

Walker, M. B. (1982). "Smooth transitions in conversational turn-taking: Implications for theory." *Journal of Psychology, **110**,* 31-37.

Watson, A. and Sasse, M. A. (1996). "Evaluating audio and video quality in lowcost multimedia conferencing systems." *Interacting with Computers, **8**,* 255-275.

Watson, A., & Sasse, M. A. (1997). "Multimedia conferencing via multicast: determining the quality of service required by the end user." *Proceedings of AVSPN '97.*

Watson, A., & Sasse, M. A. (1998). "Measured perceived quality of speech and video in multimedia conferencing applications." *Proceedings of ACM Multimedia'98,* Bristol, 12-16 Sept. 1998, pp55-60. ACM.

Watson, O. M. (1970). *Proxemic Behavior: A Cross-Cultural Study,* The Hague: Mouton.

Watson, O. M. & Graves, T. D. (1966). "Quantitative research in proxemic behaviour." *American Anthropologist, **68**,* 971-985.

Weeks, G. D. & Chapanis, A. (1976). "Cooperative versus conflictive problem-solving in three telecommunication modes." *Perceptual and Motor Skills, **42**,* 879-917.

Weisband, S. & Atwater, L. (1999). "Evaluating self and others in electronic and face-to-face groups." *Journal of Applied Psychology, **84,*** 632-639.

Whittaker, S. (1995). "Rethinking video as a technology for interpersonal communications: theory and design implications." *International Journal of Human-Computer Studies, **42,*** 501-529.

Whittaker, S. & O'Connail, B. (1997). "The role of vision in face-to-face and video-mediated communication." In K. E. Finn, A. J. Sellen, & S. B. Wilbur (Eds.) *Video-Mediated Communication*, pp 51-74, Mahwah, NJ: Lawrence Earlbaum Associates.

Whittaker, S., Geelhoed, E., & Robinson, E. (1993). "Shared workspaces: How do they work and when are they useful?" *International Journal of Man-Machines Studies, **39,*** 813-842.

Wichman, H. (1970). "Effects of isolation and communication on co-operation in a two-person game." *Journal of Personality and Social Psychology, **16,*** 114-120.

Williams, E. (1977). "Experimental comparisons of face-to-face and mediated communication." *Psychological Bulletin, **16,*** 963-976.

Williams, G. (1997). "Task conflict and language differences: Opportunities for videoconferencing?" In J. Hughes, W. Prinz, T. Rodden & K. Schmidt (Eds.) *Proceedings of the Fifth Europeann Conference on Computer Supported Cooperative Work,* 97-108. Kluwer Academic Publishers. Netherlands.

Witmer, B. G. & Sadowski, W. J. (1998). "Nonvisually guided locomotion to a previously viewed target in real and virtual environments." *Human Factors, **40,*** 478-488.

Wittson, C. L. & Benschoter, R. A. (1972). "Two-way television: Helping medical center reach-out." *American Journal of Psychiatry, **129,*** 624-27.

Yamaashi, K., Cooperstock, J. R., Narine, T. & Buxton, W. (1996). "Beating the limitations of camera-monitor mediated telepresence with extra eyes." *Proceedings of CHI '96* pp 50-57.

Zhang, Wolf, Darjaval & Touna, (1998). "Talking to customers on the web: A comparison of three voice alternatives." *Proceedings of CSCW'98,* 109-117.

Zuckerman, M., & Przewuzman, S. J. (1979). "Decoding and encoding facial expressions in preschool-age children." *Environmental Psychology and Nonverbal Behaviour, 3,* 147-163.

11 Appendices

11.1 Appendix A - Perceived Distance and Zoom - Questionnaire

Thank you for taking the time to complete this experiment. Please fill in this short questionnaire with reference to the task you have just completed.

Participant number -

The Image on the screen

1. Which setting would be most appropriate to place a call centre agent at if you were applying for a loan?

1	2	3	4	5
Closest				Farthest

2. Did you ever feel that the image on the screen felt imposing? (circle all that apply).

1	2	3	4	5
Closest				Farthest

3. If you were video conferencing with a stranger, which image would you like them to use?

1	2	3	4	5
Closest				Farthest

4. If you were video conferencing with a friend, which image would you like them to use?

1	2	3	4	5
Closest				Farthest

5. For each distance say whether you would be comfortable talking to the person (a stranger) or not.

	Uncomfortable - Too Near	Comfortable	Uncomfortable - Too Far
2.5 feet (Closest)			
4ft			
6ft			
9ft			
12ft (Furthest)			

The Face-to-Face Person

6. For each distance say whether you would be comfortable talking to the person (a stranger) or not.

	Uncomfortable - Too Near	Comfortable	Uncomfortable - Too Far
2.5 feet (Closest)			
4ft			
6ft			
9ft			
12ft (Furthest)			

11.2 Appendix B – Zoom Questionnaire Results

The results of the questionnaire were as follows. All are split by condition (far and close).

The Image on the screen

Q1. Which setting would be most appropriate to place a call centre agent at if you were applying for a loan?

Answer	Percentage of responses (LARGE WINDOW)	Percentage of responses (SMALL WINDOW)
1 (2.5 ft)	10%	10%
2 (4ft)	40%	20%
3 (6ft)	50%	30%
4 (9ft)	0%	30%
5 (12ft)	0%	10%

Q2. Did you ever feel that the image on the screen felt imposing? (circle all that apply).

Answer	Percentage of responses (LARGE WINDOW)	Percentage of responses (SMALL WINDOW)
1 (2.5 ft)	90%	90%
2 (4ft)	20%	40%
3 (6ft)	0%	0%
4 (9ft)	0%	0%
5 (12ft)	0%	0%

(10% in each condition felt no image was imposing)

Q3. If you were video conferencing with a stranger, which image would you like them to use?

Answer	Percentage of responses (LARGE WINDOW)	Percentage of responses (SMALL WINDOW)
1 (2.5 ft)	10%	10%
2 (4ft)	20%	10%
3 (6ft)	40%	30%
4 (9ft)	30%	30%
5 (12ft)	0%	20%

Q4. If you were video conferencing with a friend, which image would you like them to use?

Answer	Percentage of responses (LARGE WINDOW)	Percentage of responses (SMALL WINDOW)
1 (2.5 ft)	20%	10%
2 (4ft)	40%	30%
3 (6ft)	30%	50%
4 (9ft)	10%	10%
5 (12ft)	0%	0%

Difference between Question 4 and Question 3. What difference is there in the image participants felt they would like a friend to use compared to that of a stranger?

Condition	Place stranger closer	Place Stranger equal to friend	Place stranger further away
LARGE WINDOW	10%	40%	50%
SMALL WINDOW	0%	50%	50%
OVERALL	5%	45%	50%

Q5. For each distance say whether you would be comfortable talking to the person (a stranger) or not.

	LARGE WINDOW		SMALL WINDOW	
	Comfortable	Uncomfortable	Comfortable	Uncomfortable
1 (2.5 ft)	20%	80%	10%	90%
2 (4ft)	80%	20%	70%	30%
3 (6ft)	90%	10%	80%	20%
4 (9ft)	0%	100%	50%	50%
5 (12ft)	0%	100%	40%	60%

The Face-to-Face Person

Q6. For each distance say whether you would be comfortable talking to the person (a stranger) or not.

	LARGE WINDOW		SMALL WINDOW	
	Comfortable	Uncomfortable	Comfortable	Uncomfortable
1 (2.5 ft)	20%	80%	10%	90%
2 (4ft)	70%	30%	100%	0%
3 (6ft)	70%	30%	60%	40%
4 (9ft)	20%	80%	0%	100%
5 (12ft)	10%	90%	0%	100%

11.3 Appendix C – Questionnaire for Study 2

Financial Advice Across a Video Link

Thank you for your participation in this study. The following questionnaire is aimed at giving an insight into how participants felt about both the task and the technology used in completing it.

Please take your time to complete the questions - there are no right or wrong answers and as much information as possible is useful.

Section 1 is concerned with the technology that you just used.

Section 2 is concerned with the actual task you just completed.

Personal details

Age: _______ **Gender:** Male Female **Subject Number:**______

Section 1 - The Technology.

1. **How much experience have you with using computers?**

1	2	3	4	5
Very experienced	Experienced	Some experience	Inexperienced	Very inexperienced

2. **Have you ever used desktop videoconferencing before (like that used in the experiment)?**

   Yes No

3. **Have you ever done any banking by video conferencing?**

   Yes No

4. **Did you have any technological difficulties in completing the task?**

____
____
____
____
____
____
______

Section 2 - The Task.

Which choice did you make and why?

____
____
____
____
____
_____

1. How difficult did you find it to make your choice?

1	2	3	4	5
Very difficult	Difficult	Neither easy nor difficult	Easy	Very easy

2. How realistic did you find this task?

1	2	3	4	5
Very realistic	Realistic	Neither realistic nor unrealistic	Unrealistic	Very unrealistic

3. Would you be willing to accept financial advice in a bank using technology like this?

Yes No Unsure

4. How helpful did you find the financial advisor?

1	2	3	4	5
Very unhelpful	Unhelpful	Neither unhelpful nor helpful	Helpful	Very helpful

5. Did you find the technique of the financial advisor……. ?

1	2	3	4	5
Very pushy	Pushy	Neither pushy nor aloof	Aloof	Very Aloof

6. How much did you trust the financial advice given?

1	2	3	4	5
Trust a lot	Trust a little	Neither trust nor distrust	Distrust	Distrust a lot

7. Did you feel that the advice given by the financial advisor was totally independent (entirely reflecting your circumstances)?

1	2	3	4	5
Totally independent	Mostly independent	Neither independent nor biased.	Mostly biased by their company	Totally biased by their company

8. How much persuasion did you feel the financial advisor used to influence the decision you came to?

1	2	3	d/k
A lot	A little	None	Don't know

9. How likely do you think it is that you would have come to the same decisions if the information been available to you without the financial advisor?

1	2	3	4	5
Very unlikely	Unlikely	Neither unlikely nor likely	Likely	Very likely

10. How likely do you think it is that you would have come to the same decisions in a face to face situation?

1	2	3	4	5
Very unlikely	Unlikely	Neither unlikely nor likely	Likely	Very likely

11. How likely do you think it is that you would have come to the same decisions if the advice was given over the telephone?

1	2	3	4	5
Very unlikely	Unlikely	Neither unlikely nor likely	Likely	Very likely

12. How likeable was the financial advisor?

1	2	3	4	5
Very Likeable	Likeable	Neither likeable nor unlikeable	Not likeable	Not at all likeable

13. What did you think of the video image? (tick all that apply)

______ Person appeared too close

______ Person appeared too far away

______ Too small

______ Too large

______ Unclear

______ Slow/Jerky

______ Well positioned

______ Badly positioned

Other (please specify)____

14. Please rank the following methods of seeking financial advice in order of preference (1=most preferable, 5=least preferable).

______ Video

______ Face to Face

______ Telephone Banking

______ Postal

______ Leaflets/Application forms in a branch (no human contact)

15. How far would you be willing to travel in order to get face to face advice, if the video link was available in your local branch?

Zero 0-15min 15-30min 30min-1hour 1-2hour >2hours

16. Which financial services would you be willing to carry out via a system such as this one? (tick all that apply).

________ Apply for a loan

________ Apply for a mortgage

________ Discuss an overdraft

________ Open an account

________ Close an account

________ Customer complaints

________ Discuss financial advice

Other (please specify)
____

17. Have you any other comments on any part of the simulation?

____
____
____
____
____
______

Thank you for your participation in this investigation, please return this questionnaire to the experimenter.

11.4 Appendix D. - Results of the questionnaire by condition.

All figures are in percentages. Some data was missing in some questionnaires.

11.4.1 Personal details

1. Age (range = 17-32)

	Close	Far	Overall
Median Age (years)	20	18	19

2. Gender

	Close	Far	Overall
Female	4	11	15
Male	11	4	15

11.4.2 Section 1 - The Technology.

3. How much experience have you with using computers?

	Close	Far	Overall
Very experienced	7.1	13.3	10.0
Experienced	35.7	40.0	36.7
Some experience	50.0	46.7	46.7
Inexperienced	7.1	0.0	3.3
Very inexperienced	0.0	0.0	0.0

4. Have you ever used desktop videoconferencing before (like that used in the experiment)?

	Close	Far	Overall

Yes	14.3	13.3	13.8
No	85.7	86.7	86.2

5. Have you ever done any banking by video conferencing?

	Close	Far	Overall
Yes	0	0	0
No	100	100	100

6. Did you have any technological difficulties in completing the task?

	Close	Far	Overall
Yes	0	0	0
No	100	100	100

11.4.3 Section 2 - The Task.

7. Which choice did you make and why?

	Close	Far	Overall
PEP	28.6	26.7	26.7
Shares	21.4	6.7	13.3
TESSA	7.1	13.3	10.0
Bonus Account	42.9	53.3	46.7

8. How difficult did you find it to make your choice?

	Close	Far	Overall
Very difficult	0.0	0.0	0.0
Difficult	0.0	6.7	3.3
Neither easy nor difficult	57.1	46.7	50.0
Easy	42.9	40.0	40.0
Very easy	0.0	6.7	3.3

9. How realistic did you find this task?

	Close	Far	Overall
Very realistic	13.3	13.3	13.3
Realistic	53.3	53.3	53.3
Neither realistic nor unrealistic	20.0	26.7	23.3
Unrealistic	13.3	6.7	10
Very unrealistic	0.0	0.0	0

10. Would you be willing to accept financial advice in a bank using technology like this?

	Close	Far	Overall
Yes	53.3	73.3	63.3
No	6.7	13.3	10.0
Unsure	40.0	13.3	26.7

11. How helpful did you find the financial advisor?

	Close	Far	Overall
Very unhelpful	13.3	0.0	6.7

Unhelpful	0.0	6.7	3.3
Neither unhelpful nor helpful	13.3	6.7	10.0
Helpful	53.3	53.3	53.3
Very helpful	20.0	33.3	26.7

12. Did you find the technique of the financial advisor....... ?

	Close	Far	Overall
Very pushy	0.0	0.0	0.0
Pushy	6.7	6.7	6.7
Neither pushy nor aloof	86.7	93.3	90.0
Aloof	0.0	0.0	0.0
Very aloof	0.0	0.0	0.0

13. How much did you trust the financial advice given?

	Close	Far	Overall
Trust a lot	40.0	20.0	30.0
Trust a little	46.7	66.7	56.7
Neither trust nor distrust	6.7	6.7	6.7
Distrust	6.7	6.7	6.7
Distrust a lot	0.0	0.0	0.0

14. Did you feel that the advice given by the financial advisor was totally independent (entirely reflecting your circumstances)?

	Close	Far	Overall
Totally independent	6.7	6.7	6.7

Mostly independent	50.0	46.7	50.0
Neither independent nor biased.	36.7	40.0	36.7
Mostly biased by their company	6.7	6.7	6.7
Totally biased by their company	0.0	0.0	0.0

15. How much persuasion did you feel the financial advisor used to influence the decision you came to?

	Close	Far	Overall
A lot	6.7	0.0	3.3
A little	60.0	40.0	50.0
None	33.3	60.0	46.7
Don't know	0.0	0.0	0.0

16. How likely do you think it is that you would have come to the same decisions if the information been available to you without the financial advisor?

	Close	Far	Overall
Very unlikely	6.7	0.0	3.3
Unlikely	6.7	0.0	3.3
Neither unlikely nor likely	13.3	0.0	6.7
Likely	60	80.0	70.0
Very likely	13.3	20.0	16.7

17. How likely do you think it is that you would have come to the same decisions in a face to face situation?

	Close	Far	Overall

Very unlikely	0.0	0.0	0.0
Unlikely	6.7	6.7	6.7
Neither unlikely nor likely	6.7	20	13.3
Likely	66.7	60.0	63.3
Very likely	20.0	20.0	16.7

18. How likely do you think it is that you would have come to the same decisions if the advice was given over the telephone?

	Close	Far	Overall
Very unlikely	20.0	6.7	13.3
Unlikely	26.7	26.7	26.7
Neither unlikely nor likely	20.0	33.3	26.7
Likely	33.3	33.3	33.3
Very likely	0.0	0.0	0.0

19. How likeable was the financial advisor?

	Close	Far	Overall
Very Likeable	26.7	20.0	23.3
Likeable	46.7	46.7	46.7
Neither likeable nor unlikeable	26.7	33.3	30.0
Not likeable	0.0	0.0	0.0
Not at all likeable	0.0	0.0	0.0

20. What did you think of the video image? (tick all that apply)

	Close	Far	Overall

Person appeared too close	40.0	0.0	20.0
Person appeared too far away	0.0	6.7	3.3
Too small	6.7	6.7	6.7
Too large	13.3	0.0	6.7
Unclear	6.7	0.0	3.3
Slow/Jerky	20.0	6.7	13.3
Well positioned	66.7	60.0	63.3
Badly positioned	6.7	20.0	13.3
Other (please specify)	0.0	0.0	0.0

21. Please rank the following methods of seeking financial advice in order of preference (1=most preferable, 5=least preferable).

Mean score for each alternative

	Close	Far	Overall
Video	2.07	2.40	2.24
Face to Face	1.00	2.13	1.58
Telephone Banking	3.57	3.67	3.62
Postal	4.29	3.53	3.90
Leaflets/Application forms in a branch (no human contact)	4.07	3.27	3.66

Percentage of respondents ranking each option

	1	2	3	4	5
Video	6.9	72.4	13.8	3.4	3.4
Face to Face	75.9	10.3	0	6.9	6.9
Telephone Banking	0	0	58.6	20.7	20.7

Postal	6.9	6.9	6.9	48.3	31
Leaflets/Application forms	10.3	10.3	20.7	20.7	37.9

22. How far would you be willing to travel in order to get face to face advice, if the video link was available in your local branch?

	Close	Far	Overall
Zero	21.4	26.7	24.1
0-15min	14.3	26.7	20.7
15-30min	35.7	33.3	34.5
30min-1hour	21.4	13.3	17.2
1-2 hour	7.1	0.0	3.4
> 2 hours	0.0	0.0	0.0

23. Which financial services would you be willing to carry out via a system such as this one? (tick all that apply).

	Close	Far	Overall
Apply for a loan	64.3	40.0	51.7
Apply for a mortgage	35.7	6.7	20.7
Discuss an overdraft	71.4	66.7	69
Open an account	85.0	80.0	82.8
Close an account	100.0	73.3	86.2
Customer complaints	57.1	60.0	58.6
Discuss financial advice	85.7	73.3	79.3
Other (please specify)	0.0	0.0	0.0

24. Have you any other comments on any part of the simulation?

	Close	Far	Overall
Lack of eye contact	21.4	33.3	27.6
Other	0.0	0.0	0.0

11.5 Appendix E - Whiteboard Information Sheets

TESSA - Tax Exempt Special Savings Account

Tax free savings earning a good rate of interest. Anyone who pays income tax stands to benefit and opening and account is easy. Take a closer look to discover more.

A Choice of savings

You can choose whether you prefer to make regular savings or lump sum investments - or a combination of the two. Regular savings are made with monthly standing orders from as little as 20 pounds per month, lump sum investments can be made from 250 pounds and up to the annual limits.

Savings Limit

You can save up to 3,000 pounds in the first year, unless you have had a TESSA before, in which case you can re-invest all, or part of the capital into your TESSA, provided it's opened within six months of your first TESSA maturing.

Attractive Rates of Interest

Your money will earn a good rate of tax free interest, providing that you leave your savings untouched for five years. You can withdraw up to 80% of your interest and still benefit from the special tax priveleges

If you withdraw more than this amount then your TESSA will be closed, and all interest will be taxable, including any interest payments that have already been paid to you.

A Five Year Plan

The TESSA is a savings plan for a five year investment. There is no access to your capital during this period. A maximum of 9,000 pounds can be invested over the 5 years with with separate yearly limits

PEPS - Personal Equity Plan

A New Way to Invest

The PEP is a lower risk investment that provides a high regular income - completely free of personal tax. So, if you are an investor who wants a high regular income from your savings, or you are a cautious investor, a PEP may be just what you are looking for.

High Income

Our PEP is based on one of our established unit trusts. Over the past five years the PEP has outperformed major building society accounts significantly. Please bear in mind however, that the way an investment has performed in the past is not necessarily a to the income it will generate in the future.

No Tax to Pay

The income generated by your PEP is paid completely free of personal tax every three months. Alternatively you can reinvest the income and let your capital grow. All capital gains are also tax-free.

Lower Risk

When you invest in a High Income PEP, your money is pooled together with other investors and spread across a broad range of investments which are managed by a professional fund manager - so you don't need t make any investment decisions yourself.

Access to Your Money

PEPs are medium to long term investments, so you should ideally be prepared to leave your money in there for 5 years or more. However, it is possible to withdraw money from your PEP without losing your tax benefits, so long as you leave a minimum of £500 in the fund

Straightforward Charges

When you invest in a PEP there is a 4% initial charge. You will also be charged a fee of 1% of the value of your PEP every year. This fee will be deducted from the income your PEP generates, ensuring your capital is not eroded by charges.

Shares

Shares have the greatest potential for high returns on your investment but also carry the greatest risk.

Money in real terms

Most investment methods will give you a return on your money. Unfortunately, money is not always worth the same amount. This is due to INFLATION. Since 1945 the rise in the cost of living has been far greater than the rate of interest available to money on deposit. The real value of deposits has almost halved in that time.

The risk of deposits is that you might end up with less spending power than you started with.

Higher return investment

Savings likely to produce a higher return than inflation are usually based on stocks and shares. Since 1945, the average yearly rise in value of selection of British shares has been 6.65% greater than the rise in the cost of living

Risk

The risk you face with stock market investments is their short term fluctuations. To benefit from the potential for excellent returns, you need to take a long term view.

Timing is another issue. Buying at a high price and selling at a low price will lose you money. Alternatively buying at a low price does not guarantee the shares will go up in value.

11.6

Bonus Account

A high interest savings account offering flexibility and increasing rates of interest as the size of your investment increases. An interest bonus is available if you make less than 2 withdrawals per year.

Rates of Interest

The Bonus Account offers competitive rates of interest on a minimum initial investment of £5000. As your savings increase, the rate of interest you receive will also increase. The account has investment tiers of £5,000, £30,000, £50, 000 and £100,000.

In addition to the interest you earn, you can also earn an added interest bonus. If you make 1 or less withdrawal in a year then you will earn an interest bonus of 0.75% gross (0.60% net equiv.). The bonus rate is guaranteed above the basic rate of interest you receive on your account.

The rates of interest payable are variable. We may alter interest rates at any time if we consider it necessary to do so to reflect market conditions. When we do so then we will publish the changes widely.

When Interest is Paid

You have the choice to have interest paid either having your interest paid annually or monthly. Interest can be paid directly into your account or out to any other bank account you desire.

How Much Can I Invest?

To open a Bonus Account you may invest any amount from £5000. The maximum permitted investment is £250,000. You can make additions to your account at any time.

You can make withdrawals from your Bonus Account at any time. All we ask is for 120 days written notice of any withdrawal. Should you be unable to give that notice then you can have immediate access subject to 120 days interest penalty on the amount withdrawn.

Financial Advice Recall INTERVIEW

Subject Number:

	Questions	Answers
1.	What were the 4 options?	Bonus Account Shares TESSA PEPs
2.	Which options are tax free?	TESSA PEPs
3.	Which Options have charges on them?	PEPs
4.	Which option is the highest risk?	Shares
5.	Are immediate withdrawals allowed with the Bonus Account?	Yes
6.	What withdrawals are allowed with the TESSA?	Only 80% of the interest
7.	What does PEP stand for?	Personal Equity Plan
8.	What does TESSA stand for?	Tax Exempt Special Savings Account
9.	What are the conditions for earning a bonus with the bonus account?	One or less withdrawal over a year
10.	What are the conditions for withdrawing without penalty in the Bonus Account?	120 days WRITTEN notice
11.	What is the maximum initial investment in the first year of a TESSA?	£3,000

12.	What are the charges for the PEP?	4% initial charge 1% every year
13.	Which options are based on the stockmarket?	Shares PEPs
14.	Which options are risk free?	TESSA Bonus Account
15.	Who chooses where to invest your money in a PEP?	A Fund Manager
16.	What are the minimum and maximum investment levels of the Bonus Account?	£5 000 £250 000
17.	What is the interest rate on the Bonus Account?	No set rate - VARIES
18.	How do you close a TESSA?	It ends after 5 years **OR** Automatically closes if you withdraw
19.	What is the minimum amount of money to keep in a PEP?	£500
20.	What is inflation?	A rise in the cost of living
21.	How can you make money with shares?	Buy at one price, sell at a higher price
22.	How much above the rise in the cost have living have shares risen on average?	6.65%
23.	What is the bonus interest rate available on the Bonus Account after tax?	0.60%