The Effects of Computer Mediated Communication on the Processes of Communication and Collaboration.

Alison J S Newlands

Thesis submitted for the degree of Doctor of Philosophy in the Department of Psychology, Faculty of Social Science, University of Glasgow, February 1998.

Abstract

This thesis investigates the effects of computer-mediated communication upon collaborative problem solving. The results of three studies are presented, which explore the range of channels of communication afforded by two dissimilar forms of computer-mediated communication. The first study explores the effects of an interactive text-based form of computer-mediated communication, which provides users with a very restricted range of channels. Studies two and three examine the effects of collaborating in a video-mediated context, a technologically sophisticated communication system that affords an array of channels of communication more similar to face-to-face interactions. The effects of these communicative contexts are assessed using a multi-faceted approach. This method of evaluation is based upon analysis of task performance, the structure of the interactions, and measures of the process and content of communication.

The findings show that the novice users of these computer-mediated contexts can achieve effective communication and collaboration, but the ease and pace with which this is accomplished varies with communicative context. Users of the highly constrained text-based system initially performed less well on the collaborative tasks, but with experience adapted to the context in appropriate ways. Participants in the video-mediated conditions appeared to adjust quickly to this context. Subtle differences in the structure, process and content of their interactions show that they also had to make allowances for the restraints imposed by the technologically mediated context. These results are discussed within the frame-work of a collaborative model of communication.

Contents

Chapter 7 Conclusions of the Thesis

List of Appendices

List of Tables

List of Figures

Acknowledgements

First of all l would like to thank my supervisor, Professor Anne Anderson, for all her help, advice and support throughout my research, and the writing up of this thesis.

My thanks also go to Jim Mullin and Anne Marie Fleming for their technological assistance, without which the research in this thesis could not have been completed. I would also like to thank the Human Communications Research Centre for allowing me access to their Map Task Corpus, and to Gwyneth Doherty-Sneddon and George Smith for their assistance in coding some of the dialogues in Study 3. In addition, l would like to acknowledge the support and encouragement l have received from my husband, family, colleagues, and friends; who always seemed to assume that l would succeed in completing this work.

Declaration

I declare that this thesis is my own work carried out under normal terms of supervision.

Chapter 1. Introduction to Communication and Collaboration: Literature review

This thesis focuses on one aspect of communication, interactive communication between adults who are engaged in collaborative problem solving tasks. The research examines the impact of a range of computer mediated communicative contexts upon these interactions, exploring the effects of mediated communication upon the processes of communication and collaboration.

1.1 What is meant by Communication? Definitions and Approaches

Communication is one of the "primary means by which people affect one another" (Krauss and Fussell 1996, p. 2). Yet it is difficult to find a definition of the concept of communication that is universally acceptable. One reason for this is that the concept has been used in many disciplines, as diverse as cell biology, genetics, electrical engineering, computer science, sociology, linguistics and psychology. Even within these disciplines there is little agreement over the precise meaning of the concept.

1.1.1 Examples of Definitions of Communication

A general definition of communication is given in Chambers Twentieth Century Dictionary (Macdonald, 1972), which states that to communicate is to succeed in conveying one's meaning to others. Rather more specific definitions of communication can also be found. For example, Cherry (1957) suggests that communication can be defined as the physical signals with which one individual can influence the behaviour of another. Watzlawick, Beavis and Jackson (1967) simply state that human communication involves information exchange in context. Sperber and Wilson (1986) offer a definition of communication based upon an Information Processing model, they suggest that communication is "... a process involving two information-processing devices. One device modifies the physical environment of the other. As a result, the second device constructs representations similar to the representations already stored in the first device." Krauss and Fussell (1990) suggest a definition based upon what communication does, rather than what it is; communication is a "process by which knowledge that resides in one or more people comes to be represented in one or more others." (Krauss and Fussell 1990, p. 112). Clark (1996) bases his definition of communication upon the Latin roots of the word; to communicate is "to make common", or to make known amongst a group of people.

A common theme running through most of these definitions is that communication involves transmitting information from one system, or part of a system, to another (Krauss and Fussell, 1996). Interpersonal communication could therefore be defined simply as the transference of information from one person to another. This is insufficient, however, as the definition does not state what is meant by information and how it is conveyed.

1.1.2 The Distinction between Symbols and Signs

Information can be transmitted through a range of communicative acts, which convey information in very different ways. One way of conceptualising the differences between communicative acts is to consider whether they are forms of verbal or non-verbal behaviour, but frequently the distinction is drawn between the use of symbols or signs. A clear example (cited in Krauss and Fussell, 1996) of this distinction can be seen in the way that people respond to an embarrassing situation. One response would be to say "I'm embarrassed". Another way of responding would be to blush whilst saying nothing. Both responses convey the same information, but are produced and comprehended in quite different ways. Blushing is an involuntary action which occurs in particular situations, and can only be interpreted within the context of that situation. So 'signs' are involuntary acts that have causal significance. The use of symbols, in this case saying "I'm embarrassed", is an intentional act. The meaning conveyed by the symbols is determined by social convention.

Krauss and Fussell (1996) suggest that one of the main problems in establishing a general definition of the term 'communication', is to decide whether signs should be included or excluded from the definition. Theorists tend to adopt different stances on this issue. Some would include only symbolic behaviour in their concept of communication (for example, Ekman and Friesen, 1969; Wiener et al., 1972). Other theorists (such as Watzlawich et al., 1967) suggest that all types of behaviour can convey information, therefore both signs and symbolic behaviour should be considered as forms of communication. Both types of definitions are problematical. Definitions which exclude signs as a form of communication need to be able to distinguish between expressive and symbolic behaviour. This distinction is not always as easy to make as it might appear; signs and symbols may represent "two poles of a continuum rather than discrete categories" (Krauss and Fussell 1996, p. 15). On the other hand, if both signs and signals are included as forms of communication, the problem then is to decide the relative importance of the information being expressed.

The current research explores the effects of various communication contexts upon task-oriented dialogues, and includes analysis of the discourse. For example, in studies 2 and 3 spoken utterances generated in various Video-Mediated Contexts are examined in depth. According to Clark "utterances are prototypes of genuine signals." (Clark, 1985, p.187), and require coordination between the producer and the addressee (Grice, 1957; 1968). The emphasis in this thesis, therefore, is placed upon the signals used by people to communicate specific meaning. The signals include verbal communication (spoken utterances, or written messages), as well as non-verbal communication that is intended to convey specific meaning (such as head nods to indicate agreement, or acknowledgement).

In summary, Clark's (1996) definition of communication will be applied in the current work; communication is the process by which information is 'made common', or is made known amongst a group of people. In this thesis, the definition will be restricted to the signals used during communication. As the following section will illustrate, Clark's view of communication emphasises the point that communication is a social activity, which requires participants (speakers and hearers) to coordinate their actions; "people have to coordinate closely to make a piece of information common for them" (Clark, 1996, p. 153).

1.2 Models of Interpersonal Communication

Krauss and Fussell (1996) describe four theoretical perspectives, or classes of models, of interpersonal communication. The classes of models are labelled as Encoder/Decoder models, Intentionalist models, Perspective-taking models, and Dialogic models. This thesis will follow the Dialogic perspective of communication. There are two main reasons for choosing this perspective. Firstly, Dialogic models are based upon the premise that communication is a joint activity; conversational speech, rather than individual acts of speech production and comprehension, is taken as the basic model of communication (Krauss and Fussell, 1996). Secondly, one of the Dialogic models is supported by a substantial amount of research; this is Collaborative model of communication proposed by Clark and colleagues (for example, Clark and Wilkes-Gibbs, 1986; Isaacs and Clark, 1987; Clark and Schaefer, 1989; Clark and Brennan, 1991). Before discussing the Collaborative model in detail, some of the key differences between the Dialogic perspective and the other models mentioned by Krauss and Fussell (1996) will be outlined.

1.2.1 Dialogic models in Comparison to Other Perspectives

One of the key differences between Dialogic and other models of communication, is that Dialogic theorists view the meaning of an utterance to be 'socially situated'; that is, the meaning of a message can only be understood in the context in which it is written or spoken. Furthermore, because conversation is a joint activity, with participants collaborating to achieve conversational goals, the utterances of individuals cannot be analysed in isolation; they must be examined within the context of the ongoing interaction. The unit of analysis employed by Dialogic theorists thus differs from many other approaches. For example, researchers using an Intentionalist model of communication base their analysis upon the turns, or speech acts made by each individual. The unit of analysis used by Dialogic theorists is based upon the joint interaction of all the participants in a conversation.

Another distinguishing feature of Dialogic models is their emphasis upon conversational goals. Whilst other models assume that the prime aim of communication is to convey information, Dialogic theorists suggest that the "goal of communication is the achievement of *intersubjectivity*." (Krauss and Fussell, 1996, p. 79, original emphasis). The term 'intersubjectivity' means that participants reach similar interpretations of a message, or establish a "temporarily shared social world" (Rommetveit, 1974, p. 29,). Most models of communication assume that participants can achieve intersubjectivity; but they "differ in terms of how extensively they assume participants to rely upon intersubjectivity, and at what point in the communication process intersubjectivity is salient" (Schiffrin, 1994, p. 387). In Dialogic models of communication the attainment of mutual understanding is of paramount importance; it is the primary goal of communication, and information is exchanged in order to achieve this goal.

Krauss and Fussell (1996) state that the fullest description of a Dialogic perspective of communication can be found in the work of Clark and his colleagues, who give a detailed account of the Collaborative model of communication. This model, which has also been called the Constructionist Approach (McCarthy and Monk, 1994a), will form the theoretical basis of this thesis.

1.2.2 The Collaborative Model of Communication

The Collaborative model is based upon the premise that communication is a joint activity, "a collective activity of the first order" (Clark and Brennan, 1991, p. 128).

As in all collaborative activities, participants need to coordinate both what they are doing (the content of the activity) and the process of the activity. An example of this is given by Clark and Brennan (1991). When two people decide to play a piano duet by Mozart they need to ensure that they are both playing the same piece of music; this is 'coordination of content'. At the same time they must also coordinate how they play the music, they need to "synchronise their entrances and exits, coordinate how loudly to play forte and pianissimo, and otherwise adjust to each other's tempo and dynamics. This is coordination of process." (Clark and Brennan, 1991, p. 127).

Similar forms of collaboration are required during communication if participants are to reach mutual understanding of each others' utterances. Participants need to coordinate the process of communication. Speakers should ensure that addressees are attending to what is being said and are attempting to interpret the utterances. Addressees need to provide evidence that they are attending (Clark and Brennan, 1991). At the same time speakers and addressees collaborate on the content of the conversation, "working together in regular ways to produce evidence of a shared understanding." (Wilkes-Gibbs, 1995, p. 241). This process of establishing shared understanding, or 'common ground', is referred to in the Collaborative model of communication as the 'process of grounding' (Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1987; 1989; Isaacs and Clark, 1987)

1.2.3 Common Ground and the Process of Grounding

During conversation, speakers and addressees ensure that they have similar conceptions of the meaning of an utterance before they proceed with the conversation (Clark and Wilkes-Gibbs, 1986). In other words, they establish that they have attained a level of 'shared knowledge'. The concept of shared knowledge has been defined in various ways. For example, it has been referred to as 'common knowledge' (Lewis, 1969), or as 'mutual knowledge' (Schiffer, 1972). Schiffer defines mutual knowledge in the following manner:

A and B mutually know that $p =_{\text{def}}$

(1) A knows that *p*

(1') B knows that *p*

(2) A knows that B knows that *p*

(2') B knows that A knows that *p*

(3) A knows that B knows that A knows that *p*

(3') B knows that A knows that B know that *p*

etc., ad infinitum.

If Schiffer's definition of mutual knowledge was applied to everyday conversations, then participants would have an infinity of statements to check before they could be assured that they had understood each other. Clark and Marshall (1981) argue that this level of mutual understanding is not required in conversations; one sided definitions of mutual knowledge will suffice. This can be represented as follows: when two people (A and B) are conversing then mutual knowledge would be established for A if "A knows that A and B mutually know that *p*" (Clark and Marshall, 1981, p. 18).

Clark and Marshall (1981) suggest that each participant only requires half of the statements; A only requires the statements without the primes, whilst B just needs the statements with the primes. The Collaborative model accepts that perfect mutual understanding can never be fully achieved (Clark, 1985; Clark and Brennan, 1991); instead, participants establish a level of mutual understanding to a 'criterion sufficient for current purposes' (Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1989). The term 'mutual knowledge' appears in the earlier work of Clark and his colleagues, see for example Clark and Marshall (1981). However, because 'mutual knowledge' appeared to be open to mis-interpretation (Clark, 1992), in recent work the term 'mutual knowledge' has been referred to as 'common ground'. Common ground, as defined by Clark (1985), Clark and Schaefer (1987), Clark and Brennan (1991), refers to the mutual knowledge, beliefs, and assumptions of participants in a conversation.

Common ground is constantly updated during conversation by the process 'grounding'. This is a collective process which ensures that each participant has understood a previous utterance, to a level sufficient for their current purposes; this is termed the 'grounding criterion' (Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1989). The process of grounding occurs over a sequence of turns, and may involve the use of insertion sequences or repair techniques. Insertion sequences are a set of exchanges which are embedded within the ongoing exchanges (Jefferson, 1972; Schegloff, 1972). The example below demonstrates this point, and is taken from the London-Lund corpus (Svartvik and Quirk, 1980). The dash (-) indicates a short pause. and the asterisk (*) indicates the start of an insertion sequence.

Extract 1

Alan: Now - do you and your husband have a j- car

Barbara*: - a car?

Alan: Yes

Barbara: No -

In this example Barbara does not appear to have understood Alan's question, or is unsure she has heard it correctly; she indicates this by setting up an insertion sequence. Having obtained a positive reply to her question, Barbara now knows that she has the correctly understood Alan's question, and she can now supply the answer.

In conversation, when a person speaks an utterance it has to 'grounded' so that it becomes part of both participants' mutual knowledge and beliefs. In the Collaborative model of communication grounding occurs through the process of 'contributing' to the discourse (Clark, 1985; Clark and Wilkes-Gibbs, 1986; Clark and Schaefer, 1987). Contributing to a conversation requires the completion of two phases - presentation and acceptance phases. Clark and Schaefer (1987) define these two phases as follows:

Presentation phase

Person A presents an utterance to person B. A assumes that if B gives evidence of understanding the utterance, then A can believe that B understands what A meant by the utterance.

Acceptance Phase

Person B accepts the utterance present by A giving evidence that he believes he has understood what A meant by the utterance. B assumes that once A registers this evidence A will also believe that B understands.

This can be put more simply, as "A *presenting* an action for B to consider, and of B *accepting* that action as having been understood" (Clark and Schaefer, 1989, p. 151, original emphasis). If both of these phases are completed correctly, then A and B will both believe that they have reached the mutual belief that B has understood what A meant by the initial utterance. Both presentation and acceptance phases must be completed for A to have contributed to the discourse.

1.2.4 Evidence of Understanding

There is a range of ways in which an addressee can show - or provide evidence - that he or she has understood an utterance. Whilst both positive and negative evidence could be offered, the Collaborative model of communication suggests that "people ultimately seek positive evidence of understanding" (Clark and Brennan, 1991, p. 131). Clark and Schaefer (1989) suggest five ways in which evidence of sufficient grounding can be given by the addressee. The three most frequently used methods of giving positive evidence (or 'evidential devices') are 'continued attention', 'initiation of next relevant turn' and 'acknowledgements' (Clark and Brennan, 1991).

Continued attention

The addressee gives positive evidence that he has understood and accepted a contribution by simply continuing to pay attention to the contributor. One way of showing attention is through the use of eye gaze. Clark and Brennan (1991) suggest that this is the most basic way of giving positive evidence of grounding (sufficient for the current purpose), but this is also the weakest form of positive evidence that an addressee can display (Clark and Schaefer, 1989).

Initiation of the relevant next contribution

A stronger form of implicit evidence of sufficient grounding occurs when the addressee responds with a relevant utterance. The following example (cited in Krauss and Fussell, 1996, p. 83) demonstrates this form of acceptance:

Extract2

A: Last night John and I saw Schindler's List.

B: Did you like it? I saw it last week and was really moved.

Speaker B's appropriate response to A's contribution demonstrates that B has understood what A was referring to; that Schindler's List is a film. Other examples of relevant responses occur in adjacency pairs; such as question-answer pairs of utterances. If the speaker presents a contribution in the form of a question, then the next relevant turn would be an appropriate answer (Schegloff and Sacks, 1973). If the addressee responds with an appropriate answer to the presented question, he not only completes the adjacency pair but also gives evidence that he has understood the contributor's question. An inappropriate answer provides negative evidence of sufficientgrounding.

Acknowledgement.

A more explicit way of giving evidence of sufficient grounding is the use of acknowledgements, which are most usually indicated by use of backchannels, such as "I see" or "uh huh". Other forms of acknowledgement take the form of assessment comments such as 'gosh' or 'good God', or the use of non-verbal gestures such as head nod. (Clark and Brennan, 1991).

Two evidential devices, referred to as *demonstration* and *display*, are used less frequently (Clark and Schaefer, 1989). An addressee can demonstrate that he has understood all or part of a contributor's presentation, or he can display verbatim understanding of all or part of the contributor's presentation. Verbatim displays frequently occur when a contributor presents an addressee with material that needs to be memorised; for example, when discussing a telephone number or a postal address (Clark and Schaefer, 1987). In these circumstances the addressee often repeats verbatim the original utterance as a way of remembering the information, but this also demonstrates that he has understood the contribution.

1.2.5 Recursive Nature of the Acceptance Process

When the addressee initiates the acceptance phase in response to the presentation of an action by the contributor, the addressee's evidence is also a presentation which needs to be accepted (Clark and Schaefer, 1989). In other words, the acceptance process is recursive. At what point does the cycle of presentation and acceptance of a contribution stop? Clark and Schaefer suggest that the **Strength of Evidence Principle** provides a possible answer. This principle states that the strength of evidence required for accepting a presentation, in a succession of presentation and acceptance phases, will be reduced for each recursive cycle of the two phases involved in contributing to a discourse. That is, the strength of evidence given by the addressee will decline with each successive acceptance cycle. Eventually, usually after two or three cycles (Clark and Schaefer, 1989), the addressee will consider that the contribution has been sufficiently grounded, and he will offer one of the weakest forms of evidence; he may continue to pay attention or initiate the next relevant turn (Clark and Schaefer, 1989).

The number of recursive cycles required to establish mutual understanding may also be determined by other factors. For instance, the purpose of the conversation - whether it is mainly a social interaction or primarily an exchange of information - will determine the strength of evidence required for acceptance of contributions. Clark and Schaefer (1989) suggest that task related conversations may require stronger evidence of understanding than social dialogues. This view is supported by research into referential conversation by Cohen (1984) and Clark and Wilkes-Gibbs (1986).

1.3 Collaborative Model of the Process of Definite Reference

Clark and Wilkes-Gibbs (1986) examined the process of definite reference in conversation, starting from the premise that the act of reference is "a collaborative process requiring actions by both speakers and interlocutors." (Clark and Wilkes-Gibbs, 1986, p. 2). Their approach is at variance with previous traditional approaches to reference, which Clark and Wilkes-Gibbes termed 'literary models' of reference, in which the actions of the addressee play little part in the establishment of a referent.

In their 1986 paper Clark and Wilkes-Gibbs first offer evidence to support their Collaborative model of reference, and then explore the benefits of a collaborative model of reference. Two types of evidence to support the collaborative nature of definite reference are presented; first from examples of spontaneous conversational speech drawn from several corpora (such as Schegloff et al., 1977; Sacks and Schegloff, 1979; Cohen, 1984), and then from the results of their own laboratory experiment (Clark and Wilkes-Gibbs, 1986).

In the experiment, pairs of subjects completed a communication task which was originally devised by Krauss and Glucksberg (Krauss and Glucksberg, 1969; 1977). Participants sit on either side of table, separated by an opaque screen. Each subject has a series of cards in front of them, which show abstract figures called Tangrams. In front of one participant, entitled the 'director', the figures are laid out in a predetermined order, whilst the figures are arranged in a random sequence for the other participant (the 'matcher'). During the task the director assists the matcher to put her set of cards into the same order as the target sequence, each pair completing six trials of the task. The dialogues are tape recorded so that the acts reference can be analysed.

Clark and Wilkes-Gibbs noted that during the initial trials directors *described* the figures using indefinite references; the participants would then collaboratively identify the required Tangram. During this process participants established how they would refer to the Tangrams in the future. In subsequent trials the director would then *refer* to the figures using definite references. The following examples show how one director referred to the same card on the first and fourth trials of the task. The examples are taken from Clark and Wilkes-Gibbs (1986).

Extract 3

Trial 1. All right, the next one looks like a person who's ice skating, except they're sticking two arms out in front.

Extract 4

Trial 4. The next one's the ice skater.

As the trials progressed pairs of participants became more efficient at the task. That is, they took less time and used fewer words to complete the task. This was because they had established a mutually acceptable way of referring to the figures in the initial trials (Clark and Wilkes-Gibbs, 1986).

This process of establishing reference was of particular interest to Clark and Wilkes-Gibbs, because they found that it depended upon both participants taking an active role in the process. Examples from existing corpora and the Tangram experiment showed that speakers employed a range of communicative techniques to involve the addressee in the act of identifying the referent. For instance, contributors used noun phrases in a questioning manner (trial noun phrases) in order to establish if the addressee had understood the reference (these are referred to as 'try markers' by Sacks and Schegloff, 1979). This is demonstrated in the following example taken from Cohen (1984), and cited by Clark and Wilkes-Gibbs (1986). In extract 5, the dash (-) indicates a pause in the conversation.

Extract 5

S: Okay now, the small blue cap we talked about before?

J: Yeah

S: Put that over the hole on the side of that tube -

J: Yeah

S: - that is nearest to the top, or nearest to the red handle.

The extract shows that S draws J into the referential process, by asking a question. In this case, J feels that he has identified the object correctly and so responds with an affirmative, after which S continues with the referential process. Another method of mutually establishing a referent occurs when speakers offer information in a series of noun phrases (termed 'installment noun phrases' by Clark and Wilkes-Gibbs, 1986). In these cases the speaker pauses between each phrase to encourage the addressee to confirm that he/she has understood each utterance as it is presented. This is demonstrated in the example above, where S refers to the 'hole on the side of that tube' and then pauses for J to confirm that the reference has been understood. Installment noun phrases also occurred in Clark and Wilkes-Gibbs' Tangram experiment, an example from which is given below. In this extract, D indicates turns by the director and M turns by the matcher.

Extract 6

D: And the next one is the one with the triangle to the right -

M: Okay

D: With the square connected to it.

Clark and Wilkes-Gibbs claim that these examples show that the act of reference is a collaborative process in conversations; both participants have a role to play. The speaker initiates a noun phrase, which may be accepted or rejected by the addressee. Clark and Wilkes-Gibbs suggest that addressees can accept a referential noun phrase in two ways. They can allow the conversation to continue (*presuppose acceptance*), or acknowledge the presentation (*assert acceptance*). In either case, when addressees show that they have accepted the noun phrase, that they are confident that they have identified the referent sufficiently for current purposes, then this completes the phases of presentation and acceptance. The speaker has successfully added the referent to the pool of mutual knowledge that both participants share, and has contributed to the discourse.

If addressees do not accept the reference, or do not believe that they have sufficient common ground to uniquely identify the referent, then they may reject the presentation by the contributor. An alternative strategy is to initiate a side sequence of exchanges, during which the identity of the referent can be established. In either situation the original noun phrase will altered or 'refashioned' (Clark and Wilkes-Gibbs, 1986). This can be accomplished in a variety ways; for example, either the director or the matcher can make repairs or expansions to the noun phrase. An example of an extension by the matcher is given below, the example is taken from Clark and Wilkes-Gibbs (1986).

Extract 7

D: Um, third one is the guy reading with, holding his book to the left.

M: Okay, kind of standing up?

D: Yeah.

M: Okay.

In this extract, the matcher acknowledged what the director had offered so far; but then initiated a side sequence to offer a possible extension to the noun phrase. Once the side sequence had been completed (by D saying 'Yeah'), then the matcher showed acceptance of the refashioned noun phrase. On some occasions matchers rejected the director's noun phrase, and presented an alternative noun phrase instead. An example of a replacement by the matcher is demonstrated below.

Extract 8

D: Okay, and the next one is the person that looks like they're carrying something and it's sticking out to the left. It looks like a hat that's upside down.

M: The guy that's pointing to the left again?

D: Yeah, pointing to the left, that's it! (Laughs).

M: Okay.

In Extract 8, the matcher rejects the director's noun phrase, and presents a totally different description of the figure. Rather than expanding or repairing the noun phrase initiated by the director, a totally new noun phrase is accepted by both participants. Clark and Wilkes-Gibbs found that participants used this way of refashioning of the noun phrase less frequently than repairs or expansions.

Having demonstrated that referring is a collaborative process, Clark and Wilkes-Gibbs (1986) then suggest why such a collaborative method is preferred in conversations. They base their argument for a collaborative approach to reference upon the **Principle**

of Least Collaborative Effort which states that: "Speakers and addressees try to minimise *collaborative effort*, the work that both speakers and addressees do from the initiation of the referential process to its completion." (Clark and Wilkes-Gibbs, 1986, p 26. Original emphasis).

The Principle of Least Collaborative Effort appears to account for many of the features that occur in the initiation of a noun phrase, and in the devices used to refashion noun phrases. For instance, it has already been established that speakers tend to present noun phrases that do not adequately identify the referent. The speaker and the addressee then collaborate to refashion the noun phrase, and resolve the reference. When a speaker takes greater care in initiating the noun phrase, then less refashioning is required. In other words "there is a trade-off in effort between initiating the noun phrase and refashioning it." (Clark and Wilkes-Gibbs, 1986, p. 26).

There are several reasons why speakers encounter problems in producing noun phrases that do not require refashioning. Clark and Wilkes-Gibbs suggest three main reasons. Firstly, in most conversations speakers are trying to plan and produce utterance in a restricted amount of time. They may feel unable to produce a perfectly constructed noun phrase in the time allowed, and therefore initiate a less adequate reference which can then be repaired or expanded as required. Secondly, the speaker may be able to construct an ideal noun phrase; but it may be so complex that the addressee will not be able to understand it if it is presented in a single turn. The speaker may prefer to use installment noun phrases, giving the information in a piecemeal fashion. Thirdly, speakers are often unsure which form of noun phrase would be acceptable to the addressees, so they use a trial and error procedure; "they try out a description and leave it to the addressees to refashion if it isn't acceptable." (Clark and Wilkes-Gibbs, 1986, p. 27).

The devices used to refashion noun phrases can also be seen to minimise the amount of collaborative effort required in reference situations. For example, in the cases where repairs are deemed necessary, Schegloff, Jefferson and Sacks (1977) note that speakers prefer to repair their own utterances rather than do so at the prompting of the addressee. Clark and Wilkes-Gibbs suggest that the same preferences occur when noun phrases require expansion, or replacement, during the acceptance phase of collaborative referring. These preferences are beneficial in terms of collaborative effort; it takes less time and effort if the presenter of a noun phrase corrects him/herself than it would if the correction is made by the prompting of the addressee.

Clark and Wilkes-Gibbs have demonstrated that people collaborate during conversations to establish the identity of a referent; making sure that the referent becomes part of their common ground is a mutual responsibility in conversation. This is summarised in the **Principle of Mutual Responsibility**. This is defined as follows: "Participants in a conversation try to establish, roughly by the initiation of each new contribution, the mutual belief that the listeners have understood what the speaker meant in the last utterance to a criterion sufficient for current purposes" (Clark and Wilkes-Gibbs, 1986, p. 33).

The examples of the process of establishing definite reference given so far have all been taken from communicative settings where speakers and addresses are in face-to-face situations. Is the process similar in other forms of communication? What happens if interlocutors are not engaged in face-face communication? For example, in lectures, prepared speeches and television broadcasts, speakers are communicating with more than one addressee, and the participants may not share the same temporal or spatial location. In these communicative settings, Clark and Wilkes-Gibb suggest that a weaker version of the Principle of Mutual Responsibility will be evoked; this is called the **Principle of Distant Responsibility**, as speakers and addresses are distant from each other in time and/or space. This principle is defined as follows: "The speaker or writer tries to make sure, roughly by the initiation of each new contribution, that the addressees should have been able to understand his meaning in the last utterance to a criterion sufficient for current purpose." (Clark and Wilkes-Gibbs, 1986, p. 35).

Exactly how this is achieved will depend on a range of factors, such as modality (whether the discourse is spoken or written), or the degree of formality of the setting (spontaneous or formal). Clark and Wilkes-Gibbs suggest that in spontaneous monologues speakers monitor what they say (just as they would in everyday conversations), so it is likely that they will refashion noun phrases in much the same manner as occurs in dialogue. However, they will do so in the absence of feedback from listeners. In written communication, writers usually have considerably more time to plan and edit; they should, therefore, produce definite references that require little repair or expansion. Writers should take more responsibility for establishing reference than speakers. Clark and Wilkes-Gibbs (1986) support their ideas by referring to research by Cohen (1984), who notes that participants in a referential task employed a more collaborative method of establishing referents when conversing over a telephone link than over linked tele-typewriters.

In summary, the Collaborative Theory of definite reference in conversations states that speakers and addressees are mutually responsible for establishing that the addressee has understood the reference, to a criterion sufficient for current purposes. The methods used to establish mutual knowledge, to make the referent part of the participants' common ground, can be explained by the Principle of Least Collaborative Effort. Clark and Wilkes-Gibbs (1986) demonstrate the application of this principle in both the presentation and acceptance phases of the process of definite reference.

1.4 Research Supporting the Collaborative Model of Communication

Support for the Collaborative model of language use comes from a line of experiments. Examples of this research will be given in the next section of this chapter.

The research of Isaacs and Clark (1987) follows on closely from the work of Clark and Wilkes-Gibbs (1986). This research focused on acts of reference in conversations where participants had varying amounts of knowledge about the subject under discussion. Differences in expertise occur in many communicative situations - for instance, when people seek medical or legal advice - but they also occur in everyday conversations. In many everyday instances the discrepancy between expert and novice may be barely discernible. Isaacs and Clark (1987) propose that people adjust to different levels of expertise; they 'accommodate to discrepancies in expertise', whilst they collaboratively establish the identity of the referent.

To investigate how the processes of accommodation occur in conversations, Isaacs and Clark devised a referential task based upon the Tangram task pioneered by Krauss and Glucksberg (1969; 1977). In this new referential task two sets of postcards, showing well known scenes of New York city, were used instead of the Tangram figures. Half of the participants in this experiment were familiar with New York City, whilst the others were unfamiliar with the city. Subjects were, therefore, either experts or novices in this realm of knowledge. The participants worked in pairs, with subjects having equal or differing levels of expertise. Participants were ignorant of the level of expertise of their partners before a trial commenced. The following extract shows a typical contribution in which a postcard is identified. The example is taken from Isaacs and Clark (1987):

Extract 9

turn 1 Director. Tenth is the Cidicorp [sic], Citicorp Building?

turn 2 Matcher. Is that with the slanted top?

turn 3 Director. Yes

turn 4 Matcher. Okay.

This sequence of exchanges resembles the process of definite reference described by Clark and Wilkes-Gibbs (1986). The matcher sets up a side sequence (turn 2) and expands the original reference offered in turn 1 by the director; this is accepted by the director (in turn 3) and the acknowledgement by the matcher completes the cycle of presentation and acceptance.

Isaacs and Clark examined the process of reference by categorising the way in which the director initially described each postcard, using the following categories: proper names, proper name plus a description, and description only. The percentage use of these forms of introduction by experts and novices, and how these changed over time were then examined. As expected, Isaacs and Clark (1987) found that experts (people familiar with New York City) used proper names significantly more frequently than novices (66% versus 32%). They also observed that experts used proper names more often if their partner was an expert rather than a novice. This occurred in spite of the fact that the majority participants did not explicitly ascertain their partner's level of knowledge (Isaacs and Clark, 1987). When participants had mixed levels of expertise, directors were found to accommodate to the matcher's knowledge of New York City; using more descriptions of the postcards and fewer proper names. Novice directors gradually incorporated proper names into their references as they acquired knowledge from expert matchers, but continued to use descriptive types of references if they were working with novice matchers.

The speed with which participants adjusted to their partner's level of expertise was particularly surprising. Analysis of the first trials of the task showed that "directors adjusted to their partner before they had arranged even four postcards." (Isaacs and

Clark, 1987, p. 31). To make the necessary adjustments, partners must have assessed each other's level of expertise very rapidly. Isaacs and Clark suggest four ways in which the level of knowledge of a partner could be assessed during the conversations. Firstly, expertise was demonstrated in the use of proper names to refer to buildings and scenes depicted in the postcards. Participants who were unfamiliar with New York City were naturally unable to use this form of reference. Secondly, experts tended to use definite references (such as, 'the fountain') rather than indefinite references; as they were familiar with the scene, they knew that the scene they were referring to could be uniquely identified by this feature. The third and fourth cues were inferred from the different perspectives taken by experts and novices. Whilst novices were restricted to referring to what could be seen in the picture, experts tended to refer to the place itself, where it was located, what it was used for, and mentioned features which were not conspicuous in the postcard (Isaacs and Clark, 1987).

The paper by Isaacs and Clark (1987) provides empirical support for the Collaborative model of reference. Both participants in a conversation are mutually responsible for ensuring that the act of reference is completed, to a criterion sufficient for their current purpose. The authors also demonstrate the processes by which participants assess, supply and acquire expertise whilst they converse, and how this assists them to collaboratively establish references more efficiently as the trials progress. These processes illustrate the way in which participants accommodate to each other's level of expertise, which appear to occur remarkably rapidly. This demonstrates the flexibility of communicators, and how they adjust to various communicative settings and environments; a point that will be referred to later in the thesis.

Schober and Clark (1989) investigated how understanding is achieved by listeners who are actively involved in a conversation (addressees) compared to listeners who simply overhear a conversation (overhearers). The study examined the different predictions made by collaborative and autonomous models of communication, concerning the level of understanding achieved by these two types of listeners. An autonomous model would predict that overhearers should reach a similar level of understanding as addressees, as long as they hear all of the conversation and come from similar cultural backgrounds. The Collaborative model predicts that overhearers will not achieve the same level of understanding as addressees, because overhearers do not take an active role in establishing the common ground (Schober and Clark, 1989).

The two experiments in this study were based upon the Tangram-matching task. In experiment 1 the conversations between directors and matchers were recorded, and then played back to overhearer matchers (henceforth overhearers) who attempted to position the Tangram figures in the correct order whilst listening to the audio tape. Overhearers completed the task under several different conditions. In the first experiment, some overhearers were given control of the pace of the interaction, being allowed to pause (but not rewind) the tape whenever they wished. Another group of overhearers heard only the recordings of the last three trials of the experiment. In the second experiment the overhearers were seated in the same room as the director and the matcher, each participant being screened-off from the others.

In both experiments, Schober and Clark found that the process of collaboration followed the pattern observed by Clark and Wilkes-Gibbs (1986). During the first trial the figures were described in detail by the director; but in succeeding trials the figures were referred to with definite descriptions (Schober and Clark, 1989). As in previous repeated reference tasks, the descriptions became briefer across trials, and the amount of linguistic output required to make the reference also declined (Clark and Wilkes-Gibbs, 1986; Isaacs and Clark, 1987; Schober and Clark, 1989).

The most important measure of communication and collaboration used in these studies was task accuracy, the percentage of figures placed in the correct order. The measures of accuracy from both experiments supported the collaborative view of understanding. Matchers performed significantly better than the overhearers who heard all six of the trials. This indicated that overhearers did not reach the same level of understanding as matchers; "... being witness to the buildup of common ground did not seem to provide all the necessary information for overhearers to understand their references as well as addressees." (Schober and Clark, 1989, p. 218). Overhearers who heard only the last three trials performed less accurately than overhearers who heard the complete set of trials; so, being able to overhear the director and matcher collaboratively agreeing upon the description of the figures was of some benefit. Overhearers who were allowed to pause the audio tape tended to do so, but this did not improve their accuracy rate (Schober and Clark, 1989).

Further support for the Collaborative Model was provided by the second experiment. First of all, the accuracy scores of the overhearers were again significantly poorer than the matchers, even though the overhearers were co-present with the other participants. Secondly, video recordings of the trials showed that overhearers changed their minds more often than the addressed matcher - overhearers swapped the figures around more often - which indicated that they were not as confident about the identity of the referents as the matchers were. Thirdly, Schober and Clark compared the time between placement of a card relative to 'completion time'; this time was defined as the point at which the director and matcher verbally acknowledged that they were ready to go on to the next figure. This analysis showed that overhearers frequently placed cards into the sorting frame either earlier or later than the matcher (Schober and Clark, 1989).

There were several consequences of this discrepancy between matchers' and overhearers' completion times. Overhearers who placed a card after completion of the reference were not prepared for the start of the next reference, this often resulted in the overhearer making an incorrect choice of the next figure. An additional problem for overhearers was the varying degree of communicative efficiency of some of the director/matcher pairs; Schober and Clark observed that efficient pairs worked at a greater pace, and tended to refer to figures from perspectives that were difficult for some overhearers to follow. These findings offer support for the collaborative view of understanding: "Listeners who participate in a conversational interaction go about understanding very differently from those who are excluded from it. It is because of these differences that addressees understand faster and more accurately than overhearers." (Schober and Clark, 1989, p. 228).

1.5 Further Support for the Collaborative Model of Communication

The Collaborative Model of communication and language processing has not only received support from the findings of Clark and his colleagues. Research by other psychologists has also produced evidence of the importance of collaborative processes in communication, and how mutual knowledge and understanding is achieved on a moment by moment basis. Two examples from this supportive research will be outlined in this section; first the research by Garrod and Anderson (1987) and then the findings of Anderson and Boyle (1994).

Garrod and Anderson (1987) used a computer maze game to examine the way in which people "coordinate their use and interpretation of language in a restricted context" (p. 181). During the maze game pairs of participants are seated in front of a visual display unit (VDU) in separate rooms. The maze is displayed on the VDU, and consists of a series of small boxes which are linked by paths. Participants can move their position markers along the paths, moving from one box to the next in any one turn until they reach their goal. Each participant can see only his own start point, current position, and goal. At random points in the maze the paths are blocked by gates, which can be opened if the other participant moves into a box which contains a switch.

To complete the maze game successfully, participants need to collaborate and be able to specify where in the maze they are located, so that they can direct each other towards the switch boxes. The participants therefore need to establish a mutually satisfactory way of conceptualising the maze, and how they will refer to specific locations within the maze (Garrod and Anderson, 1987). Analysis of the transcripts showed that participants used a range of descriptions to refer to the structural lay-out of the maze and their location within the maze. Four main categories of descriptions were found to occur; these were termed 'path', 'coordinate', 'line' and 'figural' (Garrod and Anderson, 1987).

One of the interesting findings of this analysis was that partners tended to be "very consistent in their choice of description within any section of the dialogue." (Garrod and Anderson, 1987, p. 189). So if one partner used a path description to designate his present location, then his partner employed a similar description in their response. Whilst participants could have explicitly negotiated how they were going to refer to spatial locations within the maze, this was found to occur infrequently (Garrod and Anderson, 1987). When it did occur, explicit negotiation tended to occur **after** participants had encountered referential problems. These negotiated description schemes were not adhered to for any great length of time (Garrod and Anderson, 1987).

From their analysis of the maze game dialogues, Garrod and Anderson found that participants solved the problem of coordinating their use of language in an interactional manner, which was more flexible and cost effective than explicit negotiation (Garrod and Anderson, 1987). Participants tended to be locally consistent; that is, the descriptions used by one person were mirrored in the next referential turn by their partner. Garrod and Anderson (1987) called this type of collaboration the 'input/output coordination' principle; speakers formulate their next utterance depending on the way in which they interpreted the most recent relevant utterance from the other conversant. If both participants adhere to this principle, then "the chances are that they will quickly establish a mutually satisfactory description scheme with the minimum of collaborative effort." (Garrod and Anderson, 1987, p. 207).

Conforming to the principle of 'input/output coordination' could be a useful solution to many referential problems faced in everyday conversations (Garrod and Anderson, 1987). However, the authors noted that the principle has some limitations. For example, there were occasions in the maze games where speakers made slight adjustments to the description schemes. These changes increased the amount of collaborative work required to identify the spatial location being referred to, often involving several additional cycles of presentation and acceptance by both speakers in order to resolve or clarify the reference. On the other hand, strict adherence to the 'input/output coordination' principle meant that the description schemes could become inflexible (Garrod and Anderson, 1987). Speakers appeared to overcome these problems by accepting (implicitly) that one of the speakers could modify the descriptive scheme, and that the other would accept this modification. In effect, one speaker becomes the leader and the other the follower (Garrod and Anderson, 1987). As long as only one speaker modifies the scheme, and the other participant adheres to the principle of 'input/output coordination' then the process of establishing mutual ground can proceed smoothly.

The findings by Garrod and Anderson demonstrate how people collaborate and coordinate their actions and use of language during a problem solving task; and they provide further empirical evidence for adopting a Collaborative model of communication. The recent findings by Anderson and Boyle (1994), Anderson, Clark and Mullin (1994; 1991) and Anderson (1995) also provide support for this model. The research examined the maintenance of coherence during task-oriented dialogues, exploring the way in which new information or entities are introduced into the discourse. If a dialogue is to be intelligible to both speakers and listeners, then new information must be introduced felicitously and become part of the common ground of the participants.

Anderson and colleagues examined the different forms of introduction that occurred during task-oriented dialogues, and the effects that these various forms of introduction had upon communicative success. The research examined the forms of introduction employed by both children (Anderson, Clark and Mullin, 1991a; 1994) and adults (Anderson and Boyle, 1994; Anderson, 1995). The findings from the adult dialogues will be reported here, as they have greater relevance to this thesis which examines adult conversations in different communicative contexts.

The task used in the research of Anderson and colleagues was The Map Task (Brown, Anderson, Yule and Shillcock, 1984). This is a collaborative problem solving task which involves one person (the Instruction Giver) giving directions to another (the Instruction Follower) about a route to follow on a simple, schematic map. The maps for each participant are similar, but not identical; some of the landmark features shown on one map do not appear on the other map. In these experiments, participants sit on either side of a table, but are unable to see each other's maps. Further details of this task will be given in chapter 3, as The Map Task forms the basis for the first and second experimental studies presented in this thesis.

The manner in which new information was introduced into the discourse was examined by categorising the way in which speakers first mentioned each of the landmarks shown on the maps. Anderson and colleagues found that two forms of introduction were observed; question and non-question forms of introductions.

Question forms of introduction occurred when a speaker mentioned a new landmark and at the same time asked a question to probe the listener's knowledge about the feature being introduced. In non-question forms of introduction the landmarks were mentioned as "part of the instruction-giving process and were not marked off by a question about the listener's knowledge." (Anderson and Boyle, 1994, p. 106). An analysis of the frequency of use of each category of introduction was computed. The results showed that, on average, Instruction Givers introduced more features than the Instruction Followers, and that they used more question than non-question introductions (Anderson and Boyle, 1994; Anderson, 1995).

It was also noted that the communicative success of pairs of participants varied considerably (Anderson and Boyle, 1994; Anderson, 1995). Some pairs achieved higher levels of task performance than others. To see if this could have been associated with the way in which new features were introduced into the discourse, the relationship between task performance and the use of question forms of introduction was examined. This analysis revealed that question forms of introduction were significantly correlated with communicative success. Participants who used question introductions, which explicitly checked their listener's state of knowledge about a landmark feature, achieved higher levels of communicative success.

This result supports the Collaborative model of communication proposed by Clark and colleagues. Question forms of introduction not only allow the speaker to check the listener's state of knowledge, they also offer the listener an opportunity to take an active role in establishing the identity of the new referent. According to the principles of 'mutual responsibility' and 'least collaborative effort', this interactive form of collaboration should facilitate effective communicative (Clark and Wilkes-Gibbs 1986). However, interactive forms of communication can only succeed if listeners respond in an informative and appropriate manner when new entities are introduced.

The way in which listeners responded to question and non-question forms of introduction was, therefore, also included in the analysis by Anderson and colleagues.

This analysis concentrated on the exchanges which occurred when a speaker introduced a feature which was not present on the listener's map. The responses by listeners were classified as follows: an informative response, if the listener explicitly told his/her partner that they did not have the feature referred to; an inadequately informative response, if the listener indicated that there was some problem in understanding a reference to an unknown feature; an uninformative response, when the listener gave no relevant verbal feedback, or inappropriate positive feed back (Anderson and Boyle, 1994; Anderson, 1995). The frequency of each type of response by both participants was calculated. The results of this analysis showed that the form of introduction used by the speakers influenced the informativeness of a listener's response; question forms of introduction nearly always elicited informative responses, whilst non-question forms frequently evoked uninformative or ambiguous responses. The following extracts (taken from Anderson and Boyle, 1994) illustrate typical exchanges which occurred when question and non-question forms of introduction were used during the Map Task. In these extracts the initials 'IG' and 'IF' indicate turns by the Instruction Giver and Instruction Follower, and an asterisk (*) indicates a feature which appears on only one of the maps and has not previously been mentioned.

Extract 10.

IG: Have you got a picnic site*?

IF: Nope where is it? *(Informative response)*

IG: It's about an inch it's below the site of the forest fire and about an inch

IF: An inch

IG: An inch and a half above the bottom of the page

Extract 11.

IG: Well I want you to draw - uh I've got pebbled shore and above that I've got coconut palm* right

IF: Mhm *(Uninformative response)*

In extract 10, the Instruction Giver uses a question to introduce a new feature (the picnic site) and in so doing checks if the Instruction Follower knows which feature he is referring to. This form of introduction elicited an informative response from the Instruction Follower; the participants then collaboratively establish the whereabouts of the picnic site. In this manner the new feature becomes part of the common ground of the discourse. In Extract 11 a new feature (the coconut palm) is mentioned by the Instruction Giver, as part of his on-going instructions about the task. The Instruction Follower's response is uninformative; it is an ambiguous response which does not inform the Instruction Giver that the Instruction Follower has no knowledge of this feature (the coconut palm does not appear on this Instruction Follower's map). Anderson and Boyle (1994) noted that a surprisingly large number of non-question introductions elicited this type of uninformative response, by both the Instruction Follower and the Instruction Givers. As these responses did not inform the contributor of the listener's lack of knowledge about the features that had just been mentioned, this made it very difficult for the participants to collaboratively reach a state of mutual understanding.

These findings clearly support the Collaborative model of reference (Clark and Wilkes-Gibbs, 1986). Question forms of introducing new entities into the discourse not only allow the speaker to check the listener's knowledge state, but also invite the listener to take an active role in establishing the referent in the participants' common ground. Question forms of introduction also made it easier for the listener to challenge the identity, or existence, of a referent, which many participants appeared to find problematic. Communicative success depended upon the speaker using an appropriate form of introduction, and the listener giving informative responses. The findings clearly demonstrate that effective communication depends on both participants accepting mutual responsibility for establishing the identity of references, and maintaining the common ground.

1.6 Summary of Research Supporting the Collaborative Model of Communication.

The research by Clark and colleagues has demonstrated how people establish mutual understanding through grounding contributions, which is a collaborative process involving both speakers and addressees. The findings of Garrod and Anderson (1987), Anderson and Boyle (1994) and Anderson (1995) support the view of Clark and his colleagues, effective communication requires the co-operation of all parties during discourse and that this is achieved on a moment by moment basis.

The majority of the research cited so far has examined the process of communication in face-to-face settings, or when participants are screened-off from each other. One of the key issues being investigated by this thesis is how this collaboration is achieved in computer mediated communication contexts; such as text-based Computer-Mediated Communication (CMC) and Video-Mediated Communication (VMC). In CMC, participants are remotely located and communicate with each other by text-based messages. VMC allows users to see and speak to each other; but access to the visual and auditory channels of communication is often more restricted than in face-to-face interactions. How will these mediated contexts affect the process of communication? How will users establish mutual understanding when they have restricted access to the range of channels of communication afforded in face-to-face communication?

Some possible answers to these questions are given by Clark and Brennan (1991), who propose a theoretical framework to account for the different ways in which people will establish mutual knowledge when they use different communicative contexts. Before describing this theoretical framework it is necessary to discuss the range of channels of communication that are afforded in face-to-face and mediated communicative contexts. These points will be discussed in detail in the next chapter.

Chapter 2. Communication Contexts: Constraints and Effects on Communication.

2.1 Channels of Communication

Face-to-face communication is a multi-channel form of communication. It provides many different ways of exchanging information, and is often considered to be the baseline against which all other communicative contexts can be compared (Chapanis, 1988; McCarthy and Monk, 1994a; Clark, 1996). In comparison to face-to-face communication, all other contexts offer a restricted range of sources of information. There are fewer channels of communication for exchanging information in telephone conversations or computer-mediated communications than in face-to-face interactions.

The channels of communication available in face-to-face communication have been described in various ways. In this chapter, the categories of channels (or systems) of communication purposed by Ellis and Beattie (1986) will be used. This categorisation was also employed by Monk, McCarthy, Watts and Daly-Jones (1996) and McCarthy and Monk (1994a). Ellis and Beattie state that face-to-face communication provides five channels of communication; verbal, prosodic, paralinguistic, kinesic and standing features.

The verbal channel is the main linguistic way of offering and exchanging information, in the form of utterances or written messages. Meaning is conveyed by the words chosen by speakers. The prosodic channel is also part of the linguistic system; conveying prosodic information (such as intonation, rhythm, phrasing and pausing). Prosody is part of the linguistic system, because it can "carry systematic contrasts in meaning" (Halliday, 1989, p.30). The remaining three channels of communication convey different forms of non-linguistic information. Paralinguistics include voice quality (such as timbre, tempo and loudness), facial and bodily gestures. These features can be used to signal, or stress the importance of what the speaker is saying.

(Halliday, 1989; Ellis and Beattie, 1986). The kinesic channel contains visual information in the form of facial expression, eye contact, gesture and body posture. Body and facial movements can convey information about a person's personality or emotional state, and facial and eye movements can also provide an indication about how an interaction is proceeding (Beattie, 1980; Crystal, 1995). The third non-linguistic channel of communication is concerned with standing features of interaction, these are forms of non-verbal communication which are more static (less dynamic) than occur in the kinesic channel. According to Ellis and Beattie, the main standing features of interaction are interpersonal distance, orientation and appearance. These features convey a range of information. For instance, interpersonal relationships can be determined by the proximity and spatial positons of conversationalists (Hall, 1966; Argyle, 1994), whilst clothes and adornments can display aspects of personality (Ellis and Beattie, 1986).

It is apparent from these descriptions of the channels of communication that non-verbal communication is a rich source of information. Argyle (1969) suggests that non-verbal communication serves six functions during conversations. It can signal levels of attention and responsiveness of participants, assist in controlling turn taking, provide the listener with a means of offering feedback to the speaker, indicate interpersonal attitudes, and gestures can be used to illustrate or augment speech (Argyle, 1969). Face-to-face communication, therefore, provides a rich variety of sources of information, all of which can assist in the processes of establishing mutual knowledge (Clark and Brennan, 1991) and coordination of conversational activities.

2.2 Channel Specialisation and Channel Combination

Channels of communication can be 'functionally specialised' (McCarthy and Monk, 1994a); they provide sources of particular kinds of information. For instance, the distance that people stand apart when talking (proximity) can only be assessed through the visual channel. Channels of communication can also be task specific, because people will "choose a medium of communication that they feel is appropriate for the task at hand." (McCarthy and Monk, 1994a, p. 42). For instance, Reder and Schwab (1988, cited in McCarthy and Monk, 1994a) found that for one-on-one supervision people usually preferred to interact face-to-face, but brainstorming did not require a specialised channel and could occur over a range of channels of communication.

Whilst channels can be functionally specialised, they can also be used in a complimentary manner; similar types of communicative information can be carried by one or more channels of communication. For instance, negation can be indicated linguistically by saying 'no', or non-verbally with the use of gestures or body-movements (such as head-shaking or hand movement). In face-to-face interactions verbal and non-verbal channels can be used simultaneously. Rather than being overwhelmed by the richness of information available in this context, Thompson and Ogden (1995) suggest that humans make efficient use of multiple language cues. Duplication of information in different channels can be beneficial, and may be reassuring for the receiver (McCarthy and Monk, 1994a). For example, information from visual and auditory sources can be accumulated and combined in a complementary manner (Brunswick, 1955, in Thompson and Ogden, 1995). This may help to disambiguate information presented in one of the channels. For example, Thompson and Ogden (1995) found that the process of speech perception improved considerably when when listeners were able to see the speaker's face.

When communication is carried out in mediated contexts the range of channels afforded by the context is restricted, in comparison to face-to-face interactions. Telephone and computer mediated contexts (CMC, audio-conferencing and Video Mediated Communication) differ in the number of channels afforded by the medium, and in the amount (or richness) of information available to users. The channels of communication available in several mediated contexts are described below.

2.3 Channels of Communication in Mediated Contexts

The telephone is probably the most familiar form of mediated communication. Less familiar forms of audio-only communication include 'voice mail' and audio-conferencing. In all of these media the linguistic channels (verbal and prosodic) are the chief means of communication available to users. Voice quality (part of the paralinguistic channel) can, however, also convey clues about the emotional state of speakers, and some aspects of personality (Rutter, 1987).

In text-based CMC the channels of communication are greatly restricted, to just the verbal channel. The information and tone of a message are conveyed by the words chosen by the sender; messages can be made to appear more or less formal. Other types of information, especially the social and interactional aspects of communication normally conveyed in non-verbal channels, are difficult to express in CMC (McCarthy and Monk, 1994a). For instance, the use of back-channels to indicate attention and agreement are not facilitated by most CMC systems. Designers of CMC systems have considered ways of facilitating backchannels, or fast responses to messages; for instance, the use of icons for signalling agreement or attention (McCarthy and Monk, 1994a).

As in other written contexts, stress and emphasis (paralinguistic information) can be achieved in CMC. Examples include the use of underlining, exclamation marks, and upper case (McCarthy and Monk, 1994a). Other authors have commented upon the use of specialised topographical symbols in CMC which can signal social information (Spitzer, 1986; Lea and Spears, 1992). Spitzer (1986) demonstrates the use of paralinguistics in CMC; for instance, the use of a smiling face :-) to denote humour, or the frown :-(to signal annoyance or frustration. The use of paralinguistic cues to embellish text is most often found amongst regular users of CMC; the symbols have no lexical meaning but can "signify socially shared meanings" (Lea and Spears, 1992, p. 324). The use of typological symbols and frequency of typing errors may also indicate something about the personality of the individual who wrote the message (Lea and Spears, 1992). CMC may lack many of the social cues available in other communicative contexts, but humans are flexible communicators; they will "use whatever cues are available to construct impressions of each other" (Lea and Spears, 1992, p. 324). The adaptability of human communicators is one of the factors that will be explored in the experimental chapters of this thesis.

In comparison to CMC, Video Mediated Communication (VMC) is a richer communicative context, as it allows users to interact in both linguistic and non-linguistic channels of communication. In VMC the audio and visual channels are presented via computer terminals and networks, so users can both see and hear each other even though they are not copresent. This context comes nearer to face-to-face interaction than any of the other mediated forms of communication, but how closely does it replicate a face-to-face context? The answer may partly depend upon the quality of the spoken and visual channels provided by VMC systems. The available literature shows that VMC systems vary considerably in the way that the visual and spoken channels are supported by the technology. The effects that the quality of the signals have upon communication and collaboration are the focus of the research reported in chapters 5 and 6 of this thesis; the relevant literature is reviewed in detail at the start of these chapters.

2.4 Conceptualising Differences in Communicative Contexts

A variety of theories and models have been proposed to account for the effects that different communicative contexts have upon human interaction. Some theories can be grouped together because they adopt a 'Cues Filtered Out Approach'; some contexts afford more channels of communication then others, and hence more sources of information. This way of conceptualising the difference between contexts would include 'Social Presence' theory (Short, Williams and Christie, 1976) and the 'Cuelessness Model' (Rutter and Stephenson, 1979). Social Presence theory categorises communication media according to degree of 'social presence' they afford; where 'social presence' denotes "the extent to which a medium was perceived to allow psychological close, interpersonal communication." (Rutter, 1987, p. 132). The social presence of a medium is based upon users' subjective judgements of the medium along several dimensions; such as warm - cold, personal - impersonal, sociable - unsociable. The Cuelessness model was based upon the number of social cues available in a communicative context; "the fewer the social cues, the greater the cuelessness." (Rutter, 1987, p. 136). The model has been refined over the years, with the emphasis shifting from the availability of cues to the usability of cues.

Several other theories, or models, have been proposed to explain the appropriateness of different medium for a variety of organisational purposes. An example of these models is the Information Richness model (Daft and Lengel, 1984; 1986), which describes media according to a set of criteria that may influence communicative effectiveness. The criteria include speed of feedback, variety of channels employed, personalness of source, and richness of language. According to Daft and Lengel (1986) and Trevino et al. (1990) people do understand the different information capacities of communicative contexts, and when given a choice will choose a media that is appropriate to the task in hand. For instance, when handling tasks which are complex, equivocal or are concerned with emotive subjects, people prefer to use an information rich media such as face-to-face conversation.

From the point of view of the research reported in this thesis, a more relevant way of conceptualising the effects of different configuration of channels of communication is offered by Clark and Brennan (1991). This framework describes communicative contexts in terms of the 'grounding constraints' afforded by the contexts. The term 'grounding constraints' refers to the characteristics of the media which facilitate the process of grounding.

2.5 Grounding Constraints and Communicative Contexts

Clark and Brennan (1991) propose that the process of grounding changes with communicative context. The reason for this is that, by the principle of least collaborative effort (Clark and Wilkes-Gibbs, 1986) "people should try to ground with as little combined effort as needed." (Clark and Brennan, 1991. p. 140). However, whilst some methods of grounding appear to require very little effort in communicatively rich contexts, the same form of grounding may take considerably more effort in other contexts. For example, whilst it is relatively easy to show agreement and understanding in face-to-face communication (by the use of gestures or backchannels), this is not achieved so easily in VMC where the visual channel is often impoverished; attenuation of visual signals makes it difficult to time backchannels effectively, or to use non-verbal forms of communicating understanding (Heath and Luff, 1991). The effort required to maintain the process of grounding, and to constrain the number of possible interpretations of an utterance, will vary with communicative context. Clark and Brennan suggest that the amount of effort required when using different grounding constraints "changes dramatically with the communication medium." (Clark and Brennan, 1991, p. 140)

2.5.1 Grounding Constraints

Eight grounding constraints are described by Clark and Brennan (1991) to characterise a range of interpersonal media. These are listed and described as follows:

1. *Copresence.* Conversation takes place in a shared physical environment. This constraint only occurs in face-to-face conversations, which is the most familiar communicative context. In face-to-face communication participants can speak and listen to each other easily, and see what each person is doing or attending to. Both participants share the same spatial and temporal location.

2. *Visibility.* Participants are visible to each other, as in face-to-face communication, but they may not be in the same location. For example, in VMC (such as video-conferencing or video-phones) participants can view video images of each other. However, the image is often restricted to just the head and shoulders of participants which makes it difficult for users to judge other participants' focus of attention.

3. *Audibility.* This characterises all communicative contexts where people converse by speaking. These include face-to-face and telephone conversations, and most video-conferencing systems. In these interactions participants can not only hear each other, but they can also make use of vocal characteristics of speech (such as intonation, the use of pauses etc.) to assist in the process of grounding.

4. *Cotemporality.* The recipient of a message receives that message at approximately the same time as the speaker produces it. For example, in face-to-face and telephone conversations there is no delay between the production and reception of a message, but delay does occur in some media (such as letter writing, faxes or electronic mail).

5. *Simultaneity.* Participants can simultaneously send and receive messages. This frequently occurs in spoken forms of communication, where participants talk over the top of each other, or interrupt each other. Another form of simultaneity can occur when one person - say the listener - uses non-verbal communication to respond to the speaker whilst the speaker is talking. In some computer-mediated communication systems participants can write simultaneously. One such system is '*talk*', a keyboard tele-conferencing tool which allows both participants to type at the same time, each persons contributions appearing in separate parts of the screen. A similar procedure is possible in some computer-mediated shared editing tools, such as ShrEdit which was designed by McGuffin and Olson (1992, in Olson, Olson and Meader, 1997).

6. *Sequentiality:* Conversational turns cannot get out of sequence. In face-to-face and telephone conversations the turns people make form an orderly sequence of utterances, which are temporally related to each other. If person A asks a question, this is normally followed by person B responding in the next turn with an answer (or a setting up a side-sequence to enable B to answer the question). In some communication contexts, sequences of turns can be disrupted, and an answer to a question may be separated from the question by the arrival of several other irrelevant messages. This sometimes happens in email and letter writing. Werry (1996) demonstrates this phenomena in conversations over an Internet Relay Chat system, and it is a characteristic of multi-party computer systems such as MUDs[1] (Multi User Dungeons/Dimensions) or MOOs (MUDs, Object Oriented).

7. *Reviewability*: In face-to-face and telephone dialogues speech is evanescent. Written and recorded messages persist and can be reviewed at a latter time, by the recipient or the sender. Some text-based CMC systems (such as Chat, a UNIX based teleconferencing system, and *talk*) allow users to see the last couple of received messages, so that the most recent part of an interaction can be reviewed.

8. *Revisability:* The sender can revise a message before transmitting it to the receiver. This is possible in many written modes of communication, such as letter writing and sending emails; the exceptions are some text-based CMC systems which transmit each letter as it is written (for example, *talk*). Revision allows people to adjust, or repair what they have written before the message is made public, which is not possible in speech

---

[1] Deuel (1996, p. 130) defines MUDs as "programmable text-based environments wherein participants may interact in virtual spaces." MOOs are variants of MUDs, and can incorporate virtual rooms, objects and player characters.

The descriptions of the grounding constraints presented above demonstrate that some communicative contexts afford more ways of establishing mutual knowledge than others. This can be put another way; some contexts provide considerably more ways of constraining the many possible interpretations of a message or utterance than other communicative contexts. Table 2.1 shows the grounding constraints that are available in a range of communicative contexts: face-to-face (Ftf), Video-mediated communication (VMC), telephone and other audio-only contexts, and text-based computer-mediated communications (CMC).

Table 2.1 Grounding Constraints and Communicative Contexts

Constraints	Ftf	VMC	Telephone	CMC
Copresence	X			
Visibility	X	X		
Audibility	X	X	X	
Cotemporality	X	X	X	
Simultaneity	X	X	X	
Sequentiality	X	X	X	
Reviewability				X
Revisability				X

Table 2.1 illustrates the point that some contexts afford a greater range of grounding constraints than other contexts. For example, in face-to-face conversations grounding can be facilitated by six of the grounding constraints, only the constraints of reviewability and revisability are absent in this context. In face-to-face interactions the process of grounding will be facilitated by the wide range of constraints available. However, some communicative contexts only afford a few grounding constraints; such as, electronic mail and other written media. In these contexts the process of grounding can only be facilitated by the constraints of reviewability and revisability.

Therefore, the process of grounding will be easier to achieve collaboratively in face-to-face than in the written contexts. In Clark and Brennan's terminology, grounding will 'cost' more - that is require greater collaborative effort - in some communicative contexts than in others.

2.6 Costs of Grounding

There are many different ways in which medium can effect the amount of collaborative effort required during grounding. Clark and Brennan describe eleven different ways, or costs, that change with communicative context. Some of these costs effect only the speaker, whilst others effect the addressee or all participants in the conversation. Clark and Brennan suggest that people often trade-off one cost against another, "costs are not independent of each other." (Clark and Brennan, 1991, p. 142). Speaker costs include *formulation* and *production*; costs paid by the addressee include *reception* and *understanding* costs. The remaining costs are paid by both participants; these include *start-up* costs, *delay* costs, *asynchrony* costs, *speaker change* costs, *display* costs, *fault* costs and *repair* costs. Some of these costs will be described and discussed in some of the latter chapters of this thesis, but a few will be described here to illustrate the effects these costs can impose upon people when they converse in different contexts.

2.6.1 Production Costs

The effort required to produce an utterance, or a message changes with communicative contexts. Most people find it relatively easy to produce utterances in contexts that afford speech and gesture; a greater amount of effort is required to produce written messages, especially hand-written messages (Clark and Brennan, 1991). Since it is more arduous to produce written messages people tend to use fewer words in written than in spoken medium. When people are adept at using a written medium, such as expert typists, the amount of effort required to produce written messages declines, and may result in longer communications. This is an example of the trade-offs between costs predicted Clark and Brennan; there is a trade-off between the cost of producing the message and the amount of linguistic output.

2.6.2 Speaker Change Costs

Detailed analysis of social interactions has shown that usually people take turns at speaking during conversations. For instance, Sacks, Schegloff and Jefferson (1974) found that most of the time only one person is talking, though there are often short episodes of overlapping speech. Whilst it is relatively easy to take turns at speaking in informal two person dialogues, the process of turn-taking is not so effortless in communicative contexts which lack the constraints of copresence, or where the visual channel is restricted or impoverished; for instance, telephone, CMC and some VMC systems. In these contexts, non-verbal signals which facilitate turn-taking are not so accessible. For example, there are recognisable patterns of gaze which indicate when one speaker is about to finish a turn. Kendon (1967) observed that speakers tend to avoid looking at listeners whilst they are talking. As they come near to the end of a turn they gaze at the listener, indicating that they are about to complete a contribution and that the person being looked at can take the next turn at speaking. In contexts which lack visibility and copresence the cost of switching between speakers is higher (requires more collaborative effort) than in face-to-face interactions. In accordance with the 'principle of least collaborative effort' (Clark and Wilkes-Gibbs, 1986), participants should attempt to reduce speaker change costs; one answer would be to use fewer turns, accomplishing more communication in a single turn than would normally occur in face-to-face conversations. This effect has been observed and commented upon by various researchers; for example, Cohen (1984), Chapanis (1988), Oviatt and Cohen (1991)

The theoretical work by Clark and Brennan (1991) provides a useful framework for comparing the ways in which communicative contexts vary. It also demonstrates how the process of grounding, of establishing mutual knowledge, can be effected by the restrictions imposed by different communicative contexts. It is easier to establish mutual knowledge in some contexts, because these contexts provide a wider range of channels of communication, and therefore a greater number of ways of constraining the meaning of utterances or messages. In written contexts, such as CMC, there are notably fewer grounding constraints. The way in which people adapt to the restrictions of CMC, and the means by which they achieve effective communication, will be explored in chapters 2 and 3 of this thesis. One of the major restrictions of CMC is that it is a written mode of communication. In later chapters, the effects of collaboration in VMC - which offers a much richer range of grounding constraints - will also be examined. These two computer mediated contexts vary along many dimensions, but perhaps the most significant difference is that they make use of different modalities; the written and spoken modes of communication. The differences between these two modalities will be considered further in the following section.

2.7 Spoken and Written Communicative Contexts

In the following sections some of the previous research which has attempted to determine the distinguishing features of spoken and written language will be briefly outlined. The section begins by looking at the manner of production of traditional (non-mediated) speech and writing, and the linguistic characteristics of these two modalities. Secondly, the manner of production and characteristics of text-based computer-mediated communication will be discussed, to determine the difference between traditional writing and CMC.

2.7.1 Manner of Production of Written and Spoken language

One way of distinguishing between spoken and written communication is to examine their manner of production (Brown and Yule, 1983). Speakers can make use of a wide range of channels of communication, verbal and non-verbal, to convey meaning and establish understanding. In contrast, the writer has a very restricted range of channels to achieve the same task. Whilst the breadth of channels available in speech can be beneficial, the speaker has to monitor all of these channels whilst simultaneously monitoring his/her own performance, noting if it matches his/her intentions and planning what to say next. In most informal conversations speakers do not start the interaction armed with a list of items they want to mention during the discourse; thinking what to say next is planned on a moment-by-moment basis. In more formal spoken contexts, such as after-dinner speeches or lectures, speakers plan what they will say in advance, and may use written notes to ensure that they cover the contents of their speech in an appropriate order (Brown and Yule, 1983). Whilst maintaining a coherent discourse may be more difficult to achieve in speech than in writing, spoken interactions have several advantages; for instance, it is relatively easy to modify what has previously been said if the addressee is showing signs of incomprehension.

The main advantages of written communications are that writers can normally take their time whilst they construct the text, and that writing provides a more permanent record of the message than speech. These advantages allow a writer to revise messages before they are dispatched, and to review previous messages as required (these are the grounding constraints afforded by written modes, as suggested by Clark and Brennan, 1991). In non-interactive written contexts - such as letter or essay writing - the writer is free from the time constraints imposed on spoken interactions, where delay or hesitation may mean that someone else takes over the role of speaker. A writer can, therefore, take time to ensure that the most appropriate words are used to convey meaning, altering and editing the text without the addressee being aware of these changes. The main disadvantage of written forms of communication is that writers are unable to monitor readers' reactions to their messages, and it takes a greater amount of collaborative effort to establish that mutual understanding has been achieved (Brown and Yule, 1983; Clark and Brennan, 1991).

2.7.2 Linguistic Characteristics of Speech and Writing

There has been a considerable amount of research concerning the differences and similarities between writing and speech. This has included research from a wide range of academic disciples; for example, researchers have explored the differences and the effects of these two modalities from the view point of sociology, history, anthropology, psychology, comparative linguistics and education (Brown and Yule, 1983; Biber, 1988). Naturally, these disciplines have approached the subject in a variety of ways, investigating the differences between speech and writing for a variety of reasons. For instance, sociologists and anthropologists have examined the effects of literacy upon society; linguists have attempted to identify the distinguishing linguistic features of spoken and written language (for example, Chafe, 1982; Beaman, 1984; Biber, 1986, 1988; Halliday, 1989; Miller, 1993). On the basis of all this research, it might be assumed that the distinguishing features of these two modalities were already firmly established. However, Biber states that there is "little agreement on the salient characteristics of these two modes." (Biber, 1988, p. 5).

Before outlining some of the reasons for this apparent lack of consensus between linguists, some of the general characteristics of spoken and written language will be described. The following list of characteristics are given by Brown and Yule (1983), who summarise a series of descriptive studies which have explored the differences between spoken and written discourse. Brown and Yule (1983) cite studies by Labov (1972); Sinclair and Coulthard (1975); Ochs (1979); Goffman (1981). Some of the key features of spoken language are listed below:

2.7.3 Characteristics of Speech in Comparison to Writing

1) The syntax of spoken language is less structured than written language. For example, speech contains a greater number of incomplete sentences, fewer subordinate clauses, and passive constructions are used less often. Speakers also tend to repeat

syntactic forms; that is, they use the same syntactic construction on several successive occasions.

2) The relationships between clauses are marked differently and less explicitly in speech; speakers tend to link clauses using words such as “and”, “but”, “then”. Rhetorical organisers - such as “firstly”, or “in conclusion” - also occur less frequently in speech than in writing.

3) Information is presented in a less densely packed manner in spoken discourse. For instance, information may be given in a series of shorter sentences (or clauses), and noun phrases are less heavily premodified. In this way the information density is attuned to the need of the listener who has to process the information in real time.

4) If the spoken interactions occur in face-to-face encounters, then participants also have the advantage of being able to use the full range of non-verbal forms of communication.

5) In spoken interactions participants can modify what they say as they talk. For example, if the addressee shows signs of not having understood an utterance, or if the speaker feels that he has not said something as clearly as he intended, he can reproduce the previous utterance in a different way (make a repair).

6) Speakers tend to use ‘generalised’ vocabulary. For instance, quantities are often less clearly defined; the speaker may refer to ‘a lot of objects’ rather than precisely stating the number of objects.

7) Speech contains frequent pauses, hesitations and false starts. Some pauses are silent, whilst others are filled by expressions such as ‘well’, ‘uhmm’, ‘you know’.

Some of the features mentioned above can be seen in the following extracts, which show examples of written and spoken discourse. The spoken extracts are taken from the HCRC Map Task Corpus (Anderson et al., 1991), and the written texts are drawn from a corpus of written texts (monologues) collected by Traxler[2] (1993). Both corpora made use of the Map Task (Brown, Anderson, Yule and Shillcock, 1984), which is a collaborative problem-solving task. As mentioned previously, the task involves an Instruction Giver describing to an Instruction Follower the whereabouts of a route on a simple schematic map. In the spoken version of the task participants worked in pairs, seated on either side of a table. In the written context participants worked in isolation; the Instruction Givers wrote out instructions for their partners to follow at a latter date.

The following extracts illustrate how instructions were given in the spoken and written contexts, for an identical part of a Map Task. A diagram of the map being described in these extracts can be seen in appendices A9 and A10. In these extracts IG indicates the Instruction Giver, and IF the Instruction Follower; three dots (...) mark the occurrence of an unfilled pause.

Extract 1: Written instructions

IG: Now make a diagonal descend downward until you are under the twin pines

Extract 2: Spoken instructions

IG: So you're sort of like drawing towards the top of the pine grove, then stop and go from ... left to right, in a diagonal underneath the pine grove

IF: A diagonal?

---

[2] Extracts from the written corpus reproduced with permission of the author.

IG: From ... from the top of the ravine to underneath the pine grove.

Extract 3: Written instructions.

IG: Go straight down until you get to the shore of the bay. Follow the bay until you just pass the coconut palm.

Extract 4: Spoken instructions

IG: So draw from concealed hideout, vertically downwards, till you come to that jutting out bit of the coastline. Then unti- until ... On that incline draw, you know, follow the line of the coast

IF: Mhm

IG: but don't ... don't go on to Crane Bay until you come to the ... the next bit that comes up. OK, stop there. Right, do you have ehm coconut palm to your left?

The written instructions are given in grammatically complete sentences, because the writer has time to plan and construct the message. Writers also use a richer vocabulary of words than speakers; for example, the word "descend" (extract 1) is used instead of the more common phrase "go down". The spoken instructions appear to be more hesitant, with frequent pauses (both filled and unfilled), repetitions of phrases or parts of phrases; extract 4, also demonstrates an example of a false-start (unti- until). The syntactic structure of the spoken utterances are often incomplete, and the information is presented in a less compact manner. The extracts also highlight another distinction between speech and writing, the opportunity for addressees to provide instantaneous feedback. In extract 2, the addressee indicates that she has not fully understood what the Instruction Giver has said, by asking a question. The Instruction Giver then provides additional information to clear up the mis-understanding. Another form of feedback is demonstrated in extract 4; this time the

Instruction Follower shows that she has understood the instructions, using a backchannel to acknowledge the contribution from the Instruction Giver.

2.8 Distinguishing Linguistic Features of Written and Spoken Texts

Whilst broad generalisations can be made about the differences between writing and speech, these are not universally accepted by all linguists. This point is stressed by Biber as follows: "The general view is that written language is structurally elaborated, complex, formal, and abstract, while spoken language is concrete, context-dependent, and structurally simple. Some studies, though, have found almost no linguistic differences between speech and writing, while others actually claim that speech is more elaborated and complex than writing." (Biber, 1988, p. 5).

One of the major disagreements between linguists centres around the structural complexity of spoken and written discourse; which is frequently considered in terms of subordination of clauses. Biber (1988) cites the findings of O'Donnell (1974), Kroll (1977), Kay (1977), Chafe (1982), and Brown and Yule (1983), who state that written text is more complex in structure (that is, it contains more subordinate clauses) than speech. Kroll (1977) reports that 35% of clauses in written narratives are subordinate, compared to 14% of clauses in spoken discourse. Biber claims that this statement is opposed by some authors (such as, Poole and Field, 1976; Price and Graves 1980), and reports that Blass and Siegman (1975) and Cayer and Sacks (1979) found little difference in the frequency of subordinate clauses in speech and writing. Beaman (1984) suggests that the reason for the discrepancy in these findings could be due to the different types of subordinate constructions that these researchers have analysed, and the variety of communicative functions served by subordinate constructions.

Other possible explanations, for the seemingly contradictory findings of linguistic research into the differences between speech and writing, have also been offered.

These have been based upon methodological issues; such as the range of linguistic features included in the analysis and how these features are defined, the types of genres being compared, the socio-economic and level of education of participants in the corpora (for example, Beaman 1984; Miller 1993). In the following section two very different approaches to the subject will be outlined; firstly the statistically based analysis under taken by Biber (1988; 1993), and then the more theoretical work of Halliday (1989). Whether either of these two approaches would provided a suitable framework for establishing the effect of computer-mediated communication, CMC and VMC, upon collaborative problem solving will then be considered.

2.8.1 Biber: A Multi-dimensional Approach to Spoken and Written Genres

The view taken by Biber is that there is no absolute distinction between speech and writing, instead they vary along several different dimensions. Biber therefore adopted a multi-dimensional approach (previously referred to as a multi-feature/multi-dimension approach, Biber 1986;1988) to determine the dimensions which distinguish spoken and written genres. Biber statistically assessed the co-occurrence of a range of linguistic features (such as tense and aspects markers, pronouns and pro-verbs, nominal forms, and subordination features) across a wide range of genres. These patterns of co-occurrence were then "interpreted in terms of the situational, social and cognitive functions most widely shared by the co-occurring linguistic features." (Biber, 1993, p. 333). In other words, Biber starts by determining which linguistic features regularly occur together, and then looks for the "underlying functional influence that encourages their use." (Biber, 1988, p. 13). This approach differs from many previous studies, which have grouped linguistic features together *a priori*, based on the functional characteristics of these features (for example, Ochs, 1979; Chafe, 1982).

Another distinguishing feature of Biber's approach to the differences between written and spoken genres, is the sheer size of the study. The patterns of co-occurrence of linguistic features where analysed in 491 samples of discourse, representing 23 different genres. The written texts included Press reports, academic prose, science fiction, personal letters; the spoken interactions included face-to-face and telephone conversations, broadcasts, and spontaneous and speeches.

The texts were analysed using computer based programmes, which calculated the frequency counts of each of the relevant linguistic features. Factor analysis was then applied to identify the groups of linguistic features that co-occurred regularly in the texts. Using this method of analysis 7 dimensions were identified in the 1988 study, two of which were deemed to be of lesser importance and have been discarded in more recent research (Biber, 1993). The remaining five main dimensions are categorised as follows: 1, Involved versus Informational Production; 2, Narrative versus Non-Narrative Concerns; 3, Explicit versus Situation-dependent Reference; 4, Overt expression of persuasion; 5, Abstract versus Non Abstract Information.

Genres can be placed at a point along each of the dimensions, depending on their 'factor score', which represents the frequency with which the linguistic features associated with that dimension appear in each genre. The genres can then be compared by their position along each of the five dimensions. Biber (1988) compared written and spoken genres along all five dimensions, and observed that none of the dimensions clearly differentiated between all types of written and spoken text. Instead the genres were spread out along the dimensions, with a considerable amount of overlap between some written and spoken genres; it was also noted that "genres can be markedly similar on some dimensions while markedly different on other." (Biber, 1993, p. 337).

The analysis also revealed that in both written and spoken modalities there are communicative settings in which the mode is used in what might be thought of as a typical way; for example, face-to-face conversations demonstrate typical use of spoken language. Biber reports that there are also situations when a modality can be used in a non-typical manner; for instance, presenting a prepared speech. Although Biber did not include CMC in his 1988 study, this context would also be an example of a modality being used in a non-typical way. The frequency of occurrence of some linguistic features associated with non-typical spoken genres can be very similar to that of non-typical written genres, so the profiles of these genress will be very similar.

In contrast to the multi-dimensional approach advocated by Biber, Halliday examined the difference between speech and writing in a more traditional linguistic manner. Halliday examined the differences between speech and writing starts by first deciding which features of speech and writing could be of theoretical interest, and then illustrating with examples the way in which these features vary with modality.

2.8.2 Halliday: A Theoretical Approach to Spoken and Written genres

Halliday (1989) examined the distinction between spoken and written utterances in terms of two related linguistic phenomena, information density and structural complexity. Halliday's theoretical and functional approach supports the claim that spoken language has greater complexity than writing. Spoken discourse contains more clauses and sub clauses than written text. The complexity of spoken language results from the intricate way in which clauses are linked together. In contrast, the syntactic structure of writing is relatively simple.

Halliday goes on to make the point that writing is also structured in a complex manner; but this is due to the amount of information carried by a sentence, or clause. Halliday illustrates this point by showing that writing has a higher ratio of lexical items (content words) to grammatical items than usually occurs in speech. So one way of distinguishing between speech and writing is to examine the density of information in both contexts. This can be achieved by calculating the ratio of lexical to grammatical items that occur in a text; this provides a measure of information density, which Halliday termed 'lexical density'. Halliday (1989) suggests two ways in which the lexical density of texts can be calculated. The simplest measure is the 'lexical density score', which is the proportion of lexical items in an utterance. The second measure, 'lexical density per clause', is defined as "the number of lexical items as a ratio of the number of clauses" (Halliday 1989, p. 67). The latter is a more revealing measure of lexical density, as it takes into account the varying grammatical structures of spoken and written language (Halliday, 1989).

Halliday reports that the lexical density of written texts is typically between 3 to 6 lexical items per clause, whilst speech usually has a lexical density of between 1.5 and 2 lexical items per clause. Written texts therefore carry almost twice as much information in a clause than spoken discourse. This is made possible by the 'elasticity' of clauses, which can accommodate large quantities of lexical information (Halliday, 1989). The majority of the content words occur in nominal groups, in the form of noun phrases, their premodifying nouns and adjectives.

How then is meaning conveyed in spoken language, which has a lower lexical density? Or, put another way, how is low lexical density achieved in speech? Halliday (1989) uses paraphrases to demonstrate that spoken and written texts can convey the same amount of information; however, they do so in very different ways. The following examples illustrate this point, and are taken from Halliday (1989, p. 61):

Example 1. Written text.
"Investment in a rail facility implies a long-term commitment"

Example 2. Spoken text.
"If you invest in a rail facility, this implies that you are going to be committed for a long term"

In the first example the information is highly compacted, but in the spoken example the information is spread out across a sequence of clauses which are intricately linked together. This produces spoken clauses which are low in terms of lexical density, though they are grammatically complex. In the examples given above the lexical density scores of the written and spoken texts are 70% and 35% respectively. The lower lexical density of spoken texts has the advantage of presenting the listener with information at a pace that can be assimilated in real time.

Halliday (1989) concludes that written and spoken language are both highly organised and complex forms of language use. The difference lies in the kinds of complexity associated with each modality: "The complexity of the written language is static and dense. That of the spoken language is dynamic and intricate. Grammatical intricacy takes the place of lexical density." (Halliday 1989, p. 87).

The differences between the approaches taken by Biber and Halliday vary considerable. Biber carries out a statistical analysis on the frequency of occurrence of a set of linguistic features,whilst Halliday illustrates his theoretical points with examples and paraphrases. These approaches to the same topic, the distinguishing features of speech and writing, are unlikely to produce similar findings. However, the authors do agree that the differences between spoken and written modes of communication are not due to a single, absolute linguistic phenomena; a range of linguistic features, or dimensions, need to be considered when comparing these modalities.

Could either of these linguistic approaches to the distinctions between speech and writing be usefully applied to the computer mediated contexts being explored in the current research? This will be discussed briefly in the following section.

2.8.3 Application of Linguistic Approaches to CMC and VMC

Biber's multi-dimensional approach could be applied to the computer-mediated contexts (CMC and VMC) being examined in this thesis. The multi-dimensional profiles obtained for VMC and CMC could then be compared to more familiar spoken and written communicative contexts (such as face-to-face interactions, or letter writing) and points of similarity and divergence between the profiles could be noted. For instance, it could be interesting to see where, or to what degree, the profiles for VMC and face-to-face interactions diverge. Useful insights into the nature of CMC could also be gained from the use of Biber's approach; for instance by determining whether the profile for CMC was closer to that of other non-typical written contexts, or more similar to non-typical spoken genres. However, this analysis would only show differences between the genres in terms of co-occurrence of linguistic features. The reasons why these patterns occur, the 'underlying function' of the associated linguistic features, would still need to be ascertained.

Alternatively, the lexical density scores of CMC and VMC could be calculated. This would provide another way of examining the effects that these contexts have upon communication. Such an analysis could illustrate the relative complexity of CMC or VMC, and whether this was due to informational density or complex structuring of clauses. An example of this type of analysis can be seen in Yates (1993; 1996), who used a corpus-based approach to examine the differences between CMC and more traditional use of speech and writing. The samples of spoken and written discourse were taken from the Lancaster-Oslo/Bergen (LOB) corpus and the London-Lund corpus respectively. The CMC samples were taken from a collection of computer-conferences organised by the Open University, and run over the CoSy system. The results showed that the lexical density of CMC and written texts were very similar; the CMC and written texts had a higher information density than the spoken interactions. Yates concludes that "CMC users package information in text in ways that are more written- than speech-like." (1996, p. 39). This type of analysis could be expanded to include comparisons between VMC and other communicative contexts.

These forms of linguistic analysis would illuminate the similarities and differences between CMC and VMC, and hopefully illustrate how these contexts compare to more traditional forms of communication. However, this could be insufficient for the current research, which attempts to explore the effect of computer-mediated contexts upon collaborative problem solving and ascertain how efficient communication is achieved in these contexts. According to the Collaborative model of communication, efficient communication occurs when participants take mutual responsibility for establishing and maintaining the process of grounding. Multi-dimensional profiles and lexical density scores would not afford a means of exploring this process, they would not illuminate the way in which people interact during collaborative tasks. Some other form of discourse analysis will be required for detailed examination of the CMC and VMC dialogues. Possible contenders will be discussed in Chapter 4, after Study 1 (which explores the impact of CMC upon task performance and the structure of interactions) has been introduced.

2.9 Manner of Production of Text-based CMC

Recent advances in communication technology now allow people to exchange written messages very rapidly. This occurs in text-based computer-mediated communication systems, such as electronic mail (email), computer-conferencing (CC) or on-line Multiple User Dimension (MUD); the latter is frequently used for purely social interactions. In all of these CMC systems participants communicate remotely and in writing, but the degree of interactivity varies from system to system. Electronic mail (apart from its potential to deliver mail very rapidly) appears to be very similar to letter or memo writing; it is an asynchronous medium in which messages are written and then posted or transmitted. Computer conferences and MUDs allow participants to interact synchronously, and almost in real time; making these written interactions feel similar to spoken conversations. This is reflected in the way that people describe computer conferencing; for example, as 'conversing over the telephone in a written modality' (Spitzer, 1986), or 'frozen dialogues' (Yates, 1993).

Whilst the manner of production of CMC is similar to other forms of written communication, CMC shares some of the characteristics of both spoken and written communication (Schrum, 1993). Participants in real-time computer conferences are under the same kind of time constraints as people in face-to-face conversations; they need to plan, think and write all at the same time. Whilst CMC shares some of the characteristics of speech and writing, it also has characteristics of its own. One of the major differences between traditionally written and CMC messages is that the former are printed onto paper, whereas CMC messages appear on the user's screen (Spitzer, 1986; Reinking, 1992). CMC messages seem to lack the permanence of printed texts, because "words on paper are permanent and palpable; words in a (computer) conference can disappear with a flick of the on/off switch" (Spitzer, 1986, p. 19).

Thinking in terms of Biber's multi-dimensional approach, CMC could be considered a non-typical use of written language. The multi-dimensional profile of CMC would, therefore, be similar to profiles for non-typical forms of spoken text; such as prepared speeches. The characteristics of the profiles would also depend upon the purpose of the interactions (Biber, 1988). If CMC was being used for transmitting task-oriented messages, then the multi-dimensional profile could resemble those obtained for many written genres; but, if CMC was being used for interactional purposes (for example, chatting to a friend in a MUD, or sending a greeting to a colleague) then the profile might be more representative of spoken conversations.

Further evidence for the effect that the purpose of the interaction effects the characteristics of CMC is provided by Yates (1993), who examined the texts produced in three different types of computer-conferences as follows: an Academic conference between staff members of the Open University; Course tutorials between undergraduates and tutors; Social conferences between undergraduate students. Yates calculated the lexical density per clause scores (Halliday, 1989) for these contexts. The results of the analysis showed that the Academic conferences had higher lexical density scores than either the Course tutorials or the Social conferences; the average lexical density scores for each genre were, Academic 4.2, Course tutorials 3.2, Social 2.4 (Yates, 1993). These scores indicate that the Academic computer conferences were more representative of written texts, whereas the lexical density scores for the Social CMC conferences were similar to spoken interactions. The Course tutorials fell across the boundary of the spoken/written texts. Yates (1993) concluded that computer-conferencing can be used successfully for a range of purposes, and that the similarity to speech or writing will depended upon the purpose of the interactions.

The question that now needs to be considered is, what effect will the characteristics of CMC have upon the processes of communication and collaboration. Before examining the available literature on this subject, the impact of other technologically mediated communicative contexts will be briefly outlined.

2.10 Impact of New Technologies upon Interpersonal Communication

During the last one hundred and twenty-five years the range of communication technologies has expanded considerably. The invention of the telephone (patented in 1876 by Alexander Graham Bell) eventually led to the development of other telecommunication systems, such as Fax, tele-conferences, computer-mediated communication and more recently Video-Mediated communications. Many of these new technologies were first introduced into organisations and industrial institutes, providing alternative methods for exchanging information and ideas, and were only later taken up by the domestic market. For example, the telephone was originally sold to businesses as an efficient alternative to the telegraph, but by 1920 the Bell

Telephone company was emphasising the social uses of the telephone (Kiesler and Sproull, 1992).

There has been considerable interest in the impact that these new technologies have upon human communication, both at the organisational and social level. In this section, some of this research will be outlined; beginning with the effects of the telephone, then computer-mediated communication (CMC) and finally Video-Mediated Communication (VMC). The focus of this section is on the way these technologies - which afford a restricted range of channels of communication when compared to face-to-face interactions - affect interpersonal communication.

2.11 Impact of the Telephone upon Interpersonal Communication

Rutter (1987) offers a potted history of the rapid expansion of use of the telephone from its invention at the end of the nineteenth century. The first UK telephone exchange was established in London in 1879, with just half a dozen subscribers; the number of subscribers had increased by 1887 to over 26 thousand, by 1984 this had risen to 30 million subscribers with almost 78% of UK households having a telephone (Rutter, 1987). The very rapid acceptance of the telephone by the business community would suggest that there was a real requirement for this mode of communication (Aronson, 1971; Cherry, 1971). The telephone opened up new avenues of communication both between and within organisations, allowing information to be exchanged more rapidly than previously.

Prior to the early 1970s, the impact that this new means of communication could have upon society and business tended to concentrate upon two factors; the economic advantages and the 'social impact' of telephone interactions. The economic advantages of using the telephone were that it could reduce the need for face-to-face meetings; people could discuss matters over the telephone instead of travelling, saving companies travel expenses and conserving the environment (Rutter, 1987). The second line of interest was the way in which the telephone could effect society, the 'social impact' of the telephone. Research included concerns about the 'decentralisation' of organisations, and the acceptability of telephone interactions instead of face-to-face meetings (for example, Rutter cites research by Christie and Holloway, 1975; Thomas and Williams, 1975). One of the major findings from these studies was that people said they would prefer to make new acquaintances in face-to-face interactions; the telephone, like other telecommunication systems (such as closed-circuit television, audio-networks) was not the preferred mode for meeting new people.

One of the reasons for this preference was that the telephone did not afford a visual channel. During telephone conversations people have to rely on vocal cues to determine 'social context cues'; such as the age, gender, social status, ethnicity, and personality etc. of fellow conversants. Concern was also expressed by social psychologists, who wondered if an audio-only modality would disrupt the process of communication. During the 1960s researchers drew attention to the importance of visual signals in social interactions (for example, Argyle and Dean, 1965; Kendon, 1967), highlighting the importance of non-verbal forms of communication (such as use of gaze and gesture) in managing smooth transition of turns taking between speakers. Examining the way in which turn-taking was achieved in purely spoken communicative contexts, such as the telephone, became the basis for much research; for example, Jaffe and Feldstein (1970), Cook and Lalljee (1972), Stephenson, Ayling and Rutter (1976), Butterworth, Hine and Brady (1977), Rutter and Stephenson (1977), Beattie and Barnard (1979). Contrary to expectations, the majority of these studies found that turn-taking was not disrupted by the absence of visual cues. In fact, Jaffe and Feldstein (1970), Cook and Lalljee (1972) and Rutter and Stephenson (1977) report that episodes of simultaneous speech occurred less frequently in telephone than face-to-face conversations.

Several theories have been suggested to account for these somewhat counter-intuitive results. Some researchers suggest that in spoken-only interactions, participants make use of verbal substitutes for visual behaviour; for example, filled pauses enable speakers to maintain conversational control, and inhibit listeners from interrupting the speaker (Maclay and Osgood, 1959; Cook and Lalljee, 1972; Beattie and Barnard, 1979). Alternatively, Rutter (1987) suggested that the lack of visual signals in spoken-only conversations leads to a more formal, or less spontaneous, style of interaction; simultaneous speech occurs less frequently in telephone conversations because participants cannot use facial expressions or gestures to reassure the speaker that the interruption is a spontaneous contribution, rather than a threat to the continuation of the interaction. The main effect of spoken-only (telephone) conversations is that they lack the spontaneity of face-to-face interactions, interactions are more formal because of the 'cuelessness' of this medium (Rutter, 1987).

The formality of interactions in mediated communicative contexts, such as CMC or VMC, is still a current issue. For instance, it has been suggested that conversations held in a VMC context (where a visual channel of communication is provided) are less spontaneous than face-to-face interactions (Sellen, 1992; 1995). The spontaneity or formality of interactions in CMC and VMC contexts will be discussed further in the later chapters of this thesis.

2.12 Impact of Computer-Mediated Communication (CMC) upon Interpersonal Communication

In CMC participants send and receive written messages over computer networks. There are several types of CMC systems, some affording synchronous communication (such as 'talk' or MUD), whilst others are asynchronous. The latter include electronic mail (email), computer bulletin boards, asynchronous computer-conferences.

The majority of studies examining the impact of CMC upon interpersonal communication have tended to concentrate upon the use of email within large organisations. The main characteristics of email systems are that they are asynchronous, they deliver written messages almost instantaneously, and that messages frequently lack many of the 'social context cues' available in face-to-face communication (Sproull and Kiesler, 1986; Siegel, Dubrovsky, Kiesler and McGuire, 1986). The characteristics of CMC ensure that it is an excellent way of rapidly disseminating information; it is also a relatively inexpensive way of communicating (Kiesler, Siegel and McGuire, 1984), and allows receivers to read their messages at their convenience. However, the absence of spoken and non-verbal channels of communication in CMC greatly reduces feelings of social presence and the provision of social context cue (Hiltz and Johnson, 1990). Electronic mail and other CMC systems can, therefore, appear to be a rather impersonal means of communication, and this may alter the way in which interpersonal communication and group discussions are conducted.

The effect that CMC has upon interpersonal communication has been studied using a variety of methodological approaches, such as field studies, self-report studies and questionnaires, and empirical laboratory based research. The latter appears to have been employed more frequently Examples of these different approaches are given in the following sections.

2.12.1 Self Report Surveys and Questionnaires

Several self report surveys can be found in the literature on the impact of CMC upon interpersonal communication. Sumner (1988) asked managerial and technical professional users of a CMC system (the Professional Office System, PROFS) which of several media (electronic mail, telephone, memos, face-to-face communication and group meetings) they would prefer to use for a variety of tasks. Electronic mail was preferred for routine tasks such as information transference, chasing up progress reports or scheduling activities. Face-to-face interactions were considered preferable for resolving disagreements, offering criticism, or reaching consensus. A similar approach was taken by Severinson-Eklundh (1986), who asked users of a computer mediated communication system (COM[3]) to rate the similarity of COM with other communicative media. The results of this survey showed that electronic mailing (email) was consider to be more similar to note-writing than letter-writing, and least like telephoning or face-to-face. Adrianson and Hjelmquist (1988) also investigated the use of COM. In this study users were asked to complete a questionnaire which ascertained their reasons for using the system, and whether COM was used to compliment or substitute for other communicative contexts. Adrianson and Hjelmquist report that users' generally found few problems working in this context; COM was an easy, friendly and stimulating context, and users' enjoyed its spontaneity. However, there were occasions when people misunderstood each other and occasional episodes of aggressive behaviour (such as, uninhibited communication or 'flaming').

In order to understand users' choice of media in greater depth, Lea (1991) studied the use of different communicative contexts in a large commercial telecommunications company. In this study participants were asked to compare eight different communication activities using a 'repertory grid' technique. The results showed that electronic mail was considered to share some of the characteristics of both speech and writing; it is similar to note-writing because it is asynchronous, but it is also has the spontaneous characteristic of speech. Overall, email was viewed as a good medium for conversations and social interaction, which would indicate that these users did not view it as an impersonal form of communication.

[3] COM was developed by the Swedish National Defense Research Institute. It has been used regularly since 1979, and is similar to a gigantic notice-board (Adrianson and Hjelmquist, 1988).

2.12.2 Field Studies into the Impact of CMC

Field studies, in which users are observed whilst they use CMC in their working environment, have also been used to assess the impact of CMC upon interpersonal communication. An example of this approach is seen in Hiltz, Turoff and Johnston (1986; 1989). Hiltz et al (1989) used this method to study the effects of anonymity during computer conferencing. Managers of a large corporation took part in computer conferences using either pen names, or their real names. Comparisons were also made with face-to-face group discussions. Hiltz et al. expected to find that participants in pen name conferences would display more disinhibited behaviour and greater signs of 'deindividuated' behaviour. Deindividuation, or reduced self-awareness (Diener, 1979) is conceptualised in this paper as the "extent to which the individual members lose their identity or individuality and get 'caught up in' the group." (Hiltz et al., 1989, p. 225). Examination of the messages contributed to the conferences showed that this hypothesis was not fully supported; there was very little disinhibited behaviour in any of the conferences. However, participants who used pen names had a tendency to disagree less with each other over the final group decision, which indicates that they were more deindividuated than real name conference participants (Hiltz et al., 1989). Anonymous conferences also contained more contributions than real name conferences, and group participation was more equally distributed.

2.12.3 Empirical Research into the Impact of CMC upon Interpersonal Communication

There has been a considerable amount of empirical research into the impersonal effects of CMC, which has shown that this context can promote feelings of depersonalisation, disinhibition and deindividuation. These effects have been investigated and commented upon by a number of social psychologists, with special interest being placed on the way in which the impersonal nature of CMC can affect group-work and group decision-making. The Committee on Social Science Research in Computing at Carnegie-Mellon University has contributed extensively to the literature in this field; for example research by Kiesler, Siegel and McGuire (1984), Siegel, Dubrovsky, Kiesler and McGuire (1986), Kiesler and Sproull (1992). The methodological design used by Kiesler et al. (1984) has become the prototype for much of the research in this area, so this study will serve as an example of empirical research into the effects of CMC upon interpersonal communication.

Kiesler et al. (1984) attempted to assess the impact of CMC upon group interaction and decision making. In this research small groups of students were asked to reach consensus on choice-dilemma problems. A repeated measures design was employed, so that each group completed the tasks in three communicative contexts: face-face, anonymous CMC and non-anonymous CMC. The CMC systems used in this study was the Converse program, which divides the monitor screen into three portions, one for each of the participants. The effects of context were ascertained using four sets of measures: 1) communicative efficiency, such as the time taken to reach consensus, the number of contributions made to the discussion and what percentage of the group discussion was work related; 2) group participation, and whether this was equally distributed amongst the group members; 3) interpersonal behaviour, such as whether one person dominated the group and the amount of uninhibited behaviour; 4) the degree of 'choice shift' (the difference between initial and final group decisions) and how consensus was achieved.

The results of this study showed that it took groups longer to reach consensus in the CMC contexts, although these discussions contained fewer remarks. There was little difference in degree of task-orientation between face-to-face and CMC contexts. The measures of group participation and group processes revealed that was greater equalisation of contribution in the CMC contexts, but CMC groups also showed significantly greater choice shift and there was greater amounts of uninhibited interpersonal behaviour in the CMC discussions (for instance, higher frequency of swearing, insults, name-calling and hostile comments).

Kiesler et al. suggest that there could be three possible explanations of these results. Participants in the CMC contexts could have encountered difficulty in coordinating their activities due to the absence of informational feedback in this context, or the findings may have resulted from the absence of social influence cues which control group discussions (for instance, the lack of a group leader to facilitate reaching consensus). Alternatively, the results could be due to the depersonalisation effect of CMC, which redirects participants attention away from their audience and makes them "...more responsive to textual cues, more impulsive and assertive, and less bound by precedents set by societal norms of how groups should come to consensus." (Kiesler et al., 1984, p. 1130). The authors suggest that the latter explanation fits their data best.

Employing similar methods other researchers have replicated these earlier findings. Participants working in CMC tend to produce more uninhibited speech ('flaming') and greater self-disclosure than occurs in face-to-face interactions (for example, Kiesler et al., 1985; Sproull and Kiesler, 1986; Smolensky, Carmody and Halcomb, 1990). However, it should be noted that Walther (1993), Walther, Anderson and Park (1994) and Lea et al. (1992) suggest that the frequency of flaming may have been over stated, and that there is very little data to support this particular phenomenon.

Research has also verified the effects of CMC upon group processes and decision making; for instance, it takes more time and effort to reach consensus in CMC than in face-to-face meetings (Sproull and Kiesler, 1986; Gallupe and McKeen, 1990; Adrianson and Hjelmquist, 1991; Kiesler and Sproull, 1992; Straus and McGrath, 1994). Decision-making can be more extreme, or polarised, in CMC group discussions (Siegel et al., 1986). Research has also shown that participants who took part in CMC discussions were less satisfied with their decisions (Gallupe and McKeen, 1990) or the products of their group work (Galegher and Kraut, 1990). These negative effects on decision-making in CMC are found to decline if the group's identity, rather than the individual's identity, is made salient (Spears, Lea and Lee, 1990; Lea and Spears, 1991) or if a team leader or manager is appointed who actively organises the discussion (Fafchamps, Reynolds and Kuchinsky, 1991).

Several studies have recently reviewed the literature in this area. For example McLeod (1992) and Hollingshead and McGrath (1995). McLeod (1992) carried out a meta-analysis of thirteen experimental studies which had examined the impact of CMC systems upon group process and task outcomes. In all of these studies the CMC systems were Group Support Systems (GSS), which supported the interactions between group members. The meta-analysis was based upon several aspects of group process and outcomes. These included the degree of task focus, equality of participation, time taken to make decisions, quality of the decisions, degree of consensus amongst participants and member satisfaction (McLeod, 1992). The results of the meta-analysis showed that "GSS leads to increased task focus, increased equality of participation, higher decision quality, longer time to reach a decision, lower consensus, and lower satisfaction."(McLeod, 1992, p. 273).

Hollingshead and McGrath (1995) compiled an annotated bibliography of fifty research reports which had examined the impact of CMC upon group work (computer-assisted group work). Most of these studies had examined ad hoc groups who completed tasks in a laboratory setting. The overall finding was that CMC group members did not interact as much as participants in face-to-face discussions; they exchanged less information, but took longer to complete the task. Quality of task performance was also found to vary with context, but was also influenced by the type of task being undertaken. Groups working in a CMC context generated more ideas, whereas face-to-face groups performed better in problem-solving tasks and tasks of a conflictful nature (Hollingshead and McGrath, 1995).

2.12.4 Summary: the effect of CMC upon Interpersonal Communication

In summary, the results of research into the effects of CMC upon interpersonal communication have shown that there are costs and benefits of using this context. The depersonalising effect of CMC may encourage equal participation in group discussions, and allow participants to suggest a greater number of alternative ideas. However, research has shown that it takes longer to reach consensus in CMC, that the quality of the decision making may be inferior and that users are less satisfied with the final outcome than in face-to-face interactions.

2.13 Video-Mediated Communication (VMC)

The use of video as a means of communication is not a recent innovation. Wilbur (1994) and Button and Maggi (1995) state that demonstrations of VMC systems can be traced back to 1930s. Whilst this communicative context has not experienced the rapid growth originally predicted, interest in the use of VMC has recently expanded. There are several possible reasons for this increased interest. Firstly, the cost of video-conferencing (equipment and cost of 'on-line' transmission) is decreasing, and "is forecast to fall even more dramatically" (Button and Maggi, 1995, p. 60). The cost of travelling, meanwhile, appears to be increasing; and so is the congestion the traveller has to contend with. Parker and Joyner (1995) suggest that "a reduction in travel costs is often used to support a case for video conferencing" (p. 429). Secondly, the rather limited performance of early VMC systems has been greatly increased by the recent availability of higher bandwidth connections. Whatever the reason, the resurgent interest in video mediated communication has sparked off a world-wide "flurry of research on technology for collaborative work" (Sellen, 1995, p. 402).

A wide range of methods and approaches have been used to explore the effect of VMC upon interpersonal and group communication, including naturalistic studies and laboratory experiments. In addition, researchers have made comparisons either between VMC and face-to-face interactions, between various configurations of VMC systems or between VMC and audio-conferencing systems. Some examples of these various approaches will be outlined here. Further details of this research can be found in chapters 5 and 6, where studies 2 and 3 explore several aspects of VMC using an empirical approach.

2.13.1 Field Studies of VMC

Several researchers have examined the effect of VMC upon group interactions by studying the use of VMC systems in the work place. Two examples of this type of research will be outlined here, work by O'Conaill, Whittaker and Wilbur (1993) and a field study by Isaacs and Tang (1994).

O'Conaill et al. evaluated the effect of two different VMC systems upon the characteristics of speech during real group meetings. One VMC system used the Integrated Services Digital Network (ISDN), which affords low quality audio and video signals. The second VMC system (LiveNet) provided high quality audio and video signals. Group meetings in these VMC contexts were compared to face-to-face interactions. O'Conaill et al. found that speech characteristics differed between the ISDN and face-to-face contexts; speaker turns were longer, turn-taking was more formal and there were fewer episodes of interruption and overlapping speech by ISDN users. In contrast, the higher quality of signals in the LN system produced speech that was similar to face-to-face communication, though turn-taking was still more formal in the LN context. O'Conaill et al. suggest that greater formality occurs in the VMC contexts because users find it difficult to interpret non-verbal communication (especially the use of gaze to assist with turn-taking procedures) in this context. The authors conclude that users of high quality VMC systems may still have to contend with some problems during group interactions.

Isaacs and Tang (1994), following on from larger study by Tang and Isaacs (1993), also used a field study approach to evaluate the effect of VMC. In this study communication between a group of software engineers was monitored whilst they used a desktop video-conferencing system (DVC). The quality of the video images and the audio signals in this system was slightly better than that produced by an ISDN system. Comparisons were made with face-to-face and telephone conversations, based upon participants' interactional behaviour (especially the use of visual and audio cues) and the mechanics of turn-taking. Isaacs and Tang found that the DVC system had several advantages over the audio only context. The video channel facilitated the use of non-verbal communication to control turn-taking procedures, as well as expressing understanding and attitudes. The authors conclude that "video may be better than the phone for handling conflict and other interaction-intense activities." (Isaacs and Tang, 1994, p. 63). However, these advantages only occur if the audio signals are transmitted almost instantaneously.

Comparisons between face-to-face and DVC contexts revealed that the video signals did not afford the full range of cues available in face-to-face interactions; for example, users of the DVC system found it more difficult to notice peripheral cues, manage turn-taking, hold side conversations or draw other peoples' attention to objects in the real world. Isaacs and Tang suggest that collaboration in VMC contexts could be improved if other electronic tools (such as shared drawing programs, or whiteboards) were incorporated into the system.

The use of field studies to evaluate VMC has provided many useful insights. However, it is not always possible to ensure that the contexts are being compared in a fair manner. For instance, Isaacs and Tang (1994) compared two meetings in each of the three contexts, but the number of participants varied with the context and over time. Five participants took part in one of the DVC meetings, but only two in the other; this pattern of participation also occurred in the face-to-face meetings. In contrast, the phone conversations were between 3 or 4 of the engineers. One way of ensuring that comparisons are made as fairly as possible is to set up laboratory experiments, in which the conditions can be carefully controlled. The next set of studies illustrate how this approach has been used to evaluate the effects of VMC.

2.13.2 Empirical Studies into the Effects of VMC

The research by Sellen (1992; 1995) and Gale (1990) will be presented here as examples of research using an experimental approach to evaluate the impact of VMC. Sellen (1992; 1995) used an experimental approach to compare the characteristics of speech in three video-conferencing systems. The opinions and feelings about using the systems were also ascertained from interviews and questionnaires The VMC systems differed in the way that they supported selective gaze and listening. One VMC system presented the video images using a Picture-in-Picture (PIP) approach, another VMC system used the Hydra system (which has individual VMC units, one for each participant), and the third VMC system used a LiveWire system in which only the current speaker is displayed to the group. Multi-party communication over the VMC systems was compared with face-to-face and spoken only interactions.

The results showed that the different VMC designs did not effect the structure of the conversations. Turn-taking was unaffected by ability to use selective gaze, and it was not disrupted by the absence of a visual channel in the spoken-only context. Differences were observed between patterns of communication in the face-to-face and all other contexts; face-to-face interactions contained more interruptions and less frequent use of formal means of turn allocation. On the basis of this data, Sellen concludes that adding a visual channel to mediated contexts did not greatly effect the process of communication. However, the questionnaire and interview data showed that participants preferred to use mediated contexts which provided visual access to others. The Hydra system was the most preferred VMC context, because it facilitated selective gaze and parallel conversations.

Gale (1990) also used an experimental approach to determine the value added to an office system by incorporating a visual channel of communication. Participants worked in three communicative contexts: data sharing, using a shared whiteboard; data sharing plus audio; data sharing plus audio and video. The audio and video channels in this experiment were high, of near broadcast, quality. Measures of participants' performance, feelings and perceptions of work groups were taken whilst they carried out a series of cooperative tasks. The tasks included information dissemination, creative cooperative work and scheduling meetings. The results showed that adding a video channel had no significant effect of upon task outcome, which was of a similar level of performance in all of the contexts. Communicative context did, however, have some impact on the time taken to complete one of the tasks; scheduling meetings took twice as long in the data sharing context, and was fastest in the data plus audio context. Increasing bandwidth, by adding a visual channel, did not reduce task time, as groups working in this context took longer to complete the task than participants who could not see each other. However, the questionnaire data showed that participants thought that provision of a visual channel increased the level of social presence afforded by context; the ratings of social presence increased as the bandwidth of the context increased. Gale concludes that adding audio and video to communication systems may be of benefit in social interactions, but increasing the bandwidth may have little effect upon task-oriented communication.

2.13.3 Summary of the Impact of VMC upon Interpersonal Communication

The studies and experiments reported above have shown that there are a range of ways of evaluating the impact of VMC, few researchers have used exactly the same method or similar comparisons. This may partly explain why there is little agreement in the literature concerning the effects of VMC upon interpersonal communication. Finn (1997) makes the comment that "There are as many areas of discontinuity across studies as there are of overlap, making direct comparisons of studies an inexact science." (p. 4).

It is also possible that the differences between VMC and face-to-face conversations are subtle, and hence more difficult to evaluate objectively than other communication contexts. However, self-report data (often referred to as subjective data) obtained in interviews and questionnaires indicates that people prefer to communicate in contexts that allow them to see each other. This was found in the studies reported above, but has also been remarked upon by other researchers. For example, the presence of a visual channel enhanced the feeling of social presence in remote, long distance VMC conferences (Anderson et al., 1997); participants felt more satisfied with their work when the VMC system afforded visual contact, less satisfied when the visual channel was disabled (Olson, Olson and Meader, 1997); participants in a study stopped using a prototype VMC system when the video signals were removed from the system (Isaacs and Tang, 1997). These examples illustrate that people feel it is easier to communicate when they can both see and hear other participants; the exact way in which the visual channel is beneficial seems to be difficult to specify, but it may increase the feelings of 'social presence' in video-conferencing situations.

In summary, the impact of VMC upon the process of communication is still being investigated. Further reviews of the literature are presented in chapters 5 and 6, and it is hoped that the studies reported in these chapters will add something to the literature on the restraints imposed by VMC upon collaborative problem solving. The effects of different communicative contexts upon task performance are considered in the following section.

2.14 Effects of Communication Contexts upon Task Performance

This topic is one of the main themes that runs through the series of experiments reported in this thesis. In this section, some of the previous research into the impact of mediated communication upon task performance will be discussed. These studies have examined a range of communicative environments, some of which closely resemble CMC and VMC (such as linked teletypewriters, and closed circuit television, CCTV). The studies have also examined the effect of modality upon a range of tasks, some of which are collaborative whilst others require negotiation. The majority of studies have found that communicative context does not effect task outcome if the task primarily involves transmission of information between conversants, or is a simple collaborative problem-solving task (for example, Davies, 1971; Chapanis, 1988). However, if the task requires participants to negotiate, or bargain with each other, then the mode of communication will have some impact upon task performance (Short, 1974; Short, Williams and Christie, 1976; Williams, 1977). In the following two sections some of this research will be outlined.

2.14.1 Communicative Context and Collaborative Problem solving Tasks

Some of the earliest experiments in this area examined the accuracy with which information could be transmitted in a range of communicative contexts. For example, Champness and Reid (1970) asked pairs of participants to transmit the contents of business letters over the telephone, a co-present audio-only context (participants where seated on either side of a table, and were separated by an opaque screen; henceforth, audio-only context) or face-to-face. The results showed that this task could be completed just as accurately, and as quickly, in all three contexts. Davies (1971), in a similar study, expanded the range of medium explored to include audio-video content (CCTV) and teletype-writing. Davies reported that participants were just as confident that they had transmitted the information accurately in all four communicative contexts, though greater efficiency was achieved in the teletype-writing context if adequate time was allowed for completing the task (Short, Williams and Christie, 1976)

Chapanis and colleagues (Chapanis, 1971; Chapanis et al., 1972, Ochsman and Chapanis, 1974; Chapanis, 1988) concentrated upon collaborative problem-solving tasks, exploring the effects of mode of communication upon task performance. The range of contexts varied in each experiment, but usually face-to-face was compared with various versions of spoken and written (hand-writing or teletype-writing) contexts. A wide range of collaborative tasks were used in these studies; the tasks typically involved one participant seeking information from another (Chapanis, 1988). Examples of the tasks included the following: Part identification problems, such as locating a specific miniature light socket from a stock of 63 such items; University class-scheduling problems; equipment assembly problems, such as assembling a trash can toter; locating the nearest doctor to specified location in a town, using telephone directories and street maps. These tasks were chosen to be "representative of tasks for which interactive computer systems are or could sometimes be employed ... they are of recognisable and practical importance in everyday life." (Chapanis, 1988, p. 129). The tasks had only one correct and objective solution; the measures of task performance in these studies were, therefore, based upon task completion and the time taken by participants in differing communicative contexts.

As well as measuring time taken to completion Chapanis and colleagues explored a range of measures of linguistic output. These linguistic measures included the number of words in each utterance or message, the number of messages sent, the percentage of messages that were questions, the type-token ratio, rate of communication as words per minute (Chapanis et al., 1977),. In some of the experiments, trained observers noted the time spent by participants in different activities; for example, time spent sending messages, searching for parts or information, waiting for an instruction (Ochsman and Chapanis, 1974).

The findings of the entire sequence of studies are summarised in Chapanis (1988). The first result of note is that the tasks were successfully completed in all communicative contexts, so collaborative problem-solving performance does not appear to be effected by mode of communication. However, the amount of time and linguistic output required to complete the tasks varied significantly; tasks were completed nearly twice as quickly if participants could speak to each other, written contexts were considerably slower. Even when ten communication contexts were contrasted (Ochsman and Chapanis, 1974) the same effect was observed; the tasks took longer to complete as the context became progressively more impoverished, but as long as participants could converse (regardless of whether or not they could see each other) they completed the task more quickly than in written contexts. There was no overlap between the spoken and written contexts in terms of solution time (Ochsman and Chapanis, 1974).

Analysis of the linguistic measures revealed that participants who could speak to each other said significantly more than users communicating in writing. On average, participants in spoken contexts used 8 times as many messages, five times as many words, twice the number of different words (type-token ratio) and communicated at a faster rate. Similar findings can also be found in the literature on the social and organisational impact of CMC. For example, Siegel et al. (1986), McGuire, Kiesler and Siegel (1987), Galagher and Kraut (1990), Gallupe and McKeen (1990), Lea and Spears (1991), Straus and McGrath (1994). Whilst these studies did not set out to explore the effect of CMC upon linguistic output, they confirm that users of CMC take longer to complete their discussions, and tend to do so using fewer, longer turns, than is usually found in either face-to-face or spoken interactions.

According to Chapanis and colleagues, the crucial factor affecting the length and structure of the communication was the absence or presence of the audio-channel; the addition of a video channel had no significant effect on time taken to complete the task or upon communicative behaviour. This finding has been shown to hold true for all of the problem-solving tasks (ten in total) included in the series of experiments by Chapanis and colleagues. The overall conclusion is that for this type of task, which involves the exchange of factual information, the task outcome is not effected by the mode of communication. However, the tasks can be completed more efficiently (in terms of the time taken to completion) if people can talk to each other.

2.14.2 Effects of Communicative Context upon Conflictful Tasks

Various types of bargaining or negotiation tasks have been used by researchers whilst exploring the impact of communication context upon the task outcome. In some research bargaining games, such as the Prisoner's Dilemma have been used; for example, Wichman, 1970. The Prisoner's dilemma is a game in which participants can either co-operate or compete with each other. Both experimenters examined the degree of co-operation that occurred when the game was played in a range of communicative contexts; for instance, Wichman (1970) compared the game in face-to-face, video-only, voice-only and a 'no communication' context. Similar results were obtained in both experiments; participants co-operated more in communicatively rich contexts, and the degree of co-operation declined as the number of channels of communicationdecreased.

Simulated negotiations have also been employed to produce conflictful task situations. Morley and Stephenson (1969; 1970) asked pairs of subjects to take part in a simulated industrial dispute, the subjects taking the role of either a Management or Union representative. In these studies participants were given written background concerning the industrial dispute, but one participant was given a stronger case than the other; that is, they were given more points with which to argue their case. The stronger case was given to the Management representative in the 1969 study, and to the Union representative in the 1970 experiment. In both experiments, the negotiations took place either over the telephone or face-to-face. The findings (which only reached significance in the second experiment, Short et al., 1976; Williams, 1977) again showed that this task was sensitive to communicative context; the person with the strongest case was more successful in telephone than face-to-face negotiations.

Morley and Stephenson explained the results in terms of the degree of formality that each communicative context engenders. Telephone conversations are more 'formal' than face-to-face interactions, therefore in telephone negotiations "the emphasis will be placed on inter-party rather than interpersonal aspects of the interaction." (Morley and Stephenson, 1969, p. 543) This would enable participants to make their appraisal of the issues under negotiation in a more objective manner, and the case with the strongest argument would win. In face-to-face negotiations, the more interpersonal aspects of the interaction would colour the judgement of the participants, which could explain why the person with the strongest case was successful less often in face-to-face negotiations.

The effects of communicative context upon this negotiation task have been replicated by Short (1971, in Short, 1974). The results supported the earlier findings, representatives with the stronger case achieved more favourable outcomes in the telephone conditions. Short continued to explore the interaction of negotiation and communicative context; for example, manipulating the degree of asymmetry in the negotiations - giving both participants equal number of points with which to argue - and the amount of conflict between the participants own views and those he was asked to advocate (Short, 1974).

In one experiment (Short, 1974) participants took part in a different negotiation task; they were asked to agree on three out of a possible nine options available to an industrial manager. Each options had different payoffs for the manager (person A) and the Union representative (person B). Before the task, A was asked to prioritise nine options in accordance with his own opinions; this list was then reversed and given to person B. The instructions for each participant were also manipulated; the instructions for A emphasised personal conflict, whilst B's pointed out necessity of bargaining to maximise possible payoffs. Negotiations took place over three media (face-to-face, Close-Circuit-TeleVision (CCTV), and a spoken-only context). The findings showed that the person whose personal beliefs were in accordance with the view he was advocating (person A) was *more* successful in face-to-face than telephone negotiations; however, if the participant was supporting a view at conflict with his own views (person B), then his negotiations were more successful in the telephone context (Short 1974; Williams 1977). Furthermore, the results from the CCTV context were more similar to face-to-face than telephone context. From these results, Short concludes that "the lack of the visual channel rather than the isolation in the audio condition was the basis for the effect of medium on the outcomes." (Short, Williams and Christie, 1976, p. 97.)

Short explains his results in terms of Social Presence theory. Telephone interactions are low in social presence, so participants concentrate on the task-oriented (rather than interpersonal) aspects of the conversation. A person who is arguing against his own beliefs will find this easier to do in an audio-only context, as he can concentrate on his line of argument, and give less consideration to the interpersonal aspects of the conversation. Other explanations could account for the findings. For example, it has been found that people avoid eye-contact when telling lies (Exline et al., 1961), so the task of presenting a brief which conflicts with one's personal opinions may be easier to do in a spoken-only mode of communication. However, the study by Short (1974) demonstrates the important point that tasks that require negotiation, or where people are asked to advocate opinions that are in conflict with their own views, are sensitive to communicative context.

2.14.3 Summary. Effects of Communicative Context upon Task Outcome

The impact that a communicative context has upon task outcome depends upon the purpose of the interaction. Simple collaborative problem-solving tasks appear to be relatively unaffected by mode of communication in terms of task outcome, though written modalities require more time to complete the task. The effects of communicating in different medium has a greater impact when the tasks involve bargaining or negotiation, or when participants are asked to advocate a view which conflicts with their personal convictions. Some types of tasks are more sensitive to communicative context than others.

Considering the wide array of communicative contexts now available, it is important that the effects of context upon task performance are fully understood, so that the appropriate context can be chosen for a particular task. There has been some research into the match between task and communicative context. Hollingshead et al. (1993) and McGrath et al. (1993) both report on the JEMCO workshop, a longitudinal study which will be described in greater detail in chapter 3. One of the findings from this workshop was that the performance of certain types of tasks was effected by communicative context. Negotiation and intellective tasks were performed better in face-to-face than CMC contexts. In contrast, tasks that involved generating ideas and making decisions were completed just as successfully in either contexts. Further research is required in this area. The research reported in the experimental chapters of this thesis may be of assistance in furthering our knowledge about the relationship between task and communicative context, especially the effects computer-mediated communicative contexts.

2.15 Summary of Review

The review started off by considering different theories and models of communication. The Collaborative model of communication, and the research introduced in the first chapter of this thesis, would suggest that effective communication depends upon establishing common ground. The process of grounding is accomplished effectively in face-to-face dialogues if participants accept mutual responsibility for establishing mutual understanding. Clark and Brennan (1991) suggest that the process of grounding will be effected by communicative context. They devised a theoretical framework to describe the ways in which the process of grounding will change with communicative context; this is based upon the set of grounding constraints afforded by different contexts. This theoretical framework provides a basis for comparing different communicative contexts, and the effects they have upon collaborative problem solving. The framework proposed by Clark and Brennan (1991) forms the main analytical basis for this thesis.

The rest of this chapter has examined different ways in which mediated contexts can effect communication; including the various channels of communication available in different contexts, effects of spoken versus written modes of communication, the impact of a range of technologies upon interpersonal interactions and how some contexts can have a detrimental effect on certain types of tasks.

Reviewing the literature, it appears that there has been abundant research into some aspects of mediated communication. However, other areas have attracted less attention. For instance, there has been relatively little research into the effects of CMC or VMC upon the **process** of communication and collaboration. Some of the literature reviewed has shown that communication contexts can effect the process of communication, but this has mainly been in terms of the amounts of linguistic output or turn management. What appears to be missing is detailed analysis of the structure and content of communication in different modalities, and how this relates to effective communication and task performance.

A second area of research which has received little attention to date, is how do people **adapt** to new communicative contexts? What changes, or adjustments, to the process of communication and collaboration do people make as they become accustomed to a novel communication context. Study 1 will examine how novice users of a CMC systems adapted to the restraints imposed in this text-based context. This will be continued in chapters 4 and 5, where the restraints imposed by limited access to audio and visual signals are examined within the context of VMC.

Chapter 3: The Impact of Computer-Mediated Communication

3.1 Introduction

The literature reviewed in Chapter 2 has shown that technological advances have inspired a wide range of research, which has explored the impact that communication technology has upon human communication. This chapter focuses on the impact of computer-mediated communication (CMC), which is a restrictive text-based communicative context. Study 1 explores the effects of CMC upon communication during collaborative problem-solving tasks. The study examines how novice users of a CMC system adapt to this form of communication in a series of tasks that lasted over several weeks. The effect of CMC upon communication and collaboration is explored by an analysis of task performance, the process of communication and the inter-relation of these two factors.

3.1.1 Why study text-based CMC?

In recent years many new forms of communication technology have been developed; such as Fax, electronic mail, voice mail, video phones and video-conferencing. Whilst all of these technologies can be used to support collaboration, it has been reported (for example, Hiltz and Johnson, 1990; Walther and Burgoon, 1992) that the most widespread form of mediated communication is still some form of text-based Computer-Mediated Communication; such as electronic mail, or computer-conferencing. Furthermore, it is suggested that "the use of relatively basic text-based communication systems are likely to predominate for some time to come." (McCarthy, Wright and Monk 1992, p. 267). Nowadays CMC is widely available and utilised in both public and private fields of life. For example, CMC is frequently used as a means of communication in Universities (Hiltz, 1986; Mason and Kaye, 1989; Harasim, 1990), military research establishments (Perry 1992) and large industrial organisations (Greenberg, 1991; Walther and Burgoon, 1992).

Whilst there has been an abundance of research into the effects of CMC, this has tended to concentrate upon the social and organisational aspects of communication in this context. However, some areas have received little attention. Very few studies have examined how people adapt over time to the restrictions imposed by computer mediated contexts.

3.2 Adaptation Over Time to CMC

One exception is reported in a series of papers by Hollingshead, McGrath and O'Connor (1993), McGrath (1993) and McGrath et al. (1993). These researchers examined the impact of CMC upon group processes and task performances in a longitudinal study, the JEMCO workshop. This laboratory based experiment explored the effects of CMC on a range of tasks, which were accomplished by small groups of participants over a thirteen week period. The study examined four factors which could affect task performance and participant satisfaction of groups working in CMC or face-to-face contexts. The factors included: the effect of the two media (CMC and face-to-face); the effects of media change, when the groups shifted to the alternative context (from CMC to face-to-face or vice versa) for weeks 7 and 8 of the study; the effects of task differences and of task-technology interactions, and how this interacted with experience of the communicative context; and effects of changing group composition, and how this interacted with the context and different tasks (McGrath, 1993).

The CMC system used in this study was a group communication support system (GCSS), which relayed messages between members of the group. Group members could browse through previous messages on a 'common message board', the contributions of each member were coded (A, B, C etc.) so that the sender could be identified. Group members could use the GCSS to read and respond to messages at any time throughout the 13 week period. Participants in the face-to-face context met in a laboratory once a week, and the meetings were video-recorded. Each week the groups were assigned a different task. There were four categories of tasks, based on McGrath's (1984) task circumflex. The tasks used in this study were drawn from the

Intellective, Decision, Generate and Negotiation categories. Several versions of each category of task were used, one early on in the study and another similar task in the second half of the study. This allowed McGrath and colleagues to see if task performance was affected by the novelty of the context or the familiarity of group members.

The results reported by Hollingshead et al. (1993) showed that the task performance in negotiation and intellective tasks was affected by communicative context; face-to-face groups performed significantly better than groups communicating in CMC. Non significant differences were observed for the decision making and decision generating tasks. However, differences in task performance disappeared as the study progressed; in the later tasks no effect of context was observed for **any** of the tasks. These results indicate that the performance of the CMC users (who were novice users of a GCSS) was initially adversely effected by the novelty of the context. A similar effect was also observed when the groups shifted contexts in weeks 7 and 8; the task performance of the groups now working for the first time in a CMC context was again significantly poorer than the performance of those working face-to-face. The changes made to the group membership (which took place in weeks 11 and 12 of the study) also adversely effected the task performance of CMC groups for a short time; in the face-to-face groups task performance was unaffected by changes in group membership (Hollingshead et al., 1993).

The subjective data showed that group members' ratings of satisfaction with task performance and group processes reflected the objective data reported above. In particular, changes in communicative context and group membership in the CMC groups resulted in lower rates of satisfaction. Group members were also asked, on a weekly basis, whether they felt that the communicative context they were working in was hindering or assisting them. Groups using the CMC system reported that the context inhibited their task performance. This effect did not change over time, and was independent of task type.

Overall, the findings of the JEMCO workshop demonstrate that changes in task performance could be ascribed to experience of the communication context, and changes imposed on the groups. The differences in task performance were not an effect of task type (Hollingshead et al., 1993). The study makes an important point, that users will adapt to a new communicative context if they are given time to gain experience of the context. The question that the study does not answer, is **how** do people adapt to a novel environment? What changes to the process of communication are required to perform well in restrictive contexts such as CMC?

The literature reviewed in the second chapter of this thesis outlined several of the restrictions imposed by CMC. These include the restraints of a written form of communication, the lack of a visual channel of communication (and hence the absence of most forms of non-verbal communication), and restricted access to social presence cues. Probably the most dramatic adjustment users of CMC have to make is to communicate in text messages. Many studies have shown that tasks undertaken in CMC taken longer to complete than in spoken contexts. As a consequence, the amount of linguistic output in CMC is considerably reduced. Whilst these features of CMC have been frequently reported, there is relatively little literature showing how these characteristics of CMC effect the process of communication, or how users adapt to the restraints imposed by this context to become effective communicators.

According to the collaborative model of communication (Clark and Wilkes-Gibbs, 1986; Isaacs and Clark, 1987; Clark and Schaefer, 1989; Clark and Brennan, 1991) establishing mutual knowledge is essential if communication is to be effective. The way in which mutual understanding is established is through the process of grounding (Clark and Wilkes-Gibbs, 1986). In spoken interactions this process is normally achieved through sequences of presentation and acceptance phases. How would grounding be achieved in CMC? How would the predicted reduction in linguistic output effect the process of establishing mutual understanding, and hence effective communication?

3.3 Aims of Study 1

The aims of Study 1 are to compare CMC and face-to-face contexts in terms of communication and collaborative problem-solving. The study explores differences in the process of communication in these two contexts, and how novice users of a CMC context adapt over time to achieve effective communication.

In order to achieve these aims, CMC users were asked to complete a series of collaborative problem solving tasks. The task used in this study was carefully chosen to fulfil certain methodological criteria. Firstly, it was considered important to use an interactive problem solving task, which required both participants to seek and exchange information. This is in contrast to the types of problem solving tasks used by Chapanis and colleagues (Chapanis, 1971; Chapanis et al., 1972, Ochsman and Chapanis, 1974; Chapanis, 1988), in which participants were given either the roles of information giver **or** information seeker. It was hoped that a more interactive task would highlight the problems encountered in communicating effectively in the restricted confines of CMC. Secondly, it was necessary to find a task that could be repeated on a number of occasions; several versions of the task were required so that adaptation over time could be examined. Thirdly, the task need to produce an objective measure of the effectiveness of the CMC interactions, some measure of task performance that would reflect how well participants had established mutual understanding.

One task which fulfils these criteria is The Map Task (Brown et al., 1984), this is a collaborative problem-solving task that has been used successfully over a series of tasks, multiple versions are available so that it can be presented over a sequence of tasks. It is, therefore, a suitable task for examining adaptations to the CMC context over time.

The Map Task was designed to produce spontaneous and unconstrained dialogue, which occur whilst pairs of subjects are engaged in a cooperative problem solving task. The Map Task is an engaging and interesting problem solving task, as both participants have some of the information necessary to complete the task successfully. One of the advantages of The Map Task is that the communicative performance of participants can be scored objectively, the accuracy of the completed maps providing an indication of how effectively the participants communicated.. In this way the Map Task differs from many of the tasks used in previous research into the effects of communication context, where time to complete the task was often the only criterion used to determine task outcome; see for example, Chapanis et al. (1972; 1977), Cohen (1984).

3.4 Previous Research Using the Map Task

One of the original corpora of Map Task dialogues was compiled by members of the Human Communication Research Centre (see Anderson et al., 1991b). The HCRC Map Task Corpus consists of 128 spoken dialogues, comprised of 64 face-to-face dialogues and 64 spoken-only dialogues; in the latter context participants were co-present but screened off from each other.

The corpus of dialogues was collected for a range of academic interests, covering research in linguistics, psychology and artificial intelligence. Examples of previous use of the Map Task therefore cover a wide range of subjects. For instance, in the field of linguistics the Map Task has been used to research the function of intonation in dialogues (Kowtko, 1995; 1997) and to examine the function of the word 'like' in dialogues (Miller and Weinert, 1995). The Map Task dialogues have also been of assistance in research into artificial intelligence, where they have been used to demonstrate that prosodic information can constrain language models for spoken dialogue (Taylor et al., 1996). In psychological research, the Map Task has been used to examine the communicative competence of children (Anderson, Clark and Mullin, 1991; 1994; Doherty-Sneddon, 1995; Doherty-Sneddon and Kent, 1996) and aphasic adults (Anderson et al., 1997), and to examine how coherence is maintained in spoken communications by children and adults (Anderson, Clark and Mullin, 1991; 1994;

Anderson, 1995). The Map Task has also been used to study the effects of several communicative contexts; for instance, examining the effects of visibility on dialogue and task performance (Boyle, Anderson and Newlands, 1994), exploring the effects of visibility upon intelligibility of spoken dialogue (Anderson et al., 1997), and the effect of VMC upon the process of communication and collaboration (Anderson et al., 1993; O'Malley et al., 1996; Newlands, Anderson and Mullin, 1996; Anderson et al., 1997; Doherty-Sneddon et al., 1997).

The use of the Map Task in comparisons of different communicative contexts is of particular relevance to this thesis. The research related to the use of VMC will be introduced in a later chapter. The study by Boyle, Anderson and Newlands (1994) will be described briefly in the following section.

Boyle, Anderson and Newlands (1994) examined the effects of visibility on task performance and various dialogue parameters, comparing face-to-face with spoken-only interactions. The study was based upon the HCRC Map Task Corpus, in which half of the dialogues occur in a face-to-face context and half in a spoken-only co-present context. Comparisons between these two sets of dialogues enabled the authors to analyse the effects of visibility upon task performance and the process of communication.

The results of the study by Boyle et al. showed that communicative context had no significant effects on the level of task performance; participants who could not see each other completed the Map Task just as accurately as participants in the face-to-face context. The process of communication was measured in terms of the length of the dialogues (number of words, turns, and length of turns) and the management of turn-taking (number and rate of interruptions) and the use of back channels to provide listener responses. Analyses on these dialogue measures demonstrated that the process of communication was effected by communicative context; participants in the spoken-only context said more than participants in the face-to-face context, and they produced a greater number of longer turns. Spoken-only dialogues showed signs of being more interactive, as they interrupted each other more frequently and provided a greater rate of back channel responses (Boyle et al., 1994).

Overall, the findings established that being able to see one's conversational partner "improved information transfer and the management of turn taking in a transactional problem solving task." (Boyle et al., 1994, p. 1). Communication in the face-to-face condition was more efficient, as the dialogues were shorter; this was probably because in this context people can use non-verbal communication to signal agreement and acknowledge each other's contributions. Participants who could not see each other adopted a range of techniques to compensate for the lack of non-verbal communication; they interrupted each other more often, and made greater use of back channels to provide their partners with an increased amount of verbal feedback (Boyle et al., 1994).

The study by Boyle et al. (1994) has demonstrated the way in which communicative context can have an effect on the process of communication. It has also shown the utility of examining a series of measures of the structure of the dialogues and task performance. The evaluation of the impact of CMC upon collaborative problem solving will follow a similar methodology. Study 1, therefore, adopts a 'multi-dimensional' approach to examine the impact of CMC upon collaborative problem solving. The term 'multi-dimensional' is being used here to indicate that a range of dependent variables, rather than examining just one variable, will be used to analyse any differences between CMC and spoken interactions. This analysis will be based upon measures of task performance as well as detailed examination of the structure and content of the dialogues, and how these measures vary over time as the CMC users adapt to this novel context. This type of multi-dimensional approach has been advocated recently by Monk et al. (1996) as a means of exploring the impact of new communication contexts, and can also been seen in the work by Schiano, Colston and

Pederson (1994); Galagher and Kraut 1994; Olson et al. 1994., Olson, Olson and Meader, 1997).

In Study 1 the impact of a CMC context will be investigated by comparing data (on task performance and the process of communication) collected in a CMC context with data from the HCRC Map Task corpus. The samples from the Map Task corpus will consist of dialogues from the spoken-only condition.

In summary, the questions being addressed in this study are: What impact does CMC have on the process of communication? How are CMC interactions structured? Are there differences in the structure of CMC and spoken-only interactions? Do users of CMC adapt their communicative strategies as they gain experience of CMC? If so, in what ways?

3.5 Study 1. Method for Comparison of Task Performance and the Process of Communication in CMC and Spoken Contexts

The subjects in the CMC condition were tested specifically for this thesis, and the results are compared to data taken from the HCRC Map Task Corpus. Details of the CMC condition are given below, followed by information concerning the comparison material and data for the spoken dialogues taken from the HCRC Map Task corpus.

3.5.1 Design of Study 1

A mixed design was used in Study 1. The between group factors was communicative context (2 levels, Computer Mediated Communication vs spoken context) and task order as a repeated measured factor (3 levels in the CMC context, and 2 levels in the spoken context).

3.5.2 Subjects in CMC Context

Twenty students from the University of Glasgow, aged between 17 to 26 years old (mean age 19.26 years) volunteered to take part in the study. The sample consisted of thirteen female and seven male participants, who were all native English speakers. A small financial reward was offered to the subjects, which they gained on completion of their third Map Task or earlier if they did not complete the full set of tasks.

Familiarity

Participants completed the task in pairs. They were allocated to the role of Instruction Giver or Instruction Follower in a random manner. The pairs of participants were all acquainted with each other; range of familiarity was between 2 months to 7 years (mean 4.24 years)

A note on sample size.

The size of sample used in this chapter is smaller than originally intended. The experiment started out with 20 pairs of participants in the CMC condition, however, only ten pairs completed all three tasks of the experiment. The main cause for the fall-off in sample size was due to problems encountered by the participants in fitting the three Map Task tasks into their other academic commitments. Some pairs completed more than one Map Task, but because the study is primarily interested in how participants adapted over time to the CMC context, the data presented here is for subjects who completed all three Map Tasks.

3.5.3 Computer Experience of CMC Participants

Sixteen of the twenty subjects who took part in the CMC condition had some previous experience of computers, three participants had completed certificated modules in Computing Studies at Secondary school. All of the participants in the CMC context were able to type, but with varying degrees of proficiency. The majority of the participants (70%) could type with at least two fingers of each hand. Only three of the participants had previously experienced any kind of computer-mediated communication, such as email. All of the CMC participants were naive users of the communication system used in this study.

3.5.4 Task and Materials

The task used in this study was the Map Task (Brown et al 1984). The Map Task has a well-defined goal, and produces an objective non-linguistic measure of communicative success. The materials for the Map Task are pairs of schematic maps, each pair including an Instruction Giver's map and an Instruction Follower's map. The maps are reproduced on A3 paper (30 cm by 42.0 cm). The landmarks on the maps are depicted by simple line drawings, and are labelled with their intended names. An example of one pair of maps used in Study 1 is shown in figure 1.

As figure one illustrates, each pair of maps is similar as they portray the same location. However, there are several specific and intentional differences. Examples are listed below:

1) Whilst a safe route past the landmarks is already marked onto the Instruction Giver's map, this route is missing from the Instruction Follower's map.

2) Absent/Present differences: some of the features on one map do not appear on the other map. For example, in figure 1 the Instruction Giver's map has a 'tribal settlement', but this is absent from the Instruction Follower's map

3) 2:1 difference: Some features appear twice on one map in a pair, but only once on the other map. For example, in figure 1 the 'golden beach' appears twice on the Instruction Giver's map, and only once on the Instruction Follower's map.

4) Although the start is marked on both maps, only the Instruction Giver's has the finish point marked

There are four different basic maps in the HCRC corpus, which were designed to be of approximately equal complexity or difficulty. Four variations of each of the basic maps are achieved by varying the named landmarks (but not the position of the landmarks) displayed on the map, giving a corpus of 16 different pairs of map.

One of the prime reasons for using the Map Task is that it provides an objective measure of task performance, which gives an indication of how effectively the participants communicated with each other. Task performance is measured by measuring the deviation between the route drawn by the Instruction Follower and the intended route; this is measured in square centimetres.

Materials for the CMC Condition

Six maps were chosen from the HCRC Map Task Corpus. The maps were divided into 2 sets (Set A and Set B) and were alternated between pairs of subjects, so that pair 1 used Set A maps, and pair 2 used Set B maps. Each set of maps had the same three basic routes but with different landmarks. The aim of using the two sets of maps was to prevent the participants accidentally gaining knowledge of the pathways before they had completed all three tasks. An extra map was chosen, which was kept in reserve and only used when a trial had to be re-run for any reason. For example, re-starting a trial after a technical break-down. Examples of all the maps used in this study can be found in Appendix A.

3.5.5 Apparatus for CMC Context

An interactive text-based message sending system, called 'Chat', was run between two interconnected Compaq 386 PCs. Chat is similar to the Unix™ 'Talk' program. The main difference between these two programs is that Talk transmits each character after it has been typed, whereas Chat allows the whole message to be produced and edited (if required) before the message is transmitted. The PCs were equipped with 14 inch colour monitors. The input from each of the participants was colour coded (blue for the Instruction Giver and red for the Instruction Follower), to facilitate reading the interleaved messages.

The CMC system used in this study allowed participants to compose and transmit messages to each other whenever they wanted to, so no form of turn-taking procedures were imposed by the system. The computer program time-stamped each message as it was transmitted, to the nearest second. On completion of each trial the computer programme produced a transcript of the interactions between CMC users, placing the messages into sequence based upon the time they were transmitted.

3.5.6 Procedure

Before commencing the task, each subject filled-in a short questionnaire. This included details as to the subjects age, gender, familiarity with their partner, and the subject's experience of keyboard skills, word-processors, computers and net-working (such as experience of electronic mail). An example of the questionnaire in given in the appendix.

Subjects were then allocated to 2 adjacent rooms, one room for the Instruction Giver and one for the Instruction Follower. The allocation of instruction role was determined by chance; subject whose surname came first in alphabet being designated the Instruction Giver. Participants kept the same role during each of the 3 Map Tasks they completed. Once the participants had settled in their rooms they were given brief instructions on how to accomplish the Map Task, these were as follows:

Instructions for the Instruction Giver.

Thank-you for agreeing to take part in this experiment.

On the table beside the computer you will find a map, on which a route or pathway has been drawn. Your partner also has a map, but the pathway is not drawn on it. Your task is to describe to your partner where the pathway goes, so that he/she can draw the path on to his/her map as accurately as possible; this path is the only 'safe' way to get from the start of the map to the finish. The map that you have may not be exactly the same as your partner's, as they were drawn by different explorers.

Please feel free to communicate freely with each other.

Instructions for the Instruction Follower.

Thank-you for agreeing to take part in this experiment.

On the table beside the computer, you will find a map and a pen. Your task is to draw on this map the route, or pathway, which your partner will describe to you. This pathway is the only 'safe' way to get from the start of the map to the finish. The map that you have may not be exactly the same as your partner's, as they were drawn by different explorers.

Please feel free to communicate freely to each other, using the computer terminal in front of you.

Written instructions on how to communicate with each other were also provided

To communicate to each other.

Just type in your message, and it will appear on the screen in front of you.

Please use the **return** key (marked <--) when one line of text on the screen is nearly full, otherwise the screen will scroll, and your partner will not be able to read all your message when you send it to him/her.

To **send** your message, press **escape** (esc), top left-hand key of keyboard, and then press **return** key. Your message will then appear on your partner's screen, and the last few lines will re-appear on your screen with confirmation that the message was sent.

Once the subjects had understood their instructions, they familiarised themselves with the keyboard and the communication system; this was achieved by encouraging participants to transmit several messages to their partner. This process also allowed the subjects to see that they received confirmation of message transmission, and that a tail of 2 lines of the previous text was left on their screen.

When both subjects were familiar with the procedures for the task and method of communication, they were told to begin. On completion of the first and second task the subjects were given an appointment for the following day until all 3 Map Tasks had been successfully completed.

3.6 Method for Collection of Data in the Spoken Context

The design for the collection of the Map Task corpus dialogues differs from the design of the CMC study. The corpus consists of sixteen groups of subjects working in a group, there being 4 students in each group or 'quad'. Each quad produces 8 dialogues by changing over partners and task roles (see Anderson et al., for more details). Data from two other HCRC collected quads were also included so that the sample size matched the size of the CMC study. Because the aim of Study 1 was to explore communicative adaptation to novel forms of communication, the CMC participants completed all 3 Map Tasks with the same partner, maintaining the same task role across the tasks; this is different from the HCRC corpus where the participants changed roles, acting twice as the Instruction Giver and twice as the Instruction Follower. Whilst the sample of dialogues taken from the Map Task corpus are not a perfect match for the CMC dialogues, they provide a base-line set of measures against which the effectiveness of communication in the CMC context can bejudged.

The sample taken from the HCRC corpus consists of 20 dialogues; these being the first and the second time that ten participants acted as Instruction Givers. Because of the way in which the Corpus was designed, the second occasion on which these subjects gave the instructions was either their second or third attempt at the task, depending on whether they had acted as an Instruction Follower in the interim.

3.6.1 Subjects in Spoken Sample

The data for twenty students was taken from HCRC Map Task Corpus, which is stored on the Human Communication Research Centre dialogue database, along with other corpora. Overall, participants in the HCRC Corpus were aged between 17 - 30 years, with a mean age of 20 years. This sample consisted of 14 female and 6 male subjects, who were all native English speakers. Subjects in this sample were aged between 18 - 27 years, with a mean age of 20.6 years.

Familiarity

Five of the pairs participating in the spoken condition were familiar with each other, on average these participants had known each other for about 2 years (Boyle et al., 1994). The other five pairs were unfamiliar with each other at the start of the experiment.

In the 'audio-only' condition participants were sitting in the same small room, seated one on each side of a double-sided easel, but prevented from being able to see their partner, and their partner's map, by a cardboard screen. Four different pairs of Maps from the HCRC Map Task Corpus were used in this 'audio-only' sample.

3.6.2 Apparatus

The dialogues were recorded on a DAT (Sony DTC 100EC) using Shure SNIOA close-talking microphones. Separate DAT channels were allocated to each speaker. Split-screen video recordings were also made.

Participants work in pairs, being randomly designated the role of Instruction Giver or Instruction Follower. Each participant has a copy of a schematic map, as described above. The Instruction Giver is told that her task is to describe the route shown on her map so that the Instruction Follower can draw it on his map. The Instruction Follower is told that his task is to draw the route on his map, as described by the Instruction Giver. Both subjects are told that their maps are of the same place, but may not be exactly the same as they have been drawn by different explorers, and differences in landmarks could occur. Subjects are encouraged to talk freely. The ensuing dialogue between participants in the spoken Map Tasks was audio and video recorded.

3.6.3 Transcriptions

Full orthographic transcriptions of the spoken dialogues were made from the DAT recording. These were available from the HCRC Map Task database. Transcriptions of CMC participants interactions were available from the merged time-logged contributions saved by the CMC program.

3.7 General Comments on presentation of Results throughout this thesis

Throughout this thesis the experiments were based on mixed designs. The between groups and within groups factors were defined as follows:

Communicative context was treated as a between groups factor. For example, in study one there were 2 levels of communicative context (CMC vs spoken).

Task order was treated as a repeated measures factor. This factor only occurs in Study 1, there are 3 levels of this factor in the CMC context and 2 levels in the spoken context.

Role of participants was treated as a within dialogue factor (2 levels, Instruction Giver vs Instruction Follower).

In line with the Collaborative model of communication, the unit of analysis is the dialogue; this represents the joint product of each pair of participants, either in terms of linguistic output or task performance.

Parametric statistics have been applied where-ever the data satisfied the required assumptions; that is, parametric statistics were used if the data was at interval or ratio level of measurement, if the sample data were drawn from a normally distributed population, and if the variances between samples did not significantly differ (homogeneity of variance). When the data could not meet these assumptions non-parametric statistics were applied instead. Homogeneity of variance was tested before each analysis of variance (ANOVA) was computed, by calculating F_{max} using the procedure laid out by Cohen and Holliday (1984).

3.8 Results of Study 1

The aim of Study 1 was to examine the impact of communicative environment upon collaboration, to see if different environments necessitate changes in communicative behaviour, and what adjustments are required for effective collaboration. The multi-dimensional approach applied in this study was carried out in the following order:

1) Comparison of task performance in CMC and spoken environments

2) Analysis of the process of communication, and how it is managed, in each communicative context. These analyses include comparisons of the structure of the dialogues (length of the dialogues in words and turns, time taken to completion), and collaboration over turn-taking procedures.

The initial impact of communicative context can be observed by comparing data from the first attempt at the Map Task in each context (comparing data from CMC and spoken task 1). Adaptations over time will be assessed by comparing data taken from the third attempt at completing the Map Task in CMC with the second attempt to complete the task in the spoken context. The third CMC task and the second spoken tasks are functionally equivalent, as they represent the final occasion on which participants acted as the Instruction Giver. For ease of reference, the second task in the spoken context will be referred henceforth as 'task 3' and the data for this task will be displayed alongside the data from the third attempt at the Map Task in the CMC context.

3.8.1 Task Performance

In the Map Task corpus task performance is assessed by calculating how accurately the route is reproduced on the Instruction Follower's map. Route accuracy is defined as the deviation in centimetre squares between the expected route, as shown on the Instruction Giver's map, and the route drawn by the Instruction Follower. The route accuracy scores for each trial can be computed using a procedure devised by Anderson, Clark and Mullin (1991). An acetate is made of the original Instruction Givers map, this shows the intended route and is overlaid with a grid of one centimetre squares. The acetate is placed on top of the Instruction Follower's completed map, and the area (in square centimetres) between the two routes can be computed (an example of this is given in Appendix B). A low deviation score indicates that the Instruction Follower's route drawn closely resembled the original route, whilst a high deviation score correlates with poor task accuracy.

The route accuracy scores were calculating for each map in the CMC and spoken samples. A square root transformation was applied to the measures of task performance in order to obtain homogeneity of variance. The group means (raw and transformed scores) for each set of tasks in the CMC and spoken context are presented in table 3.1, with the standard deviations given in brackets.

non significant in the spoken context ($p>0.1$). Task accuracy increased by over 36% in the CMC condition as the novice CMC users gained experience of the context.

In fact, by the third task the CMC Instruction Followers were drawing the routes as accurately as the spoken Instruction Followers. The route accuracy scores for these later tasks in both contexts did not differ significantly [$F(1,18) = 2.92$, $p>0.10$]. The group mean deviation scores were 7.66 cm and 6.52 cm for CMC and spoken modes respectively. Therefore, whilst the task performance of CMC users was initially poor, by the third map they were reproducing routes as accurately as participants in the spoken-only context.

The data in the previous analysis was based upon the route accuracy scores for tasks 1 and 3 in the CMC context. To determine if there was a significant difference in route accuracy between each of the successive CMC Map, a one-way ANOVA for the three CMC tasks was calculated on the transformed route accuracy scores, with task order as a within subject-pair repeated measure. A significant main effect of task order was obtained [$F(2,18) = 12.58$, $p<0.001$]. Further analysis of the group means, by post hoc t-tests (Tukey HSD) showed a significant increase in accuracy between task 1 and task 3, and between task 2 and task 3 ($p<0.05$); but, the difference in route accuracy scores between the first and second task was non-significant ($p>0.1$). The accuracy of the routes drawn in the CMC context gradually improved as the CMC users gained experience of the context, by the third Map Task performance in the CMC context was equal to performance in the spoken context.

In the Map Task, accurate route drawing is more likely to occur if the Instruction Giver and Instruction Follower communicate efficiently, ensuring that they have understood each other and establishing mutual knowledge. From the task performance results, it would appear that effective collaboration was harder to achieve in the initial CMC Map Tasks than in the first spoken Map Tasks. The fact that task performance in the HCRC sample did not significantly improve over time, suggests

Table 3.1 Group Mean Route Accuracy scores CMC and spoken Map Tasks.

	Raw deviation scores		Transformed scores	
	CMC	**Spoken**	**CMC**	**Spoken**
task 1	152.30 (64.29)	68.50 (41.90)	12.07 (2.73)	7.91 (2.56)
task 2	113.80 (77.48)		10.10 (3.34)	
task 3	60.90 (24.39)	45..30 (25.77)	7.91 (2.56)	6.52 (1.78)

A striking difference can be seen between the routes drawn by Instruction Followers in the two communication contexts. In their first attempts at the Map Task the CMC Instruction Followers route's appear to be considerably less accurate than the routes drawn in the spoken condition. However, the accuracy of the routes in the CMC context seem to improve over time. An ANOVA was carried out on the transformed route accuracy scores, with communicative context (CMC *versus* spoken) as a between groups factor, and task order (task 1 vs task 3) as a within-subject repeated measure. The analysis showed that task performance was significantly affected by context [F(1,18) = 12.22, p<0.01]. On the whole, participants in CMC context drew their routes less accurately, the overall mean deviation being 9.86 cm and 7.21 cm respectively.

Route accuracy improved with experience of the task [F(1,18) = 20.44, p<0.001]. However, further analysis of the significant interaction [F (1,18) = 5.51, p <0.05] by Simple Effects analysis showed that task performance increased only for the CMC context [F(1,18) = 23.54 p<0.001]; differences in task performance over time were that the degree of improvement in route accuracy for CMC users was not just a matter of learning how to do the task. So, how did the CMC users adapt their communicative behaviour to allow them to communicate more effectively? The next stage of the analysis, examines the process of communication, using a series of measures to determine any differences in the structure of CMC and spoken interactions. These measures have been shown to be in the literature to be sensitive indicators of the impact of communicative contexts (for example, Boyle et al., 1994; McCarthy and Monk, 1994a).

3.8.2 Process of Communication in CMC and Spoken Context

Number of Words (Word length)

To examine the relationship between task outcome and the process of communication, an analysis of the amount of communication required to complete the task in each mode was carried out. Using the transcribed dialogues, the number of words and turns occurring in the CMC and spoken interactions was calculated. As expected, the mean number of words per dialogue varied considerably with communicative context. The Mean words per dialogue in CMC and spoken condition are given in table means are 3.2, the standard deviations for these Means are shown in brackets

Table 3.2. Mean Number of Words per Dialogue in CMC and Spoken contexts.

context	Task 1	Task 2	Task 3
CMC	818.00 (326.91)	685.60 (225.33)	606.40 (279.88)
Spoken	1930.40 (1577.08)	-	1549.80 (992.29)

A Mann Whitney test (two-tailed) showed that the mean number of words required to complete the first task in CMC and spoken context differed significantly [$U(10,10) = 14$, $p < 0.005$]. Speakers used 5 times more words than CMC users on their first task.

Comparisons over time. Number of words per dialogue

In order to determine if experience of the CMC context had any effect upon the word length of the interactions, separate analyses of the Mean word lengths for the CMC and spoken contexts were computed.

First an ANOVA was applied to the data for the three CMC tasks; the dependent variable was the number of words required to complete the task, with task order as a repeated measures factor. The results of this one way ANOVA revealed a significant effect of task order [$F(2,18) = 5.25$, $p<0.05$]. Further analysis by post hoc t-tests (Tukey HSD) showed significant differences between the word length of the first and third CMC interactions ($p<0.05$). Other comparisons between the Means were non-significant ($p>0.1$). Participants in the CMC context required 26% fewer words to complete the third Map Task (818 words versus 606.4 words).

Analysis of the Mean word length for the first and third spoken dialogues was also computed, using a related t-test (two-tailed). A non-significant difference was observed between the group means [t (df 9) = 1.209, $p>0.1$], the means being 1930.40 and 1549.80 respectively for task 1 and task 3.

These results confirm that CMC participants used significantly fewer words to complete the Map Tasks than participants in the spoken context. Furthermore, the CMC interactions became shorter over time; but a similar pattern was not observed in the spoken dialogues. To see if the length effect was also reflected in the number of turns taken by participants in the CMC Map Tasks, an analysis of the number of turns per dialogue was computed.

Process of Communication: Number of turns or messages per dialogue

Before reporting the results of this analysis it is important to define what is meant by a 'turn' in this thesis. Various ways of defining turns have been reported in the literature. For instance, O'Conaill et al. (1993, p.404) suggest that turns are "attempts to take the conversational floor". Feldstein and Welkowitz (1978, p.335) provide a more technical definition, stating that "a turn begins the instant one participant in a conversation starts talking alone and ends immediately prior to the instant another participants starts talking alone". This appears to be the basis of the definition of turn employed by Chapanis et al. (1977, p 106) who defined turns in spoken and written communicative contexts as beginning "...when a subject began to talk, to write, or to type and ended when he had finished and relinquished control of the communication channel to his partner or when he was interrupted".

The definition of turns provided by Chapanis et al. will be applied in this thesis. In the spoken dialogues a turn begins when one speaker takes over the conversational floor, and ends when that speaker relinquishes the floor to another speaker, or is interrupted. Turns in the CMC interactions begin when participants begins to type a message and end when the message is transmitted to their partner. Turn-taking procedures were not imposed on the CMC users; they could send messages whenever they wanted to, and sometimes a sequence of turns was transmitted by one participant before a response came back from their partner.

The number of turns of spoken interaction was calculated from the transcribed dialogues, and the number of turns in the CMC interactions were obtained from the time-stamped sequence of messages made available by the Chat programme. The mean number of turns and messages per dialogue are presented in Table 3.3, the standard deviations are shown in brackets.

Table 3.3. Mean Number of Turns or Messages per Dialogue.

Context	Task 1	Task 2	Task 3
CMC	89.30 (57.03)	85.40 (34.98)	83.80 (38.67)
Spoken	256.00 (182.85)	-	229.40 (135.15)

The Means presented in table 3.3 suggest that the CMC participants took fewer 'turns' to complete the Map Tasks than the spoken participants. Comparison of the number of turns used for the first task in the spoken and CMC modes using Mann Whitney two-tailed test confirmed that the difference in mean number of turns per dialogue was significant [U (df 10,10) = 14.15, $p<0.005$]. The CMC interactions consisted of significantly fewer turns (65% fewer messages) than the spoken dialogues.

Analysis of the effect of task order was carried out, using the same procedure as applied to the word length of the CMC and spoken interactions. These analyses showed that the number of turns did not change significantly over time in either of the two communicative contexts: A one-way ANOVA was computed on the number of turns per dialogue in the CMC context, with task order as a repeated measure factor. This resulted in non-significant difference [$F(2,18) = 0.16$, $p>0.1$]. A related t-test (two tailed) was applied to the mean number of turns per dialogue occurring in the first and third spoken Map Tasks; the result was also non significant [t(df 9) = 0.62, $p>0.12$]. Whilst CMC users took fewer turns (messages) to complete the Map Task than participants in the spoken context, the number of turns did not change significantly over time.

3.8.3 Time taken to completion

The analyses so far have shown that users of the CMC context used considerably less linguistic output to complete the Map Tasks than participants in the spoken context. The amount of writing was reduced even further over time whilst the level of task performance increased. So, how did the CMC participants manage to package the necessary information into so little linguistic material? One answer could be that CMC users were taking more time to formulate their messages in comparison to the spoken dialogues. The following analysis examines the time taken to complete the tasks in the CMC and spoken contexts, and any changes that occurred over the tasks.

The time taken to complete each of the CMC and spoken interactions was calculated. This information was available from the time-stamped messages in the CMC context, and the time length of the spoken dialogues was made available from the HCRC Map Task corpus. The length of time (in minutes) taken to complete the tasks are given in table 3.4, along with the standard deviations (in brackets).

Table 3.4 Mean Time (in minutes) to Complete The Map Tasks.

Context	Task 1	Task 2	Task 3
CMC	61.20 (19.40)	51.10 (15.33)	46.50 (19.63)
Spoken	9.61 (6.51)	-	7.83 (4.43)

The Means presented in table 3.4 indicate that the CMC users took considerably longer to complete the tasks than participants in the spoken context. To see if these observations were significant, the data was analysed; first examining the time taken to complete the first Map Tasks in the spoken and CMC contexts, and then separate analyses of the effects of task order for each communicative context.

Comparison of the time taken to complete the first Map Task in both communicative contexts was computed using a Mann Whitney two-tailed test. The results showed that there was a significant effect of context [U (10,10) = 0, p<.005]; in the CMC context completing the first tasks took (on average) over an hour, in the spoken context the first tasks were completed in less than 10 minutes. A one-way ANOVA was computed on the time taken to complete each task in the CMC context, with task order as a repeated measures factor. The analysis showed that the CMC participants became quicker at completing the task as they gained experience of the context [$F_{(2,18)} = 4.59$, $p<0.05$]. Post hoc analysis of the three means involved (Tukey HSD) showed that the only significant difference occurred between the first and third Map Tasks ($p<0.05$). The results of a related t-test (two tailed) on the time taken to complete the first and third tasks in the spoken context was non-significant [t (df 9) = 1.76, $p>0.1$]. So whilst the CMC participants where able to reduce the time it took them to finish the Map Tasks (reducing the time taken for the third task by approximately 24%), the time taken by spoken participants to complete the tasks did not change significantly over the tasks.

The results obtained so far indicate that interactions in the CMC context took far longer than spoken dialogues, but gradually CMC users reduced the length of time taken to complete successive Map Tasks. This reduction over time could have been due to the decreasing number of words required to complete successive tasks, as writing less would take less time. The improvement in CMC task performance over tasks is not a result of the CMC users taking more time to complete the task. The improved level of task performance with shorter dialogues could indicate that the CMC users gradually adopted different ways of communicating and collaborating over time. So the next stage of this analysis will examine the process of collaboration, beginning with how turn-taking was managed in the CMC and spoken interactions.

3.8.4 Process of Collaboration: turn-taking and overlapping turns

In the CMC context users had no means of knowing whether their partner was composing a message, or waiting to receive a message. So taking turns at sending messages, without interrupting the activities of the other participant, could be difficult to achieve in the CMC context. It was anticipated that there could be a high rate of interruption in the CMC context, which might disrupt the process of communication and collaboration.

One of the problems with examining interruptions or episodes of over-lapping speech is that there are difficulties in defining the use of these terms, especially when considering these phenomena in spoken dialogues. Various ways of defining and identifying interruptions and overlaps have been applied by researchers (O'Connell, Kowal and Kaltenbacher, 1990). For example, Jaffe and Feldstein (1970) based their definition of interruptions on analysis of the characteristics of speech waves, whilst Bennett (1981) used a more subjective measure of whether a turn had been interrupted. More recently, computer programmes and speech tracking devices have been utilised to establish where interruptions occur. For instance, Sellen (1992; 1995) used a Speech Tracking system to analyse the on-off patterns of speech; Monk et al. (1996) used the 'Action Recorder' to time-stamp the on-off patterns of speech, and hence determine where one utterance overlaps or interrupts another utterance. The HCRC corpus of Map Task dialogues (which originally used the subjective judgement of transcribers to ascertain where overlapping speech occurred) has recently been updated in respect to the occurrence of overlaps. Since the participants' voices were recorded in separate channels (onto digital audio tape) it is now possible to determine very precisely when *any* degree of overlapping speech occurs. Examples of overlapping speech using this very precise technique are available on the HCRC web site (http://www.cogsci.ed.ac.uk/~matthewa/cgi/test.html). These objective means of determining when overlapping speech occurs will certainly be of assistance in future work in this area.

In this thesis, the working definitions of overlapping speech and interruptions are taken from analysis of the HCRC Map Task corpus. Episodes of overlapping speech are defined as occurring if "one or more words of the second speaker's contribution were perceived to overlap the first speaker's contribution." (Boyle et al., 1994, p. 8). Interruptions occur when one person starts to speak whilst another is already talking.

Whilst episodes of overlapping speech always contain an interruption, there can be occasions when an interruption does not involve overlapping speech; for example, if two people start to talk at exactly the same time (Boyle et al., 1994). To clarify the definitions being used, the following extracts demonstrate examples of 1) an interruption, and 2) an episode of overlapping speech which contains two interruptions. The examples are taken from Boyle et al (1994). The brackets (< >) show the start and end of the overlapping speech, and the forward slashes (/) represent where interruptions occur:

Extract 1. Example of an interruption.

IG: < Em, if you go/

IF: But where is it your trout farm?

IG: Em the right hand side of the page>

Extract 2. Example of an episode of overlapping speech

IF: < Yeah I'm at the lost steps/

IG: Well go, go over... go right... round them/

IF: Just no. Mm

IG: to the fallen pillars. That's to the left.>

Applying these definitions to occurrences of overlapping writing in the CMC context is not straight forward. In the CMC context participants could compose and transmit messages to each other whenever they wanted to; no forms of turn-taking procedures were imposed by the CMC system. The computer program time-stamped each message as it was transmitted, to the nearest second, and placed the messages into a sequence (transcript) based on the time that they were transmitted. Examination of the CMC transcripts showed that for most of the time CMC users appeared to take turns in sending messages, and then wait for a reply before sending another message.

On some occasions, however, one participant (A) would send a message, and before receiving a reply dispatch another message. The second message would arrive on the other participant's (B) monitor, disrupting any message that B was composing at the time. Messages which arrived as a result of one participant taking a turn 'out of sequence' had quite a disruptive effect on the composition of messages.

These episodes of 'out-of-sequence' turns were incorporated into the transcripts of the interactions by the computer program, which marked them as an INT (short for an 'interruption') at the appropriate point of the transcript. It is difficult to interpret these out of sequence messages in within the usual definitions of interruptions and overlaps suggested above, especially as participants could not be sure that they were interrupting the composition of text on their partner's screen. These INTs were, however, extremely disruptive; they literally split up any text being composed on the partner's computer monitor. This is illustrated in the following extract. In this extract the Instruction Giver (IG) sends a message (turn 1), and then starts to compose a second message to amend the first message. Meanwhile the Instruction Follower (IF) has replied to the first message (turn 2), but his incoming message disrupts the text being composed by the Instruction Giver; this incoming message is marked as an INT by the computer programme.

Message 1) IG: the lake is 6 cms above the left edge of the cairn

Message 2) IG: right edge so/

Message 3) IF: [INT] okay got that

Message 2) IG: (continuation) rry

Once the Instruction Giver has completed and transmitted his message, the computer programme reconstructs the ordering of the messages (depending upon the time at which they were sent). The extract above will now appear on the participants' monitors as follows:

Message 1) IG: the lake is 6 cms above the left edge of the cairn

Message 3) IF: okay got that

Message 2) IG: right edge sorry

The Chat programme did not provide users with a means of knowing what their partners were doing (waiting for a message, or composing and editing a message on the monitor), so these episodes of could have occurred quite frequently and might have been one reason why the CMC users performed less well on the initial Map Tasks than the participants in the spoken context. In order to see how frequently CMC users interrupted the orderly alternation of turn-taking, the transcripts of the CMC interactions were examined and the number of INTs computed for each of the three tasks. The mean number of INTs in each of the three tasks undertaken in the CMC context are presented in table 3.5, the standard deviations are given in brackets.

Table 3.5. Mean Number of Overlaps (INTs) per Dialogue in CMC context

CMC Task 1	CMC Task 2	CMC Task 3
12.70 (14.37)	14.60 (14.86)	10.50 (8.72)

In order to determine how the amount of overlapping turns in the CMC interactions compares to overlapping speech in spoken dialogues, the percentage of turns which occurred out of sequence (were INTs) was computed for the CMC dialogues. This data is presented in table 3.6 below.

Table 3.6 Percentage of turns out taken out of sequence in CMC interactions.

CMC Task 1	CMC Task 2	CMC Task 3
11.70 (9.05)	14.67 (10.91)	11.19 (8.29)

The data presented in tables 3.6 shows that between 11 % and 15% of written messages were interrupted by incoming messages from the other participant; across the three CMC tasks approximately 12% of messages being composed by one participant were disrupted by incoming messages from the other participant. Using the newly available data on the frequency of overlaps in the HCRC corpus, the percentage of turns which contained episodes of overlapping speech in the face-to-face and spoken only dialogues was computed; on average 24 % of turns in the face-to-face dialogues and 21% of turns in the spoken context contained episodes of overlapping speech. The CMC interactions, therefore, contained substantially fewer areas of overlapping speech than occurred in the spoken and face-to-face dialogues.

The low frequency of interrupting messages in the CMC context is slightly surprising, considering the potential problems involved in taking turns in this communicative context. Users of the CMC context could not see each other, and had no way of telling if their partner was writing a message or waiting to receive a message. An additional point made by many of the participants in the de-briefing session, was the slow pace of the interactions; participants felt that they had to wait a long time before they received a response to an earlier message. Some of the overlaps could be attributed to the frustration of waiting quite a considerable amount of time for the next message. However, the main finding of this analysis is that the rate of overlapping writing in the CMC context appears to be small in comparison to more familiar, spoken forms of communication. One possible reason for the low rate of overlaps is that the CMC users *may* have been cooperating to a high degree over the turn-taking process. Since disrupting the composition of messages could delay the communicative process, the CMC users may have adapted a cautious approach to turn-taking; reducing the risk of interrupting and disrupting the composition of messages in a collaborative manner.

3.9 Discussion

The first interesting result reported in this chapter concerns the level of task performance achieved by CMC users. The initial task performance of CMC users was very poor compared to the task performance of participants in the spoken context. Considering that the Map Task is a cooperative, non-competitive, task this result is surprising. Previous research has reported that collaborative tasks are not usually effected by variations in communicative context (Williams, 1977; Chapanis, 1988; Chalfonte, Fish and Kraut, 1991; Hollingshead et al., 1993). It is possible that the finding reported in Study 1 is a consequence of the measure of task performance we used; task performance in the Map Task is measured in terms of *task accuracy,* rather than the time taken to complete the task so often used in earlier research (for example, Ochsman and Chapanis, 1974; Cohen, 1984; Chapanis, 1988). This Study has shown that performance in the Map Task can be effected by communicative context, especially a restrictive context like CMC.

However, the results also demonstrated that task performance improved over time in the CMC context, as users gained experience of the novel communication environment. The accuracy of route drawing by CMC users in this study improved considerably over time, and equalled the performance of participants in the spoken context by the third trial. The remainder of this chapter concentrated upon examining how this was achieved by the participants in the CMC context; by examining the structure of the CMC interactions in comparison to spoken dialogues.

The analysis of the process of communication began by comparing the lengths of communication required to complete the Map Tasks in the CMC and spoken contexts. As expected, we found that written interactions in the CMC context were very much shorter (fewer words and turns per dialogue) than occurred in the spoken dialogues. The size of this effect was smaller than that reported in previous research. For example, Ochsman and Chapanis (1974) and Chapanis (1977) found a five-fold difference in word length between text-based and spoken modes; almost twice the difference reported in Study 1. However, the main focus of the analysis was on the effects of experience and adjustments over time, to determine how CMC users adapted to this novel context. The results showed that CMC interactions became shorter over time, participants gradually using fewer words to complete the succession versions of the task. Interestingly, the reduced linguistic output did not have a detrimental effect upon task performance; accuracy of route drawing increased over the tasks in the CMC context even though the number of words required to complete the later tasks declined.

Even with this significant reduction in word length across the tasks, the CMC users were still taking much longer to complete the Map Tasks than the participants in the spoken context. This is not surprising, as it takes much less time to say than to write or type a sentence. Reducing the amount that had to be written could have accounted for the decreased amount of time taken to complete the third tasks, the interesting question is what sort of reductions were made. How did the CMC users become more communicatively effective, whilst saying less?

One possible answer is that they adopted a more interactive style of communication, interrupting each other frequently to exchange information and establish grounding. However, the results of the analysis on the rate of overlapping writing in the CMC context suggests that CMC users interrupted each other relatively infrequently; only 12% of messages contained overlaps. There are two possible explanations for the lower rate of overlaps in the CMC context. First, the CMC context did not provide a visual channel, so the low rate of overlaps could be a result of the interactions being less spontaneous - more formal - than the spoken interactions. Greater formality in contexts that lack visibility has been commented upon; for instance by Rutter (1987) and Sellen (1992; 1995). Secondly, the communicative consequences of disrupting a partner whilst he or she is composing a message were quite considerable. Participants stated that they found the disruptions very disorientating and annoying, so they tried to avoid sending messages out of sequence. It could be that CMC participants were collaborating to a fairly high degree in order to achieve a relatively smooth flow of communication.

A more exhaustive analysis of the content and structure of the CMC dialogues could illuminate the changes in the process of communication in the CMC. This will form the focus of chapter 4, discourse analysis of the CMC dialogues.

Chapter 4. Impact of CMC. Analysis of the Structure and Content of CMC Interactions

4.1 Introduction

The findings reported in the previous chapter showed that CMC users improved the accuracy with which they completed the Map Tasks as they gained experience of the novel context. The question being explored in this chapter is how did they achieve this improvement? Since there was a decrease in the amount of linguist output required to complete the CMC tasks over the trials (that is, fewer words were used in successive trials though the number of turns per interaction did not change significantly), the improvement in performance was not simply a result of CMC participants using a greater number of messages, or longer messages. A possible explanation could be that as the users became familiar with the CMC environment they adopted different communicative strategies; finding more efficient ways of exchanging information and establishing mutual understanding in this restrictive communication context. One way of determining if such changes occurred would be to analyse the content and function of the CMC messages, and see if these varied over time.

The next step in examining the effect of CMC upon communication and collaboration is, therefore, to find a way of analysing the interactions which will illuminate the function and content of the CMC messages. Many different ways of analysing discourse (the term 'discourse' is being used here to refer to both spoken and written language) have been devised over the years. These have originated from a range of academic backgrounds - such as, philosophy, psychology, sociology, and linguistics - and focus on different aspects of discourse. For example, some analyse the way that discourse is structured whilst others concentrate on the functional characteristics of the content of utterances. In the following section some of the more promising candidates for the current analysis will be described, and their suitability for the task in hand ascertained.

4.2 Examples of Discourse Analysis

4.2.1 Speech Act Theory

One of the most influential philosophical approaches to discourse analysis is Speech Act theory, which originates from the work of Austin (1962) and was later extended by Searle (1969). The basic premise of this theory is that language is used by speakers to perform actions, or 'speech acts'. According to Austin an utterance consists of three types of speech acts: *locutionary acts* which convey the meaning of utterances, by the productions of sounds and words; *illocutionary acts* which carry the conventional force of an utterance (such as, asserting, warning, or undertaking); and also *perlocutionary acts* which are the effects achieved by saying an utterance (such as, convincing, persuading, or deterring).

Of these different types of speech acts, illocutionary acts have been most extensively investigated; most notably by Searle (1969), who extended Austin's work into a systematic framework for analysing the use of illocutionary acts in discourse. Searle segmented utterances in a slightly different manner to Austin. He separated the locutionary act into an *Utterance* act, performed by uttering words etc., and a *Propositional* act performed by referring and predicating. Searle retained Austin's definitions of illocutionary and perlocutionary act.

An important aspect of Searle's work is that he established the conditions which govern the use of a number of frequently occurring illocutionary acts; such as, promising, requesting, asserting, and questioning. The four classes of conditions (or 'constitutive rules') are termed the *Propositional content* conditions, *Preparatory* conditions, *Sincerity* conditions and the *Essential* condition.

An example will demonstrate the application of these different conditions. In asking a Question the utterance must be in the form of a proposition, or a propositional function; this is the *propositional content condition.* The *preparatory conditions* are that the speaker does not know the answer to the question, and that it is not obvious to either of the participants that the hearer will provide the information without being asked. The *sincerity condition* is that the speaker wants the information, and the *essential condition* (the point of the illocutionary act) is that the utterance acts as an attempt to elicit this information from the hearer. A speech act can only count as a specific type of illocutionary act if all of these conditions are satisfied.

Searle's work on illocutionary acts suggests one way of analysing the content of the CMC interactions obtained in Study 1. This type of discourse analysis could illuminate the writer's intentions in issuing an utterance, and the function of the written messages.

An example of the use of Speech Act theory in the analysis of CMC interactions is found in Cohen (1984). This study compared the distribution of a set of illocutionary acts that occurred in telephone and keyboard interactions. Participants in the keyboard (linked Cathode Ray Tubes, CRTs) context typed their messages on to computer terminals, which were "linked" so that whatever was typed onto one terminal appeared on the other after a short delay of approximately 1 - 2 seconds (Cohen, 1984). The linked keyboard context bears a close resemblance to synchronous CMC. Pairs of participants taking part in Cohen's study were asked to complete a task, during which one participant (the Expert) told his or her partner (the Apprentice) how to assemble a toy water pump.

The aims of the study were to identify the goals that speakers attempt to achieve in the two communicative contexts, to compare the discourse structure used to achieve these goals, to conduct an in-depth analysis of how acts of reference are expressed and achieved in the two contexts. In addition, the study evaluated the efficiency of a plan-based theory of communication, proposed by Cohen, to uncover speakers' intentions from the surface structure of utterances. It was hoped that this plan-based theory of communication could then be applied in speech recognition systems (Cohen, 1984).

The written and spoken dialogues (five from each context) were examined and the occurrence of eight categories of illocutionary acts was noted. Most of these speech acts took the form of requests; for example, requesting that the Apprentice identify a piece of the pump (Identify-Request), or assemble a part of the pump (Request Assembly-Action). Utterances which performed illocutionary acts were also coded with a set of 'operators' and propositions which described the assembly task. For example, if the Expert asked the Apprentice to "pick-up the blue tube base" this was coded as a REQUEST(PICK-UP (TUBE-BASE). Coders also noted utterances which *explicitly* requested the hearer to identify parts of the water pump, or to acknowledge that they had done so.

The results of the analysis showed that communicative context affected the referential process (Cohen, 1984). There were twice as many illocutionary acts in the spoken dialogues than the written interactions; this again demonstrates the length difference observed between speech and writing, as commented upon earlier in chapter 3. The distribution of the categories of speech acts was also found to vary with context. For example, a greater proportion of the spoken dialogues was concerned with Identity Requests than the written interactions; this category accounted for 35% of the illocutionary acts in the spoken context, but only 10% in the written interactions. Although the variations in the distribution of speech acts was not tested statistically, Cohen (1984) suggests that the difference in use of Identity Requests was the main factor that discriminated between the spoken and written contexts.

To complete his analysis of the act of reference in the two communicative contexts, Cohen tried to ascertain why Identification Requests occurred more frequently in the spoken dialogues. This analysis was based on the way in which new parts of the pump were introduced into the discourse; it bears a close resemblance to the work of Anderson, Clark and Mullin (1991a; 1994), Anderson and Boyle (1994) and Anderson (1995), on coherence in discourse which was described earlier in chapter 2. Cohen (1984) found that speakers habitually used Identification Requests to introduce new objects into the discourse, and that these were *explicit* attempts to get the Apprentice to identify a new piece of the pump. The Identity Requests in the spoken interactions were usually indirect speech acts, in the form of utterances such as "Do you see a little black rubber ring?"[1]., rather than direct requests such as "Find the rubber ring shaped like an O". Experts in the spoken context did not give instructions about the next stage of the task until the Apprentice had correctly identified the object. In contrast, Experts in the keyboard interactions usually introduced new objects into the discourse during their instructions on how to assemble the pump; which suggests that the writers tried to achieve a larger set of goals in a single message than speakers (Cohen, 1984). Explicit Identity Requests occurred very rarely in the written context, usually **after** referential problems had already been encountered. When explicit identification requests did occur in the written context they frequently took the form of direct speech acts; for example "Find the rubber ring shaped like an O" (Cohen, 1984, p. 111).

Cohen concludes that the way in which communicative goals are achieved differs between speech and writing in assemble tasks. Speakers are more inclined to ensure that a referent has been correctly identified before they give further task related instructions. In effect, "speakers attempt to achieve more detailed goals in giving instructions than do users of keyboards." (Cohen, 1984, p.97). This finding explains

[1] this example is taken from Cohen 1984.

why Identity Requests occurred frequently in spoken dialogues, but very rarely in the written modality. The study also showed that acts of reference in the spoken context were usually accomplished by indirect speech acts, using utterances which did not explicitly convey the speakers' intent.

Other references to Speech Act theory can be found in the literature on the design of CMC and Computer-Supported-Cooperative-Work (CSCW) systems. Insights from Speech Act theory have been used to assist and inform the design of new CMC and CSCW systems. For example, Winograd and Flores (1986) used Speech Act theory as the basis for their 'conversation for action'; this is a network model of speech acts which represents the history of the conversation and relationships between speech acts. Conversation for action formed the basis of the Conversation Manager in a CSCW system, the Coordinator. This is a work-group productivity system, which uses templates to assist users in generating and transmiting written messages. The templates are based upon the types of speech acts considered necessary for coordinated work within an organisation; such as, requests, counter-requests, promises, and assertions. The system also facilitates the transmission, storage and retrieval of the messages (Winograd, 1986). The Coordinator has been successfully used by organisations to assist in the organisation of individual activities, and promote collaboration with other users. Winograd (1986, p. 649) states that use of this CSCW system has "improved work capacity and effectiveness in a variety of settings."

This approach to modelling communication has not been without its critics. For example, Bowers and Churcher (1989) question whether it is possible to map utterances (or text messages) onto illocutionary acts in a one-to-one manner; the performance of one speech act may require several utterances, but there are also occasions when "one utterance can be used to perform many acts." (Bowers and Churcher, 1989, p. 201). A further general criticism of 'conversation for action' concerns the lack of attention to hearer uptake, which according to Austin (1962) is one of the conditions necessary for completion of a speech act. More specifically, the way in which requests can be completed in Coordinator is queried. Bowers and Churcher point out that the minimal pathway through the network would require four different speech acts (Request, Promise, Assert and Declare), but that this is hardly ever the case in everyday conversation where requests are frequently achieved in two steps using adjacency pairs; such as 'Request' followed by 'Compliance' or 'Rejection'. In order to capture the characteristics of conversation more fully, Bowers and Churcher based the design of their CSCW system (Cosmos) upon an augmented version of 'conversation for action', which was influenced by another form of discourse analysis, Conversational Analysis.

4.2.2 Conversational Analysis

A widely used form of discourse analysis from a very different academic tradition is Conversational Analysis, which was derived from ethnomethodology. It has been pioneered extensively by sociologists such as Sacks, Schegloff and Jefferson. The theoretical and methodological approach taken by these sociologists is quite different from any other form of discourse analysis. According to Taylor and Cameron (1987) the aim of Conversational Analysis is to illuminate the system which makes it possible for participants to carry out conversational activities in an orderly way. Analysis of the organisation of conversations is achieved by "extracting, characterising, and describing the interrelationships of the various types of sequential organisation operative in conversation." (Sacks, Schegloff and Jefferson, 1974, p. 698). In other words, the analyst carries out an in-depth examination of the *structure* of specific processes that occur in conversation, and this is achieved by extracting examples of the phenomena and finding ways of describing these processes.

Conversational Analysis has provided many insights into the way that everyday conversations are organised. For instance, Sacks et al. (1974) examined how turn-taking was achieved in conversations. They demonstrated that conversations consist of an orderly sequence of turns, that for most of the time only one person is talking, and that transitions between speakers are finely co-ordinated and occur at recognisable places in a turn (transition relevance places). Sacks et al. (1974) proposed a set of rules to account for the orderly characteristics of conversation. For instance, there are rules which apply to turn allocation (such as 'current speaker selects next'), and rules for dealing with violations of turn-taking procedures (such as 'first starter has rights'). When participants in a conversation apply these rules both long pauses between turns of talk and areas of overlapping speech occur infrequently.

Conversational analysis has been used to examine a range of processes that occur in conversation; for example, the analysis of opening and closing sequences of turns in conversations (Schegloff and Sacks, 1973), and the preference for making self-corrections in conversation (Schegloff, Jefferson and Sacks, 1977). The work by Schegloff and Sacks (1973) introduced the notion of 'adjacency pairs'. These are pairs of utterances which "properly have the following features: (1) two utterance length, (2) adjacent positioning of component utterances, (3) different speakers producing each utterance." (Schegloff and Sacks, 1973, p. 295). The first part of an adjacency pair is related in a specific way to the second pair part. Examples of adjacency pairs include 'question-answer', 'greeting-greeting' and 'offer-acceptance/refusal'. Other forms of adjacency pairs, such as 'summons-answer' are used to co-ordinate the opening and closing of conversations (Schegloff and Sacks, 1973). The importance of adjacency pairs is that they form recognisable patterns of interaction; the first pair-part initiates the response of the second-pair part in a predictable manner, which assists in the structuring and organisation of conversation.

The research cited above by Schegloff and Sacks (1973) and Sacks et al. (1974) demonstrate clearly that Conversational Analysis concentrates on describing structural aspects of conversation, and how conversation is managed in a pair-wise (or 'local') manner. More recently, these ideas have been incorporated into research exploring the impact of technologically mediated communication.

For example, a small study by McKinlay et al. (1994) used notions taken from Conversational Analysis, such as turn-taking and transition relevance points, to examine the impact of several CMC systems on turn-taking management and task performance. McKinlay et al. compared several synchronous CMC systems with face-to-face interactions. The CMC systems were provided with explicit or implicit forms of turn management. In the explicit turn-taking condition (ETT) participants used icons to indicate whether they were ready to talk or listen, but no form of turn-management was enforced in the implicit turn-taking condition (ITT). Analysis of the length of pauses between turns, and the number of areas of overlapping speech revealed that turn-taking in the ETT context was smoother than in the ITT condition; there were shorter delays between participants' contributions in the ETT context.

A more extensive application of Conversational Analysis is presented in the work of Frohlich, Drew and Monk (1994), who used a modified form of this analysis to examine the initiation and management of 'repair' in human-computer interactions. The term 'repair' refers to the ways in which conversational participants address problems in speaking, hearing and understanding each other. In the current context, a repair occurred when novices made a request of the computer which was not immediately granted (that is, there was a pause in the proceedings) or when they tried to undo a prior request. The analysis illustrated typical sequences of events that occurred in these situations, highlighting "some systematic repair-management features that had gone unnoticed in more quantitative studies of usability." (Frohlich et al., 1994, p. 419). For example, it was observed that repairs could be achieved within a few turns, but frequently took a considerable number of turns which might include several unsuccessful attempts at dealing with the problem. Frohlich et al. (1994) also noted that novice users tended to repeat a previous request before they attempt to modify it; this can complicate the repair procedure, as the repeated request may have to be cancelled before a new or modified request can be granted by the computer. Frohlich et al. suggest that novices seem to have problems anticipating how their requests will be interpreted by the computer, so establishing intersubjectivity is harder to achieve in Human-Computer Interactions than in everyday conversations.

4.2.3 Discourse Analysis

The term 'Discourse Analysis' is used here to refer to a specific approach to discourse which was proposed by a group of sociolinguists working at the University of Birmingham. The analysis is the result of several research projects carried out in the 1970s, which examined linguistic aspects of classroom interactions (Sinclair and Coulthard, 1975).

Sinclair and Coulthard (1975) proposed a method for describing interactions based upon 5 levels (or ranks) of discourse. These units of analysis are defined in terms of the interactive functions that utterances serve in speech. For example, whether an utterance was intended to evoke a response, was itself a response to an earlier utterance, or marked the boundary between one set of utterances and another. The five ranks of units - from highest to lowest- are called Lesson, Transaction, Exchange, Move (equivalent to a sentence) and Acts (equivalent to a grammatical clause). These discourse levels form a rank scale, units from a lower rank combine to form one unit in the rank above; for example, Acts are combined to form a Move, and a group of Moves can be combined to form an Exchange.

Sinclair and Coulthard classified 22 acts, which can be divided into 3 major categories; meta-interactive, interactive and those principally involved with turn-taking. Examples of acts from each of these major categories are given below. As the accompanying definitions demonstrate, the acts are defined by their discourse function; that is, in terms of the Acts they predict will follow them.

Marker: Realised by a closed class of items (such as 'well'). Its function is to mark boundaries in the discourse.

Elicitation: Realised by question. Its function is to request a linguistic response.

Bid: Realised by a closed class of verbal and non-verbal items (such as 'Sir', 'Miss', raised hand, finger clicking etc.) Its function is to signal a desire to contribute to the discourse.

Acts are then combined to form Moves, in a similar way that clauses are combined to make sentences. Sinclair and Coulthard list five categories of Moves: Framing, Focusing, Opening, Answering and Follow-up. These Moves can then be combined to form Exchanges. There are two major categories of Exchanges, Boundary and Teaching, which are distinguishable by the types of Moves that they contain. For example, a typical Teaching Exchange consists of Initiation (an Opening move), a Response (an Answering move), and finally Feedback (a Follow-up move). Boundary Exchanges consist of Framing and Focusing Moves. The latter Moves signal the end of one stage of a lesson and the beginning of another stage, and tell the class what is going to happen next (Sinclair and Coulthard, 1975). Exchanges are then combined to form Transactions, which equate to changes in topic or focus and are combined to form the complete Lesson.

Although Discourse Analysis was originally devised for analysis of class-room interactions, it has been applied to other communicative contexts. For example, Burton (1980) applied Discourse Analysis to dramatical texts (a short extract from Pinter's sketch 'Last to Go'). Burton found that the Initiation-Response-Feedback structure of Exchanges was inappropriate in this context, and suggested exchanges comprised of Opening, Supporting and Challenging Moves.

In contrast to Speech Act theory and Conversational Analysis, there appears to be no literature on the use of Discourse Analysis in CMC, CSCW or Video-mediated contexts.

4.2.4 Conversational Acts

An approach to discourse analysis which was developed for task-oriented dialogues was proposed by Traum and Hinkelman (1992). Conversational Acts originates from computational approaches to dialogue. It can be viewed as 'a generalizaton of Speech Act theory' (Traum and Hinkelman, 1992, p. 577) which pays special attention to the methods by which people establish mutual understanding during task-oriented spoken dialogues.

The approach taken by Traum and Hinkelman can be seen as a development of the work of Clark and colleagues (Clark and Wilkes-Gibbs, 1986; Isaacs and Clark, 1987; Clark and Schaefer, 1989; Clark and Brennan, 1991) and the Collaborative model of communication. As mentioned previously, the Collaborative approach is based upon the premise that communication is a joint activity. One of the central concepts arising out of this model is that both speakers and listeners play an active role in establishing the meaning of utterances in conversation; this is achieved through the process of grounding (Clark and Wilkes-Gibbs, 1986). In line with this earlier work, Traum and Hinkelman view discourse as a "collection of joint speaker-hearer actions" (1992, p. 578), which enable participants to ensure that utterances are grounded to a criterion sufficient for the current purpose.

Conversational Acts distinguishes between four levels of action which are "necessary for maintaining the coherence and content of conversation." (Traum and Hinkelman

1992, p. 578). The levels of acts progress from low-level acts to high-level acts in the following manner: turn-taking acts, grounding acts, core speech acts and argumentation acts.

The core speech acts and their relationship to grounding acts are of particular interest. The set of core speech acts include Inform, Request, Reject, Promise and are defined similarly to traditional speech acts. Traum and Hinkelman (1992) suggest that core speech acts form a unit of discourse that they term a 'Discourse Unit'. A Discourse Unit is an exchange of grounding acts required to initiate and ground a core speech act. Seven types of grounding acts are defined. For example, *Initiate* which is an initial utterance component of a Discourse Unit, and *Acknowledge* which shows understanding of a previous utterance (Traum and Hinkelman 1992, pp. 579-580)

Discourse Units represent the grounding process described by Clark and Schaefer (1989), in which contributions to the discourse occur through cycles of presentation and acceptance phases. A minimal Discourse Unit can consist of an initial presentation and its acceptance, which can be explicitly or implicitly expressed. Discourse Units can, however, contain many more grounding acts; these will include the initial grounding act and any continuations or repairs that are required to ground that act. Traum and Hinkelman accept that there are occasions when it is difficult to establish that an utterance or core speech act has been successfully grounded. They suggest that some utterances and Discourse Units may be grounded indirectly, and that the grounding of a later utterance can indirectly demonstrate understanding of a previous one (Traum and Hinkelman, 1992, p. 587).

Conversational Acts has been applied to dialogues taken from the TRAINS 91 project (Gross, Allen and Traum, 1992). According to Traum and Hinkelman, the aim of the TRAINS project was to build a computer system which could assist a human Manager to construct and execute plans in a simulated transportation and manufacturing domain; for instance, assisting the Manager to plan the transportation of boxes of oranges by train from a warehouse to a factory. The computer system would take the role of an 'intelligent planning assistant' that could communicate with the Manager in natural language. In the dialogues analysed by Traum and Hinkelman (1992) the role of the computer system was taken by a human agent, who conversed with the Manager over an audio link. It was hoped that analysis of the human-human interactions would provide a model of the communication processes required in this task-oriented domain, which would assist in the development of the computerised 'intelligent planning assistant'. Whether Conversational Acts can be applied to other task-oriented conversations still needs to be determined.

4.2.5 Conversational Games Analysis

A similar approach to discourse analysis is found in Conversational Games Analysis (Kowtko, Isard and Doherty-Sneddon, 1992; Carletta et al., 1995; Doherty-Sneddon et al. 1997). This approach provides a framework for looking at the communicative functions that speakers attempt to convey in their contributions. Conversational Games Analysis is derived from Artificial Intelligence models of communication, specifically from the work by Power (1979), Houghton (1986) and Houghton and Isard (1987) who developed this approach to code speaker intentions in simple task oriented dialogues between robots. More recently, Conversational Games Analysis has been applied as a means of analysing the goal directed exchanges that occurred during the Map Task. The analysis involves coding every utterance in terms of what the speaker is attempting to achieve, based upon the **function** of the utterance rather than its linguistic form or content.

Conversational Games Analysis is based upon two hierarchically related levels of dialogue structure, Conversational Moves and Conversational Games. Conversational Moves are Initiation and Response utterances, such as questions and answers. An utterance can be made up of more than one Conversational Move, or alternatively a

Move can consist of several utterances. There are thirteen categories of Conversational Moves, six categories of Initiating Moves and seven categories of Response Moves. For example, *Instruct* Moves are Initiating Moves which communicate a direct or indirect request or instruction. A *'reply-y'* Move is an elicited response which gives an affirmative reply to a previous question.

Conversational Games are defined by the goals that they achieve in a dialogue. They are units of linguistic interactions, consisting of the series of Initiating and response Moves required to a fulfil the purpose, or 'conversational goal' of the interaction. A Conversational Game takes the name of the Move that initiated the Game. For instance, an INSTRUCT Game is initiated by an *Instruct* Move, and consists of the Moves required to complete the instruction. Although some Games can be accomplished with just an Initiating Move, others may require several Conversational Moves to complete a Game. In addition, embedding of same level units is allowed in Conversational Games Analysis (cf. Sinclair and Coulthard, 1975), so that one Conversational Game can be nested within another Game. For instance, during an INSTRUCT Game a participant may wonder if she has correctly understood an instruction, so she may ask a question to seek clarification of the instructions; this initiates a new Conversational Game (in this case a CHECK Game) within the already existing INSTRUCT Game.

Due to the hierarchical structuring of the discourse units, Conversational Games Analysis bears a family resemblance to the Discourse Analysis devised by Sinclair and Coulthard (1975). Conversational Games are approximately equivalent to Exchanges; both Exchanges and Games are made up of smaller units of discourse, called Moves. At the same time, Conversational Games Analysis is similar to Conversational Acts as both can be used to illuminate the process of grounding. Several of the Moves and Games in Conversational Games Analysis illustrate the processes used by participants as they try to establish that they have understood each others' contributions. For example, a CHECK Game is initiated when speakers query that they have understood a previous utterance, an ALIGN Game is initiated when speakers query their partner's understanding of a previous utterance.

Conversational Games Analysis has been used to explore the impact of various communicative contexts; for example, it has been used to compare the structure and content of dialogues in face-to-face and spoken-only contexts (Anderson et al., 1997; Doherty-Sneddon et al., 1997) and in two Video-Mediated Communication contexts (O'Malley et al., 1996; Anderson et al., 1997; Doherty-Sneddon et al., 1997).

The most detailed report of the application of Conversational Games Analysis to different communicative contexts is presented in Doherty-Sneddon et al. (1997). In a series of studies the impact of context upon collaborative task performance and the structure of interactions was evaluated. The first study was an analysis of the Map Task dialogues from the HCRC Map Task corpus, applying Conversational Games Analysis to dialogues in the face-to-face and spoken-only context (henceforth, spoken context). The conditions under which these interactions took place have been described earlier on in this thesis (see Boyle, Anderson and Newlands, 1994). Comparisons of the route accuracy scores showed that communicative context did not significantly effect task performance, but the dialogues were significantly longer in the spoken context than in the face-to-face interactions (Boyle et al., 1994).

The results of the Conversational Games analysis highlighted two major differences between the structure and content of face-face and spoken dialogues; speakers elicited a greater amount of listener feedback (ALIGN Games) and checked their own understanding of previous messages (CHECK Games) more often in the spoken dialogues than in face-face interactions. Doherty-Sneddon et al. (1997) suggest that the increased use of ALIGN and CHECK Games in the spoken dialogues could account the for length differences between the two contexts. Spoken dialogues were longer than face-to-face dialogues because participants in the spoken context made greater use of ALIGN and CHECK Games to ensure that they were establishing mutual understanding (Doherty-Sneddon et al., 1997). In the face-to-face context participants could make use of non-verbal signals (such as eye gaze) as well as verbal forms of grounding; hence the face-to-face dialogues were shorter than the dialogues in the spoken context (Doherty-Sneddon et al., 1997).

The second study was designed to determine if VMC provided the same benefits as face-to-face interactions in the first study. In this study participants completed the Map Task in one of three communicative contexts; VMC with possibility of making eye contact, VMC without eye-contact, and audio-conferencing. In the two VMC conditions participants were provided with high quality visual and audio signals; in the audio-conferencing context participants could speak to each other over the audio-channel, but the visual channel was disconnected (fuller details of the VMC systems are presented in chapter 6). In these three conditions participants completed a computerised version of the Map Task, which was presented to them on the computer monitor. The first finding was that communicative context did not significantly effect the level of task performance (Doherty-Sneddon et al., 1997). The findings from the Conversational Games Analysis of the VMC and audio-conferencing dialogues showed just one significant difference between the three contexts; there was a significant increase in use of ALIGN Games (the speaker eliciting feedback from the listener) during audio-conferencing (Doherty-Sneddon et al., 1997). So here again, when participants were unable to see each other they made greater use of verbal methods of ensuring that they were in agreement, and that they had established mutual understanding.

4.2.6 Conclusions and Choice of Analysis for the CMC protocols

The previous discussion has briefly outlined several forms of discourse analysis. In considering which of these analyses would be most appropriate for the current research, the aims and objectives of Study 1 should be kept in mind. To recap briefly, this Study aims to explore the ways in which novice users of CMC adapt to the restraints imposed by this context, and whether the observed improvements in task performance were related to changes in communicative strategies. The analysis should, therefore, be based upon an examination of the content and function of the CMC messages. It would also be preferable to use a form of discourse analysis which has proven reliability. With these considerations in mind, the advantages and disadvantages of the different forms of discourse analysis introduced in the previous section can be considered.

Applying Speech Act theory to the CMC interactions would be one way of examining the function of the messages, as the illocutionary acts that form the basis of this analysis would portray the writers' intentions. The frequency and distribution of the speech acts, and how this varied over time, could give an indication of how the CMC users were adapting to this context. There are a few disadvantages of using Speech Act theory, some of which were touched upon in the previous section. For instance, Cohen (1984) noted that his coders encountered some difficulties in identifying speech acts in conversation. Schiffrin (1994) also comments on this difficulty, suggesting that the problem arises because "some utterances bear little surface resemblance to their underlying illocutionary force." (Schiffrin, 1994, p. 60). A related problem concerns the lack of a simple one-to-one mapping between utterances and speech acts, a point made by Bowers and Churcher (1989) and Schiffrin (1994). These difficulties could effect the reliability of a coding scheme based upon Speech Act theory.

In contrast to the other forms of analysis, Conversational Analysis examines the way in which sequences of turns are structured. The analysis tends to focus upon the local structure of pairs of utterances (such as adjacency pairs), whilst exploration of the CMC interactions will probably need to incorporate a larger unit of analysis if it is going to capture the way in which understanding is achieved in this restricted context. In summary, whilst Conversational Analysis is a useful way of exploring how people interact in everyday conversations, the emphasis on structural organisation would appear to rule it out from further considerations for the current purpose.

The Exchanges, Moves and Acts used in Discourse Analysis (Sinclair and Coulthard, 1975) are defined in terms of the interactive function they perform, so this form of analysis could provide some useful insights into the function of the CMC messages. Discourse Analysis would also facilitate the exploration of larger units of analysis, because of its hierarchical arrangement of units. There are two main disadvantages of using this scheme. Firstly, Discourse Analysis does not allow units of the same level (such as Exchanges) to be embedded within each other. Sinclair and Coulthard (1975) proposed that the initiation of a new Exchange ruled out the possibility of returning to a previous one. There is evidence, however, that embedding does occur in everyday conversations. Levinson (1983) suggests that 'insertion sequences' (Schegloff, 1972) are examples of embedding that occur frequently in conversation. The second disadvantage associated with Discourse Analysis is that it was devised to explore communication in a particular setting (classrooms); whilst it has been applied to a few other contexts, the generalisability and reliability of this form of analysis has not been extensively studied.

Conversational Acts, which is based upon a generalisation of Speech Act theory, could provide one way of analysing the CMC interactions. It was specifically designed to be a way of examining the process of grounding in the CMC dialogues, and could highlight any changes in this process that occurred as CMC users gained experience of the novel context. The first disadvantage of using this analysis is again based on the reliability of the coding scheme, which will encounter the same problems identified in the section on Speech Act Theory. In addition, identification of the turn-taking acts could also be problematical. Conversational Acts relies heavily on intonation to define and code the different turn-taking acts, and this information will not be available in the text-based CMC messages. The second disadvantage of using Conversational Acts is that it has only been applied to a small sample of dialogues, taken from the restricted domain of the TRAINS project; the reliability and generalisability of the scheme have not been addressed by Traum and Hinkelman (1992).

Finally, there are several advantages in applying Conversational Games Analysis to the CMC interactions. The analysis is based upon the function and content of utterances, and how they are used to achieve Conversational Goals. The distribution of the Conversational Games and Moves (some of which specifically highlight some of the processes of grounding), could illustrate the ways in which CMC users adapt to the context. The other main advantage of this scheme is that it has been applied to a large corpus of dialogues (The Map Task Corpus, Anderson et al., 1991a), and the reliability of the coding has been demonstrated (Kowtko, Isard and Doherty-Sneddon, 1992; Carletta et al., 1997). The only disadvantage of using this form of analysis for the current purpose, is that it has not previously been applied to written interactions. This disadvantage will probably be outweighed by the reliability of the coding scheme, and its focus upon the function and content of utterances.

Taking these factors into account, Conversational Games Analysis appears to be the most appropriate way of analysing the CMC interactions. The discourse units are defined in terms of the function speakers are trying to achieve, and the reliability of the coding scheme has been demonstrated. In addition, Conversational Games Analysis has been shown to illuminate changes in structure and content of dialogues in different communicative contexts; such as face-to-face, audio-only and Video-Mediated contexts (Doherty-Sneddon et al., 1997).

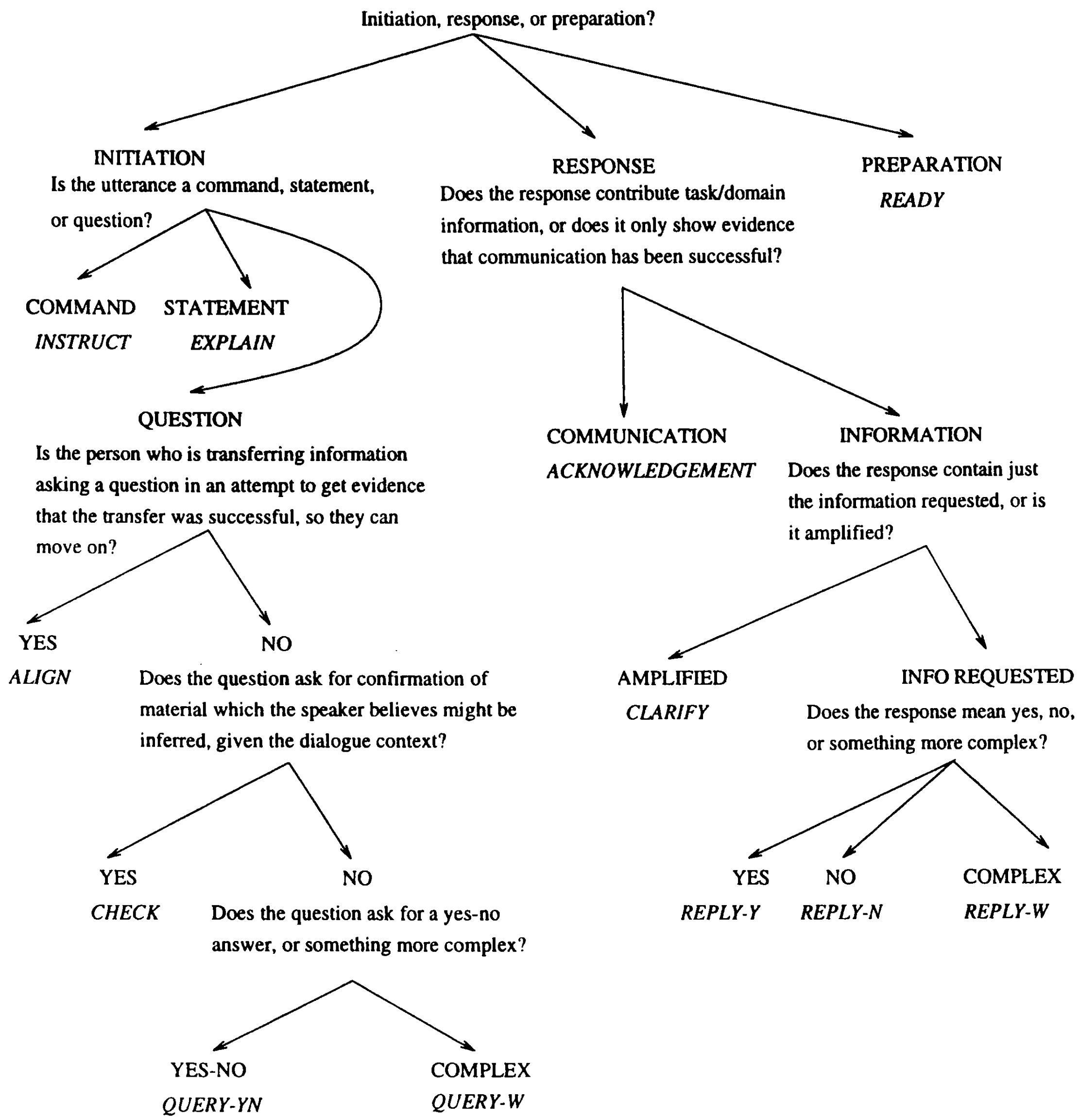
4.3 Study One, part 2: Conversational Games Analysis of CMC and Spoken Interactions

4.3.1 Conversational Games Analysis

As mentioned previously, Conversational Games Analysis examines the structure and content of dialogues on two functional levels; Conversational Games and Conversational Moves. Conversational Games consist of an Initiating Move and any other Moves required to fulfil the purpose of the Game. The coding scheme is applied by coding each utterance either as an Initiating Move or as a response; Initiating Moves which start a new Conversational Game are labelled with the name of the Game, and the beginning and end of Conversational Games are noted. Coding an utterance as a Move requires taking account of several sources of information; the semantic content of the utterance, the location of the utterance within the dialogue, prosodic and intonational information. A Conversational Move is defined by the perceived conversational function that the speaker intends to accomplish with that utterance, or utterances as a Move can contain more than one utterance.

Figure 4.1 illustrates the steps taken in deciding the Move category of utterances in Map Task dialogues. The figures is taken from Carletta et al. (1997 p. 15).

Figure 4.1 Decision tree for defining Conversational Moves in Conversational Games Analysis



The six Initiating Conversational Moves and seven response Moves are defined by (Kowtko et al., 1992) as follows:

Initiating Moves

Instruct: communicates a direct or indirect request or instruction, to be done immediately or shortly.

e.g. "You then go down south two inches."

Check: checks self-understanding of a previous message or instruction

e.g. "So you want me to go down two inches?"

Query-yn: Query-w: question (yes-no/open-ended) which asks for new or unknown detail about some part of the task;

e.g. " Do you have a rockfall"

Explain: describes status quo or position in task with respect to the goal; freely offered, not elicited; provides new information.

E.g. "I've got a cairn."

Align: checks the other participant's understanding or accomplishment of a goal; also checks alignment, attention, agreement and readiness of participants.

e.g. "Okay?" meaning "are you with me"

Response Moves:

Clarify: clarifies or rephrases what has previously been said; usually repeats given or know information; is elicited by the other person.

e.g. "South, two inches."

Reply-y, Reply-n: Affirmative (*Reply-y*) or negative (*Reply-n*) elicited response to *Query-yn, Check* or *Align*; also indicates agreement, disagreement, or denial.

e.g. "Yes I have" , or "No I can't do that."

Reply-w: An elicited reply to *Query-yn* or *Check*; can be a response to a *Query-yn* that is not easily categorised as positive or negative response.

e.g. "Down"

Acknowledge: Vocal acknowledgement of having heard and understood; not specifically elicited but often expected before the other speaker will continue; announces readiness to hear the next move - in essence a request to 'please continue'; may close a Game.

e.g. "All right" or " Oh right, I see what you mean"

Ready: Indicates intention to begin a new game and focuses attention on oneself, in preparation for the new move; an acknowledgement that the previous games have just been completed, or leaving the previous level or game; consists of a cue-word.

e.g. "Now" or "Right"

The manner in which Conversational Moves are grouped into Conversational Games is illustrated in below. In the following extract *G indicates the beginning of a Conversational Game, *M indicates the associated Initiating Move, the brackets indicate the beginning and end of Games. The role of the speakers are indicated by the following abbreviations: IG is the Instruction Giver, and IF the Instruction Follower.

Example One.

*G INSTRUCT Game

IG: Now go straight down about 5 centimetres

*M *Instruct*

*G CHECK Game embedded.

IF: Towards crane bay at the bottom?

*M *Check*

IG: Yes

*M *R*eply-yes

stop a few millimetres from crane bay

*M *Clarify*

*End of CHECK Game embedded.

IF: Okay done that

*M *Acknowledge*

*End of INSTRUCT Game

In this example the Instruction Giver initiates an INSTRUCT Game, instructing the Instruction Follower to draw the route on the map downwards for 5 centimetres. The Instruction Follower initiates a CHECK Game to check her self-understanding of the Instruction Giver's previous message, and this is embedded in the INSTRUCT Game. The Instruction Giver responds to the Instruction Follower's CHECK Game, confirming that going towards Crane Bay was the intended direction (reply-yes) and responds further to the Instruction Follower's CHECK by offering further information to clarify the earlier instruction; this closes the Instruction Follower's CHECK Game. The Instruction Follower then draws the route down to the Crane bay, acknowledging that she has done so completes the Instruct Giver's INSTRUCT Game.

Conversational Games analysis was applied to the protocols of the first and third Map Tasks undertaken in the CMC context. The dialogues were generated by the CMC system from the messages sent between pairs of participants; each message was time stamped by the computer programme, and then placed in a temporal sequence. As the CMC interactions were purely text-based, the Conversational Games Analysis had to be completed purely on the basis of the textual messages that the participants had generated.

4.3.2 Level of Analysis for this Thesis

Throughout this thesis the results of Conversational Games Analysis will be reported at the Move level. This level of analysis includes more of the available data than analysis at the Conversational Games level, which may be beneficial considering the greatly reduced amount of linguistic output in the CMC dialogues. However, Conversational Moves which are a response to Initiating Moves are excluded from the analysis as they may be associated (correlated) with specific types of Initiating Moves. For example a *Check* Move often elicits a *Clarify* response Move, and a Yes or No question frequently elicits a yes or no reply Move. Therefore, the following analysis is based upon the frequency of occurrence of Initiating Moves (*Instructs, Aligns, Query-yn, Query-w, Explains and Checks*) and any subsequent re-occurrence of these Initiating Moves within a Game.

4.3.3 Reliability of coding of CMC Interactions

In this study reliability at both levels of coding (Conversational Games and Conversational Moves) was assessed by having 2 expert coders independently code a CMC dialogue. Interjudge reliability was found to be 82% agreement at the conversational Move level, and 75% at the Conversational Games level; where coders have not only to agree that a new Game has begun but also have to agree on which of the six types of Games has been initiated. To ensure that interjudge reliability was not just due to chance factors, a Kappa coefficient of agreement for nominal scales (as described by Cohen, 1960) was also calculated at the Conversational Move level of analysis; it was found that agreement on classification of Moves was good [$K = .89$, $N = 65$, $k = 2$], indicating that there was significantly more agreement between the two expert coders than expected by chance factors alone.

Previous research using Conversational Games Analysis has reported similar levels of interjudge reliability; for example, Kowtko et al. (1992) report interjudge reliability as assessed at the Conversational Move level between expert and novice coders, and found that coders achieved agreement on 78% of the Conversational Moves. Carletta et al. (1997) carried out an extensive Reliability study on the use of Conversational Games Analysis of the Map Task corpus. Carletta et al. examined the interjudge reliability of the four experienced coders, reporting the reliability of the coding scheme at both levels of the analysis. The overall reliability of Move classification was high [$K = .83$, $N = 563$, $k = 4$], and reliability of Move segmentation was very high [$K = .92$, $N = 4{,}079$, $k = 4$] Carletta et al. (1997) noted that the reliability of classifying Conversational Games was good, but there were disagreements over where Games ended between coders; which affects where coders place the start of the next new Conversational Game. However, where the coders did agree on the beginning of a Game they also agreed on what type of Game it was [$K = .86$, $N = 154$, $k = 4$] (Carletta et al. 1997). This is one of the reasons why the analysis of the CMC interactions was carried out at the Conversational Move level of analysis, as this level of analysis has been shown to be more reliable.

The kappa coefficient obtained in the current study (κ = .89) shows that agreement between coders was very high, which indicates that the coding was reliable; differences in use of Conversational Moves in the CMC and spoken dialogues can be explored with confidence.

4.4 Results of Conversational Games Analysis of CMC and Spoken Interactions

Conversational Games Analysis was applied to the first and third trials of the Map Task undertaken in the CMC context. The full set of Conversational Games described by Kowtko et al. (1992) were found to be necessary and sufficient to code the CMC interactions. No new categories of Games or Initiating Moves were required.

4.4.1 Results of Comparison of CMC and Spoken Contexts

The first analysis of the coded interactions examines the distribution of Initiating Moves in the first trial of the Map Task by CMC users, comparing this to the distribution of Initiating Moves that occurred in the first attempt at the Map Task in the spoken (audio-only) dialogues. The same set of spoken dialogues were used for comparative purposes in chapter 3, and were taken from the HCRC corpus. The aims of the current analyses are two-fold. Firstly, to see if the distribution of Initiating Moves varied with communicative context, and whether these differences could account for the poorer task performance which occurred in the initial attempts at the Map Task by CMC users. Secondly, the analysis aims to explore any changes that occurred in the CMC interactions as novice users gained experience of the CMC context, to examine adaptations in communicative strategy over time.

Due to the significant differences in length of the CMC and spoken dialogues, the frequency with which each category of Initiating Move occurred was standardised to represent the number of Initiating Moves in every 100 turns of dialogue. The mean frequency and mean standardised frequency of Initiating Moves is presented in table 4.1, for the first CMC and spoken Map Tasks. The total mean frequency is given in the final row of the table, and the standard deviations for each mean are given in backets.

Table 4.1. Mean Frequency and Standardised Frequency (per 100 turns of dialogue) of Initiating Moves in the first CMC and Spoken Map Task Dialogues.

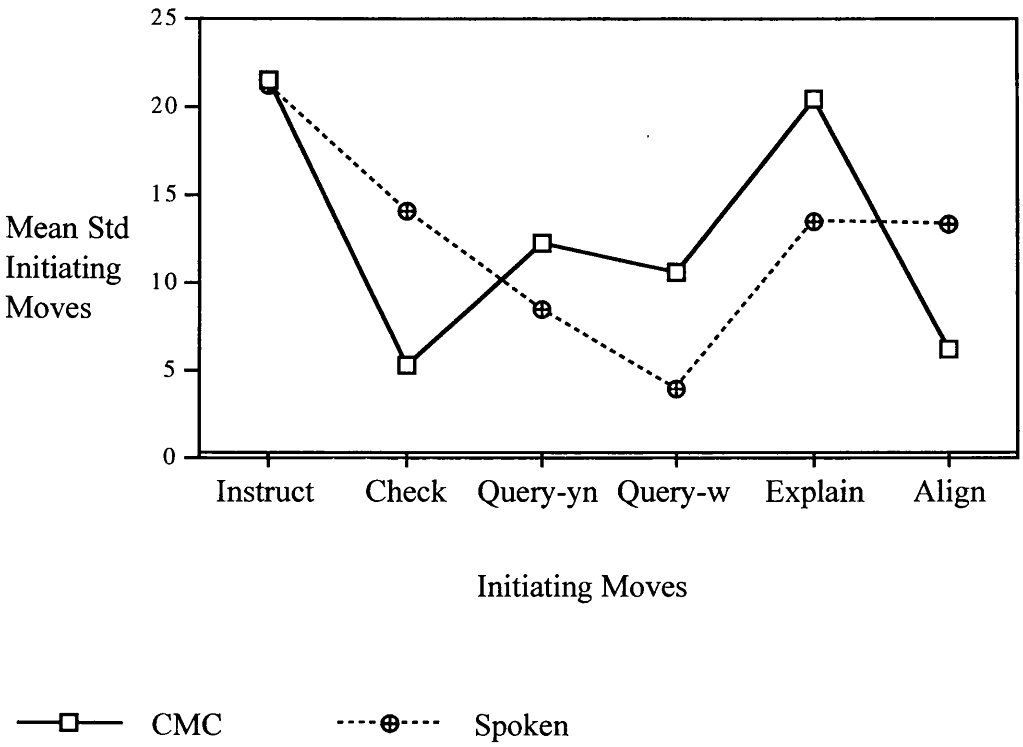
	Frequency		**Standardised frequency**	
	CMC	**Spoken**	**CMC**	**Spoken**
Instructs	16.20 (6.60)	42.90 (18.91)	21.61 (7.35)	21.26 (8.18)
Checks	5.50 (3.48)	33.80 (32.21)	5.24 (3.54)	14.07 (10.72)
Query-yn	8.40 (3.59)	15.60 (6.09)	12.24 (9.61)	8.53 (4.89)
Query-w	9.60 (7.25)	8.60 (6.48)	10.53 (5.51)	3.95 (2.85)
Explain	15.50 (9.07)	30.80 (26.27)	20.84 (17.59)	13.44 (7.92)
Align	7.00 (5.90)	29.30 (25.80)	6.17 (3.36)	13.39 (12.18)
Total	62.20	161.00	76.18	74.64

The mean frequencies (columns 1 and 2) given in table 4.1 show that fewer Moves were initiated during the CMC interactions than in the spoken dialogues. This is probably due to the fact that the CMC interactions contained significantly fewer turns, and words, than the spoken dialogues. The opportunity to initiate Conversational Moves was significantly reduced in the CMC context. However, once the length of the CMC and spoken dialogues are taken into account, there appears to be little difference in the mean overall number of Initiating Moves per 100 turns of dialogue.

The data in table shows that some categories of Initiating Moves were used more frequently than others. This is to be expected given the nature of the Map Task, and has been commented upon in previous research (Doherty-Sneddon, 1995; Kowtko, 1997; Doherty-Sneddon et al., 1997). For example, *Instruct* and *Explain* Moves occurred significantly more frequently than any other Initiating Moves. *Instruct* and *Explain* Initiating Moves accounted for approximately 28% and 23% respectively of all Initiating Moves; so more than half of the total number of Initiating Moves involved one participant telling the other participant where to draw the route on the Map, or offering un-elicited information. Other Initiating Moves occurred less often, for example, Open-ended questions (*Query-w* Initiating Moves) occurred less frequently, accounting for 10% of all Initiating Moves.

More interestingly, the standardised frequency of some categories of Initiating Moves appears to vary with context; in other words there appears to be an interaction between communicative context and frequency of use of some categories of Initiating Moves. For example, the standardised data suggests that CMC users initiate proportionally fewer *Check* Moves, but proportionally more *Explain* Moves and *Query-w* Moves. This is illustrated in Figure 4.2 below.

Figure 4.2 Graph of the Interaction of Communicative Context and Proportional Initiation of Moves.



To test the significance of these observations, and to observed any changes that occurred over time as the CMC users adapted to the novel context, a series of 2 way Analysis of Variance (ANOVA) were computed on the standardised frequency scores for each category of Initiating Move. In each of these mixed design ANOVA Communicative context (CMC *versus* spoken) was treated as a between subject factor, and task order as a repeated measure.

The data (using the standardised scores) for these ANOVAs is presented in Table 4.2 below, the standard deviation for each mean are given in brackets.

Table 4.2 Mean Standardised Frequency of Initiating Moves in first and last CMC and spoken tasks

	CMC		Spoken dialogues	
	CMC task 1	CMC task 3	Spoken task 1	Spoken task 3
Instructs	21.16 (7.35)	25.74 (11.41)	21.26 (8.18)	19.60 (5.94)
Checks	5.24 (3.54)	5.72 (2.85)	14.07 (10.72)	9.89 (4.70)
Query-yn	12.14 (9.61)	9.74 (5.68)	8.53 (4.89)	8.47 (4.06)
Query-w	10.53 (5.51)	4.48 (2.75)	3.95 (2.85)	3.61 (2.30)
Explain	20.84 (17.59)	14.15 (8.85)	13.44 (7.92)	8.28 (3.25)
Align	6.17 (3.36)	6.59 (4.20)	13.39 (12.18)	10.00 (6.69)

Instruct Initiating Moves

A 2-way ANOVA on the proportional use of *Instruct* Moves, with communicative context as a between group factor and task order as a repeated measure revealed non significant main effects ($p>0.1$). A significant interaction between these two factors was observed [$F(1,18) = 7.49$, $p<0.05$] indicating that the proportional use of *Instruct* Moves changed over time. Further analysis by Simple Effects showed that there was a significant increase in the use of *Instruct* Moves by CMC users [$F(1,18) = 8.05$, $p<0.01$], but the proportion of *Instruct* Moves did not differ over tasks in the spoken context.

Check Initiating Moves

A 2-way ANOVA on the proportional use of *Check* Moves, with communicative context as a between group factor and task order as a repeated measure revealed a significant main effect of context [$F(1,18) = 13.52$, $p<0.01$]. Users of the CMC system initiated *Check* Moves less frequently (almost 63% fewer *Check* Moves) than participants in the spoken context. The main effect of task order and the interaction between context and task order were non-significant ($p>0.10$).

Query-yn Moves

A 2-way ANOVA on the proportional use of *Query-yn* Moves, with communicative context as a between group factor and task order as a repeated measure, revealed non significant main effects and interaction effects ($p>0.1$).

Query-w Moves

A 2-way ANOVA on the proportional use of *Query-w* Moves, with communicative context as a between group factor and task order as a repeated measure, revealed a significant main effect of communicative context [$F(1,18) = 10.55$, $p<0.01$]. Participants initiated more than twice as many *Query-w* Moves in the CMC context than in the spoken interactions. The main effect of task order was also significant [$F(1,18) = 8.16$, $p<0.05$], with more *Query-w* Moves being initiated in the first trial of the task than the last trial (means being 7.24 vs 4.04 respectively). A significant interaction between communicative context and task order also occurred [$F(1,18) = 6.50$, $p<0.05$]. Further analysis by Simple effects revealed that users of the CMC system initiated more than twice as many *Query-w* Moves in the first Map Task than participants in the spoken context [$F(1,18) = 14.61$, $p<0.01$]. Simple effects analysis also showed that there was a significant reduction (50% reduction) in the use of *Query-w* Moves by participants in the CMC context over time [$F(1,18) = 11.24$, p0.01]. By their third attempt at the Map Task the frequency of *Query-w* Moves did not vary with communicative contexts ($p<0.10$).

Explain Moves

A 2-way ANOVA on the proportional use of *Explain* Moves, with communicative context as a between group factor and task order as a repeated measure, revealed non significant main effects ($p>0.1$). The interaction between context and task order was also non-significant ($p<0.10$)

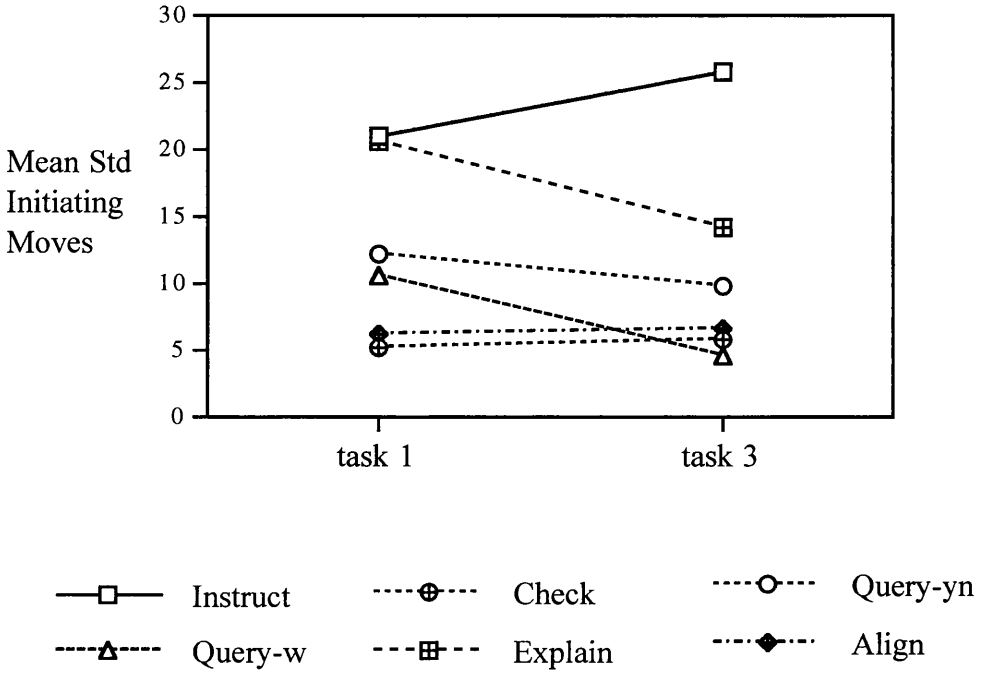
Align Moves

A 2-way ANOVA on the proportional use of *Align* Moves, with communicative context as a between group factor and task order as a repeated measure, revealed non significant main effects ($p>0.1$). However, there was a trend ($p<0.07$) for more *Align* Moves to be initiated in by participants in the spoken context than in the CMC context (11.70 vs 6.38 *Align* Moves per dialogue). The interaction between context and task order was non-significant ($p<0.10$).

4.4.2 Summary of Results of Conversational Games Analysis

The results obtained from the analysis of Conversational Moves illustrate two points. First, communicative context had a significant effect on the proportional use of some Conversational Moves; users of the CMC system initiated fewer Check Moves but proportionally more Query-w Moves (these main effects of context are illustrated graphically in Figure 4.2 above). Second, as participants adapted to the novelty of communicating in a CMC context proportional use of some Initiating Moves; for the proportional use of Instruct Moves increased over time, whilst the frequency of Query-w Moves declined. The affect of adapting over time to the CMC context is illustrated in Figure 4.3 below.

Figure 4.3 Adaptation over time in the CMC context.



4.5 Discussion: Results of Conversational Games Analysis

The analysis has shown that participants in the CMC dialogues used a similar range of Conversational Moves and Games as participants in the spoken context. However, the shorter CMC interactions contained fewer Initiating Moves than the spoken dialogues. The reduction was not a uniform decrease across all Move types. Some types of Moves were used less frequently in CMC than in speech, demonstrating that the communication context did have an impact on the types of Conversational Goals attempted in the dialogues.

The main differences between the use of Initiating Moves in spoken and CMC interactions were primarily a decreased use of the Initiating Moves used to establishing mutual understanding: *Align* Moves and *Check* Moves. Speakers use *Check* Moves to check that they have correctly understood something the other person has said or written. This interactive ways of establishing mutual understanding occurred less frequently in the CMC context, which may partly explain why CMC users performed the Map Task less accurately than participants in the spoken dialogues. CMC participants also tended to make less use of *Align* Moves, which are used to check the listener's understanding of a previous message, to ensure that both participants were in agreement over the position of a landmark, or were ready to continue with the task.

According to the Collaborative model of communication (Clark and colleagues) effective communication is achieved when participants establish mutual understanding, through the process of grounding. In the CMC interactions, there appears to be very little checking of mutual understanding; CMC users make little use of the interactive means of grounding. Previous research has shown that economising on the amount of checking of understanding is a potentially risky communication strategy; this has been demonstrated in the child and adult spoken dialogues (Anderson and Boyle, 1994; Anderson et al., 1994). The low levels of checking that occurred in the CMC Map Tasks could explain why the CMC participants performed less well in the initial Map Tasks. The question then is, how did the CMC users manage to complete the later tasks so accurately? What adaptations did they make to the way they communicated as they gained experience of the CMC context? How did they manage to communicate more effectively in the later Map Tasks?

The results of the Conversational Games Analysis showed that the distribution of Initiating Moves changed over time; by task 3, participants in the CMC context initiated more *Instruct* Moves but fewer *Query-w* Moves. Could these changes, or adaptations to the CMC context, account for the observed improvements in task performance?

One possible answer to these question is suggested by the Collaborative model of communication, and the work on grounding in different contexts by Clark and Brennan (1991). CMC users may have adapted to the slower pace of the written context and taken greater care in the way they constructed their messages. The more leisurely pace of the written interactions could have enabled the participants to plan, and edit, their contributions before they transmitted them to their partner; resulting in shorter and more precise contributions.

Clark and Brennan (1991) predicted that participants should use the grounding techniques afforded by a medium which require least collaborative effort. The grounding constraints available in the CMC system used in this study would be the constraints of 'reviewability', ' revisability' (Clark and Brennan, 1991). As written messages have more permanence than spoken utterances, the CMC users would be able to review (re-read) incoming messages; at least for a short time, until further messages removed them from the screen. This would mean that complex instructions could be deciphered in CMC. Equally, CMC users were able to revise messages before they posted them, which would facilitate the composition of more precise and concise instructions. In summary, it is suggested that the grounding constraints afforded by the CMC context could have enabled CMC users to review and edit their messages before transmitting them, and so produce shorter, more precise instructions and communications. Was there any support for this view in the results of the Conversational Games Analysis?

It has already been noted that the main difference between the first and third tasks was the increased initiation of *Instruct* Moves. This effect is mainly due to the Instruction Giver, who plays the more dominant role in the Map Task. The CMC and spoken dialogues were searched, and all Initiating Instruct Moves were extracted for further investigation. Several differences were observed between the written and spoken instructions. The CMC instructions appeared to contain a greater number of exact measures of distance (inches, centimetres and even millimetres) than the spoken dialogues. There also appeared to be some very precise ways of stating direction in the CMC dialogues. Examples of the preciseness of instructions in the CMC dialogues are given below. The examples are taken from different pairs of participants in the CMC context. In these extracts the role of the speakers is indicated as follows; Instruction Giver (IG) and Instruction Follower (IF).

Examples of precise instructions in the CMC context

Extract 1

IG: From bottom left of highest viewpoint go 1 cm to left and 2.5 cm down

Extract 2

IG: Is there anything to left of cottage?

IF: No

IG: OK then draw a line about 6 cm long at a downward slope of 45 degrees to the left.

Extract 3

IG: Make a mark 1 cm directly above the left end of the bakery, now draw a curved line, like a semi circle, and stop at the mark

IF: Okay done that.

Extract 4

IG: Start at Forest Fire and go to Picnic Site which is south, distance is 10 to 12 cm.

Extract 5

IG: Are you near the Collapsed Shelter?

IF: No, 3 cm south of Bridge

IG: Go east 7 cm.

Extract 6

IG: You come around the left side of the Forest Site and proceed south until you are south west of the play area on a 180 degree scale.

In contrast, the spoken instructions were frequently rather imprecise. The following extracts from the spoken Map Tasks illustrate this point. In these extracts the role of the speakers are designated by the initials IG and IF.

Examples of Imprecise Instructions in the Spoken Context.

Extract 7

IG: Go round underneath the cattle stockade and curve up on a wee kind of hill.

Extract 8

IG: from the start you're gonna go left and you're going to go down, down you like you're gonna go across a bit and down to the bottom of the extinct volcano

Extract 9

IG: Right, you're going to have to, like, curve out to the right as though there's something there, because I've got farmed land here, so

IF: right . Is it a sort of straight line out, or just a , you know, to the right or just curved

IG: Eh, just sloping very slightly.

Extract 10

IG: Go past the river, a few centimetres past the river.

As these examples show, Instruction Givers in the CMC context used exact measures of distance and direction in their instructions to the Instruction Followers. Precise indications of distance were also found in the spoken instructions, but not as frequently. In general, the spoken instructions appeared to be less precise than the written ones. To see if there was a significant difference in the precise nature of instruction giving between the two contexts, the *Instruct* Moves from the CMC and spoken Map Task dialogues were searched, and the number of Instruct Moves that contained at least one exact distance measure was calculated. These scores were then standardised to allow for the difference in length of the CMC and spoken dialogues. In this instance, the standardisation was the proportion of *Instruct* Moves which contained precise distance measures. The standardised scores for the first and third CMC and spoken tasks were calculated.

An ANOVA was performed on the standardised scores, with communicative contexts as a between groups factor, and task order as a within subject pair repeated measure. The analysis showed that significantly more of the CMC *Instruct* Moves contained precise measures of distance [$F(1,18) = 4.75$, $p>0.05$]. On average 43% of *Instruct* Moves in CMC contained at least one precise distance measure, compared to 17% of *Instruct* Moves in spoken Map Tasks. A main effect of task order showed that the proportion of *Instructs* with precise distant measures also increased over time

[F(1,18) = 8.865, p<0.01]. The interaction of context and task order was non-significant [F>1]. However, further analysis of the means involved in this interaction was performed, to see if the significant effect of task order was due to changes in both communicative contexts. Simple Effects analysis showed that task order had a significant effect in CMC [F(1,18) = 6.87, p<0.05], but not in the spoken dialogues. CMC users made greater use of precise distance measure than participants in the spoken context, and they increased the use of these distance measures as they gained experience of CMC.

In order to determine if the increased use of explicit distance measures over time could explain the observed improvement in task performance by users of the CMC system, a correlation was computed between the number (non-standardised) of precise distance measures and route accuracy scores for the first and third CMC Map Tasks. A Spearman rank correlation coefficient was computed, but the result was non-significant (rho (df 20) = -.29, p>0.1). The accuracy of the routes drawn by the Instruction Followers in the CMC context could not be shown to be associated with the increased use of explicit measures of distance by the Instruction Giver.

Despite this non-significant result, the use of precise instruction giving in the CMC context seems to be a sensible way of adapting to the restrictive nature of this text-based communicative context. It may take less collaborative effort to ground precise instructions in this written context than the rather under-specified messages often found in spoken interactions; this would reduce the number of turns (or cycles of presentation and acceptance phases) required to establish mutual understanding.

4.6 Conclusions

These results support the view that CMC participants were adopting a strategy of using more precise and concise instructions. These precise instructions would be easier to construct in the slower paced CMC context than a spoken dialogues, as the text-based messages could be edited and reviewed before being transmitted; a point predicted by Clark and Brennan (1991). Whilst the slower speed of composition in CMC is a communicative cost, it can be used beneficially. The CMC context provided participants time to plan clearer, less ambiguous instructions and messages than are usual found in spoken dialogues, where there is a constant pressure to keep the interaction going. The result is that spoken messages in general, and instruction giving in particular, are often rather unspecified and potentially ambiguous. Speakers overcome these problems by the iterative process of checking and reformulating contributions, as this kind of interactive sequence is a low cost option in speech. In text-based communications, such as CMC, checking and recasting messages is a relatively effortful procedure. Explicit checking of previous contributions and their intended interpretations is less frequently undertaken, and may be required less often if the contributions are more precise.

These results suggest that the participants in the CMC context adapted to the novel environment in an appropriate manner. CMC users compensate for the limitations imposed by the context by producing instructions which are more precise and concise, making good use of the slower pace of the interactions. By making these adjustments to the CMC context they eventually achieved task outcomes comparable to task performances of spoken dialogues.

This study has demonstrated the impact of a CMC context upon task performance and the process of communication. The study has shown that CMC users adapted quickly to the novel context, and demonstrated the flexibility of human communicators. However, it should be noted that the CMC users commented on the awkwardness of the context, and the irritatingly slow process of communicating in this medium. The next two studies investigate how people complete similar tasks in a less restrictive communicative context, Video-Mediated Communication which enables people to interact in a context that more closely resembles normal conversation.

Chapter 5. Video-Mediated Communication: effects of constraining the audio channel of communication

The first study reported in this thesis demonstrated that people could collaborate using a text-based form of computer mediated communication. Effective use of CMC, however, required participants to adjust their style of collaboration to suit the communicative context. One of the major constraints of CMC was highlighted in study one; the length of time it takes to complete a task in a written medium. Whilst the slower pace of CMC appears to facilitate adoption of a more precise and concise form of communication, it was very clear from the comments made by participants that they found text-based CMC to be a tedious and frustrating way of interacting.

Recent advances in communication technology now facilitate spoken and visual channels of communication, such as video-conferencing systems or Video-Mediated Communication (VMC). Two of the major questions being addressed in the next part of this thesis are how similar is VMC to face-to-face communication, and what, if any, communicative adjustments do people make when they collaborate in this context? Study 2 begins to explore the effects of VMC upon communication and collaboration by concentrating on one of the benefits of VMC, the ability to use spoken interactions. Study 3 will concentrate on the effects of the visual channel in VMC. A general introduction to VMC will help to set the scene for both these studies.

5.1 Introduction to Video Mediated Communication

In this thesis VMC is taken to indicate communication using both the visual and the audio channels, which are transmitted and presented via computer terminals and networks. VMC may also include the use of shared screen facilities, which either or all of the participants can write to, or edit etc. Other terminology found in the literature in this area includes terms such as 'Multimedia' and 'Computer Supported Cooperative Work' (CSCW). 'Multimedia' can be broadly defined as "two or more remotely located people electronically sharing audio, video, and data via either desktop PC or a group room system" (Fussel and Benimoff, 1995, p. 229). The term 'CSCW' is more generally used when referring to a group of people (i.e. more than two) communicating over a distance. It should be noted that not all CSCW systems provide a video channel. CSCW has been defined as the way "people work together to some purpose, using computer technology" (Connolly and Pemberton, 1996, p. 2). It is apparent from these definitions, that Multimedia and CSCW can be synonymous with the term VMC. For the sake of clarity, the term VMC will be used in this thesis as the research reported here is based on dyadic, rather than group, interactions where the visual channel is always present.

5.1.1 Expanding interest in the Application of VMC

Recent literature shows that VMC is now being used for a wide range of applications. For example, VMC is being used in distance education, the provision of health services to peripheral areas, and facilitation of group meetings in the research and business sectors. Some examples of the diverse use of VMC are described below.

Educational uses of VMC

VMC is currently being used in a range of educational distance learning settings. Examples of such uses include Wood and Parham (1996) who describe the application of live interactive video-conferencing as a means of educating people in the health profession about gerontology and geriatrics; more than 22 thousand people took part in a series of videoconferences. Similarly, Knowles and Dillon (1996) found that video-conferencing was an effective way of informing American architects about new legal requirements; the example given is of 19 thousand people who attended one of three satellite video-conferences, which explained the requirements set out in the 'Americans with Disabilities Act'. Robinson (1993) details the use of VMC for in-service training of teachers in three different areas of Northern Ireland; despite some technical problems, the video conferences enabled both educational and community co-operation.

Health

Video-conferencing has been found to be a useful means of disseminating information in the area of public health. Reiss et al. (1996) examined the use of an interactive video-conference, which was broadcast to Public health nurses in eleven USA states. The aim of the video-conference was to increase the nurses' awareness of a Public law act (Public Law 99-457 Part H), which deals with the special health requirements of infants and young children. Furnace et al. (1996) explored the potential of video-conferencing to relay lessons to medical students of Royal Aberdeen Children's Hospital, who were on placement at an Inverness Paediatric unit. A different application of VMC is described by Tröster et al. (1995), who investigated the effectiveness of provision of neuropsychological services to rural areas, using a video-conference system run over digital telephone lines. The VMC system enabled clients to be interviewed 'on line' by a neuropsychologist based in the Kansas Medical Centre, who decided which psychometric tests were required for a diagnosis. The tests were administered and scored by the rurally-located psycho-metricians who then used VMC to transmit the scores back to Kansas for evaluation by the trained neuropsychologists.

All of these studies reported that the application of VMC was successful. These innovative uses of VMC prove to be very cost-effective ways of disseminating knowledge and expertise. The video-conferences also provided remotely located health workers a means of networking and interacting with colleagues.

Group meetings

Another popular application of VMC is the expanding use of this context as a means of communication between remotely located group members. Lehrman (1991, in

Button & Maggi, 1995) carried out a survey of the use of VMC in a sample of USA commercial organisations, and found that this had doubled between 1985 and 1991.

The research by Gale (1990), O'Conaill et al. (1993), Isaacs and Tang (1994), Sellen (1992; 1995) which was described earlier are all examples of research into the effect of VMC upon group interactions. Other examples are also found in the literature. For instance Olsen, Olsen and Meader (1994) examined the use of VMC (amongst other forms of mediated communication) by small groups of people who were involved in a design problem. Olsen et al. point out the advantages of using VMC and other mediated contexts; whilst these contexts may not replace face-to-face interactions they allow organisations '...additional flexibility in how teams are structured and deployed' (Olsen et al., 1994, p. 33).

Video Mediated communication systems are increasingly being used as an alternative to face-to-face interactions. Heath and Luff (1991; 1992) report on the use of a multi-media office environment used by researchers at Cambridge EuroPARC for informal interactions and collaborative work. They found that this VMC system changed the way which participants used non-verbal and verbal communication. Analysis of video-recorded naturalistic VMC interactions showed that the use of non-verbal communication was less effective in this context; for example, participants found to difficult to initiate a conversation or controlling turn-allocation using gaze, and had to make more exaggerated body and hand movements to indicate gestures in VMC than is normal in face-to-face interactions. Heath and Luff suggest that VMC introduces certain communicative asymmetries into interpersonal interactions, as the impact of visual conduct is reduced users rely more heavily upon the vocal channel of communication.

In summary, this section has demonstrated the expanding use of VMC in various communicative settings. VMC can be an effective form of communication, in that it can facilitate dissemination of information to a large number of people, or discussion between groups of people, without the cost of travel expenses. Whilst discussions and transfer of information can be achieved by text-based CMC, VMC has probably attracted increased interest of late because it offers users the opportunity to both see and hear other participants. VMC more closely resembles the communication context that we are most familiar with, face-to-face conversation.

5.2 VMC versus Face-to-face Communication

It has often been assumed that the provision of both visual and audio channels will provide a communicative context that is similar to face-to-face interactions; for example, Sellen (1995) reports that "the explicit goal of video-conferencing is often stated as one of simulating face-to-face meetings" (p. 403). Although VMC affords both visual and spoken channels of communication it cannot replicate face-to-face interactions, as the participants do not share the same local space. It is claimed that a sense of 'social presence' is hard to achieve when you can only see the head and shoulders of other interlocutors (O'Malley et al., 1996).

So what impact does VMC have upon communication and collaboration? How closely can VMC resemble face-to-face interactions. At present it is difficult to answer this question, despite the growing body of research. As stated earlier, this could be due to the different approaches used by researchers to evaluate the effect of VMC and how it differs from other communicative contexts.

5.3. Previous Literature on Effects of VMC upon Communication and Collaboration

The early comparisons between different communicative context (such as, Ochsman and Chapanis, 1974; Williams, 1977; Cohen, 1984) included some mediated communication systems that were analogous with VMC, such as Closed Circuit TeleVision (CCTV). These studies concluded that the task performance was not effected by communicative contexts, as long as the task was of a collaborative nature. The process of communication was not greatly effected by the absence or presence of a visual channel, but availability of a spoken channel significantly altered the quantity of linguistic output and the amount of time it took to complete collaborative tasks; dialogues in spoken contexts contained more words, but were accomplished in less time, than interactions in writing. However, as these early comparisons were made "in the absence of advanced computer technology" (Masoodian et al., 1995, p. 238) it may be inappropriate to relate this earlier work to current technologically mediated communication contexts.

More recent research into the effect of VMC includes the studies mentioned earlier (in chapter 2) by Sellen (1992; 1995), O'Conaill et al. (1993), Isaacs and Tang (1994) and Gale (1990). Another example of a naturalistic study is described by Fish, Kraut and Chalfont (1990), who examined the extent to which informal conversations in VMC and face-to-face were similar. The VMC system used in this study was Bellcore's VideoWindow (VW) which provides life-size images and high quality audio signals. Social interactions over VW were assessed from video-taped recordings and questionnaire data. The results showed that participants found it more difficult to initiate and maintain conversations over VW. Fish et al. suggest that this was partly because VW system did not provide full 'reciprocity' of audio and visual signals.

Anderson et al. (1996) report the results of three experimental studies of collaborative problem solving, where communication between pairs of participants is supported by a video link and shared multimedia tools. This VMC context was compared to face-to-face and audio-only interactions in terms of the task performance, the process of communication, decision making and user satisfaction. Anderson et al. found that task performance was not affected by communicative context. However, analysis of the process of communication and decision making showed that VMC did not deliver the same benefits as face-to-face interactions; the addition of a video channel did not improve communicative efficiency or more optimal decision making. Further details of the research by Anderson et al. will be presented in chapter 6, as this work forms the basis for Study 3.

The research by Anderson et al. is unusual, as it is one of the few studies of VMC which has attempted to examine the relationship between the process of communication and task outcome. Many other studies of this context have tended to focus on either the product of the collaboration (the task outcome) or the process of communication (Isaacs and Tang, 1994). To capture the interaction of these two factors a multi-dimensional approach to the evaluation of VMC is required. The usefulness of such an approach is seen in the work of Boyle et al. (1994), who used the HCRC Map Task to compare face-to-face and spoken-only communicative contexts. This study reported that similar levels of task success were achieved in both contexts. If the process of communication is also considered, then significantly more dialogue was required to achieve the same level of success in the spoken condition, and the structure of dialogues also varied with communicative context (Boyle et al., 1994). Furthermore, Doherty-Sneddon et al. (1997) found that visual signals (such as gaze) could be substituted for verbal signals without penalising task outcome. The point is, if only one dimension of communication outcome is analysed then subtle differences caused by communicative context may go unnoticed. This may be important when investigating the impact of VMC upon collaboration and communication, as this context appears to share many of the features of face-to-face communication and differences between contexts may be quite subtle.

A full appreciation of the impact of VMC, or other forms of computer mediated communication, can best be achieved by assessing how well the cooperative task was accomplished, accompanied by an analysis of the way the result was achieved; that is, in terms of how the communication process was structured and how the co-operation was managed. In addition, the subjective assessment of users' satisfaction with the communication process and task outcome can provide useful additional information. The benefits of using subjective data are illustrated by Tang and Isaacs (1993) and Sellen (1995), who found that users of conferencing systems preferred to communicate when the visual channel was made available.

5.3.1 The Audio Channel in VMC

Whilst research has shown that people prefer to see who they are talking to, it has also been acknowledged that "audio is relatively more important than video in supporting collaboration" (Tang and Isaacs, 1993, p. 194). In a series of studies covering a range of multimedia systems (video-conferencing rooms and table-top VMC) Tang and Isaacs found that the quality of the audio channel can have a profound effect upon communication. This is especially so if the audio and video signals are not synchronised. Problems in synchronising the visual and audio signals are chiefly a result of the fact that video images contain more bits of information than audio signals. Since it takes more time to digitise and compress a video image than an audio signal (Angiolillo et al., 1997) it is difficult to keep the two sets of signals together when they are transmitted over the network. Unless the audio signal is delayed it will arrive ahead of the visual image, resulting in the two signals being out of synch. One answer is to build-in a delay in the audio channel so that the signals arrive at approximately the same time. Nonetheless, it is difficult to get the synchronisation perfect. Isaacs and Tang (1994) observed that deliberately delaying the audio can make it difficult for VMC users to make use of the information carried by the visual channel; "small delays in the audio can disrupt the participants' ability to reach mutual understanding and reduce their satisfaction with the conversation." (Isaacs and Tang, 1994, p. 68). VMC users also commented that they found it more frustrating to cope with delay in the audio signal than a lack of synchronity between the visual and audio channels (Isaacs and Tang, 1994).

Audio delay is one way in which the audio channel in VMC differs from the spoken channel in familiar face-to-face interactions. Another difference concerns the type of network being used to transmit the audio signals, and whether this can support full duplex or ('open channel") or half-duplex audio. Full-duplex channels can create more technological problems than half-duplex channels, but both of these forms of audio have benefits and drawbacks. One of the main problems with using open channel audio is that it is can suffer from the effects of 'echo' and 'feedback' (O'Conaill et al., 1993). These problem can be overcome by using uni-directional microphones, and by 'dampening' the rooms in which the video-conferencing occurs. Whilst this form of audio may be more problematic and expensive to set-up, it does allow people to talk to each other as if they were conversing over the telephone or face-to-face.

Many video-conference systems use half duplex audio tools, which transmit only one voice at a time. The chief problems associated with this type of audio tool are that users need to take more care about how they interact. For instance, greater care and attention is required to manage turn-taking procedures or interruptions will occur frequently. Some VMC systems provide half-duplex audio tools which are 'voice activated'; in these systems the first person to talk gains control of the audio channel. When other participants want to speak, they have to wait for the first speaker to come to the end of their turn, or for a pause in the speech, before they can take-over the audio channel. Some of the problems associated with this kind of system are described by O'Conaill et al. (1993) who found that a half-duplex audio channel resulted in participants using longer turns, with fewer interruptions and backchannels when compared to interactions over full duplex audio (supported by VMC or face-to-face).

An alternative method of controlling turns in a half-duplex audio channel is for participants to manually activate the audio channel. For example, in some VMC systems an audio-tool dialogue box is presented on the monitor, and speakers activate this by clicking in the audio tool box with the mouse. Participants have to hold down the mouse button whilst they are speaking to keep control of the audio channel. This way of controlling the a half-duplex audio channel is often referred to as 'click to speak'. Despite having to be manually operated, a 'click to speak' audio tool is often preferred because voice-activated switching can be disrupted by back-ground noise. Machines cannot distinguish between intended contributions to the conference and side-conversations or background noises (such as sneezes or loud machinery) and therefore control of the audio channel can be switched unexpectedly between speakers. This results in a high degree of unpredictability in voice switched audio conferences. A speaker may begin to say something and then find that the channel has been taken from them, and part of their contribution appears to be 'lost'. There are also occasions when background noises can be loud enough to cause the channel to switch away from the current speaker, before that person has completed their turn (O'Conaill et al., 1993).

In summary, currently the audio channel in many VMC systems may not provide the quality of audio we are accustomed to in speaking over the telephone or in face-to-face interactions. Delay in the audio signal, a lack of synchrony between the audio and visual signals, and the use of click to speak audio tools are all factors that can effect the quality of the audio link between participants.

5.3.2 The Visual Channel in VMC

It is most often the visual channel provided by VMC systems which intrigues novice users of a video-conference systems; even when the quality of the visual signal is not of a high standard, as was often the case in earlier research. There are a number of factors which effect the quality of the video signal; these include the temporal resolution (the rate at which the video frames are updated, or 'frame rate') of the video image, size of the image, spatial resolution, type and capacity of the computer being used to display the images and the type of network being used to transmit the image.

In a series of experiments to determine the effects that the quality of video images have upon communication, Barber and Laws (1994) found that temporal resolution had the largest impact on the quality of the visual signals; people found it increasingly difficult to make out what others were saying as the frame rate was reduced.[1]

An idea of the impact that these factors have upon the quality of the visual image can be obtained by comparing the frame rate provided in broadcast television with many early VMC systems; broadcast television usually has a temporal resolution of 25-30 frames per second, whilst early VMC systems frequently had a frame rate of 5 frames per second (Watson and Sasse, 1996). The low visual refresh rate afforded by many early VMC systems resulted in video images which had a stilted appearance, as if the image had frozen and then moved on again.

Improvements in methods for compressing and transmitting visual signals, along with increased available bandwidth on computer networks, has meant that some VMC systems can now deliver images of similar quality to television broadcasts. Many VMC systems in current use, however, are restricted in the service they can deliver to users, who are faced with systems that provide low frame rate video signals and half-duplex audio channels. What impact would these characteristics of VMC have upon the process of communication and collaboration?

5.4 Establishing Mutual Knowledge in VMC

For people to collaborate, they need to communicate effectively. Clark and colleagues state that effective communication is achieved through the process of grounding (Clark and Wilkes-Gibbs, 1986; Isaacs and Clark, 1987; Clark & Schaefer, 1987; 1989). The process of grounding ensures that each participant has understood previous

[1] Further details of the experiments by Barber and Laws (1994) will be presented in Chapter 6, where the quality of the visual channel provided by VMC systems will be discussed in greater detail.

contributions to the conversation to a level sufficient for their current purpose. Clark and Brennan (1991) propose that the way in which grounding is achieved will be determined by the purpose of the conversation. More importantly in the context of this thesis, the process of grounding also changes with the medium of communication. Clark and Brennan (1991) suggest that different communication contexts can be described in terms of the grounding constraints that they afford.

Two of these grounding constraints have already been discussed and illustrated in Study 1 of this thesis. Participants adapted to the restraints of CMC by changing the way in which they established mutual understanding. Since it took more collaborative effort to check that a previous message had been understood in this context, participants wrote messages that were as precise and unambiguous as possible. This concise style of communication was facilitated by the relatively slow pace of CMC interactions, which enabled users to edit and review messages as required. In other words, in Study 1 the CMC users made effective use of the grounding constraints afforded by this context which have been labelled by Clark and Brennan (1991) as the constraints of 'reviewability' and 'revisability'.

Within the framework suggested by Clark and Brennan, the process of grounding in VMC should be facilitated by the grounding constraints of visibility, audibility, cotemporality, and sequentiality. That is, participants will be able to make use of the visual and spoken channels to ensure that they have understood each other. The quality of the video and audio links provided by a VMC system may restrict the use of these grounding constraints. For example, low quality video images may allow participants to see each other, but may not afford the use of gestures, eye contact or facial expressions to signal attention or agreement. The audio channel in VMC could also effect the process of grounding, as restricted access to the audio channel (for instance, using half-duplex audio channels) has been shown by O'Conaill et al. (1993) to change the way in which conversations are structured. It may be more difficult to establish mutual understanding when speakers use longer turns and listeners provide fewer backchannels.

With these grounding constraints in mind, we would predict that participants will need to expend greater effort to establish common ground in VMC than in face-to-face interactions; VMC will impose different costs of grounding than other communicative contexts. A couple of examples will illustrate this point. Firstly, it takes more time and effort to set up a video-conferencing session than to start a face-to-face conversation; video-conferences often have to be arranged in advance to ensure that all of the participants are available at the same time. Secondly, it takes more effort to manage turn-taking and avoid episodes of overlapping speech in VMC, especially when the channels of communication lack synchronisation. Thirdly, the quality and scope of the visual image provided by many VMC systems can make it difficult for participants to use non-verbal communication (such as gesture) to draw attention to an object, or allocate the next speaker turn. These three examples illustrate the kind of 'costs' that may be incurred when grounding contributions in VMC. These costs are referred to by Clark and Brennan (1991) as 'start up costs', 'speaker change costs' and 'display costs'.

If the methods of establishing common ground vary between VMC and face-to-face contexts, then there should be differences in the way that the interactions are structured. Some of these changes will be observable by comparing the structure and content of the dialogues. The impact that impoverished visual signals have upon communication and collaboration will be discussed in greater detail in chapter six. In this chapter the impact of two types of audio channel, full duplex and half-duplex will be explored to see how restraining access to the spoken channel effects collaborative communication. Before introducing Study 2, some of the relevant literature will be described.

5.5 Comparison of Different Audio Tools in VMC Systems

A review of the literature shows that there appears to have been very little research comparing the effects of the different audio tools provided by VMC systems. This seems to be a historical fact as Williams (1977) reported a similar observation: "Research that has explicitly compared different systems of the same type is rare" (p. 965). Williams comments on two studies, one by Christie and Holloway (1975), and one of his own earlier studies, which compared four types of audio conference systems (Williams (1974) in Williams, 1977). The four audio tools were microphone-speaker system, voice-switched telephone with loudspeakers, manually switched telephone with loud-speaker and boom microphone-headphone system. Williams' earlier study produced the now familiar results; there were differences "in attitudes to systems, and in speech patterns, though not in task outcome" (Williams 1977, p. 965). He concluded that there were a more differences between face-to-face and audio-only contexts than occurred between the four audio-only systems.

More recently, research has drawn comparisons between different VMC systems; for example, O'Conaill et al. (1993); O'Conaill and Whittaker (1997), and Sellen (1992; 1995). Some of these studies have varied the type of audio tool afforded by VMC systems. However, at the same time the quality of the visual signal, speed of transmission and the lag between audio and video signals also vary, so that the impact of the various audio tools is difficult to ascertain. The study by O'Conaill, Whittaker and Wilbur (1993) is a good example of this type of comparative research.

5.5.1 An Evaluation of the Spoken Aspects of VMC

O'Conaill et al. (1993) applied a naturalistic approach to the evaluation of VMC systems. The study evaluated the impact of two VMC systems of communication when they were introduced into the workplace to support real work meetings. One of the VMC systems was run over an Integrated Services Digital Network (ISDN) between United Kingdom and USA. ISDN is a readily available network, but "suffers the limitations of transmission lags, a half-duplex line, and poor quality video" (O'Conaill et al., 1997 p. 389). The second VMC system was run over an analogue network, the London Interactive Video Education Network (LIVE-NET), which connects eight colleges in the London area. The system provides full duplex audio with broadcast quality video, it has negligible transmission delay as it uses "optical transmission and video-switching technology" (O'Conaill et al., 1993, p. 390). In essence the two VMC systems provide audio and visual signals of very different quality. The ISDN system provided half-duplex audio, with voice-activated switching (O'Conaill, private communication, 1996) and low quality video. LIVE-NET afforded full-duplex (open channel) audio and high quality video.

The impact that these dissimilar VMC contexts had upon real group discussions was evaluated by comparing the speech characteristics that occurred in five ISDN, four LIVE-NET and five face-to-face group discussions. On average six people participated in each discussion. A simplified form of Conversational Analysis was applied to the middle sections of the transcribed group discussions, taking a twenty minute section of the conversations which occurred in each of the three communicative contexts. The speech characteristics analysed in this study included the frequency and length of turns, frequency of interruptions, rate of overlapping speech, and use of explicit handovers. It should be noted that O'Conaill et al. (1993; 1997) make a distinction between interruptions and episodes of overlapping speech. Both are examples of simultaneous speech, but interruptions occur when the first speaker gives no indication that they are about to relinquish the floor, whereas overlapping speech occurs after speakers signal that they may relinquish the floor.

O'Conaill et al. (1993; 1997) report significant differences in speech characteristics between the two VMC contexts, and between the VMC and face-to-face interactions. Group discussions in the ISDN context contained significantly fewer, but longer turns than either the LIVE-NET or face-to-face interactions. O'Conaill et al. suggest that turns in ISDN mediated meetings appeared more 'lecture-like' in character, as they contained very few interruptions and backchannels. Contributions in LIVE-NET meetings were shorter than in ISDN conversations, and contained significantly more interruptions and backchannels. However, conversations in both VMC contexts contained fewer interruptions than face-to-face interactions; in face-to-face discussions "almost 10% of turns were interruptions compared with less than 2% in ISDN" (O'Conaill et al., 1997, p. 120).

The overall rate of overlapping speech did not differ significantly with context, but further analysis of the three subtypes of overlaps examined in this study showed that overlaps occurred for different reasons in the three contexts. The subtypes of overlaps were defined by O'Conaill et al (1993) as follows: 'projections', which occur when the next speaker anticipates that the current speaker is about to complete their turn; 'floorholding', where overlapping speech occurs when the next speaker tries to take the floor whilst the current speaker attempts to hold the floor; 'simultaneous starts', which arise when two participants attempt to begin a new turn concurrently. Overlaps occurred in "face to face mainly because of projections and floor holding, in LIVE-NET because of projections, and in ISDN because of simultaneous starts" (O'Conaill et al., 1993, p. 413). The higher frequency of simultaneous starts in ISDN meetings were probably due to the transmission lags in the ISDN system (O'Conaill et al., 1993).

One way of overcoming the problems of turn-taking in ISDN is to signal who should speak next. O'Conaill et al analysed the use of explicit handovers in each communicative context. Explicit handovers included the use of questions, tagging, or naming the next speaker. This analysis showed that explicit handovers occurred more often in the ISDN context than in face-to-face meetings. However, explicit handovers were also used more frequently in LIVE-NET than in face-to-face discussions; so even when transmission lags were absent, mediated communication led to greater formality in turn-taking procedures.

Overall, the findings by O'Conaill et al. (1993; 1997) show that the speech patterns were more disrupted in the ISDN than LIVE-NET meetings. However, it is difficult to know whether this effect is due to the lag in transmission of signals, to the half-duplex audio channel, to the impoverished video images, or indeed to the joint effect of all of these features of the ISDN system.

As O'Conaill et al. (1993) point out, further laboratory based research is required to tease apart the relative impact of low quality video and audio signals upon VMC collaborations. This is one of the aims of Study 2, which examines the effects of two types of audio channels (full duplex and half-duplex) by varying the audio tools but not the quality of the visual signals.

5.6 Study 2. Impact of VMC on Collaborative Communication: the Audio Channel

5.6.1 The Aims of Study 2

In Study 1 it was noted that text-based CMC affected both the process of communication and the process of collaboration. The Study also demonstrated that novice users quickly adapted to CMC. Study 2 examines the communicative constraints of two VMC systems, which vary only in their audio channel. One VMC system affords a full-duplex 'open channel' (OC) audio link. The second VMC system provides a 'half-duplex' audio link, in which participants use the mouse to activate the audio tool when they wish to speak; this is a 'Click-to-Speak' (CTS) audio system. The aim of the study is to determine if these different audio tools cause speakers to alter their communication. The way that people communicate in these two VMC contexts will be compared to face-to-face interactions.

The restraints imposed by a click to speak (CTS) audio tool could have impact on the process of communication; especially management of smooth interactions, as participants using a CTS tool will need to take it in turns to use the audio channel. Users of the open channel VMC system should be able to talk to each other easily, in a similar manner to face-to-face interactions. Study 2 will explore how these differences effect processes of communication and collaboration.

The Study uses an experimental approach, varying the type of audio link whilst keeping other aspects of the environment and the task as constant as possible. One of the advantages with using this approach is that the effects of the different VMC systems on task performance can be ascertained. O'Conaill et al. (1993) comment that it was difficult to find an objective measure of the successfulness of communication in their study of real-life meetings, so this aspect of the impact of VMC could not be judged. The task being used in this study is The Map Task, which provides an objective measure of task performance.

The research questions being addressed in this study are, therefore, what effects do full-duplex and half-duplex audio channels have upon the processes of communication and collaboration in VMC? These questions can be answered by comparing measures of task performance and the process of communication in the two VMC systems, and contrasting these with measures taken from face-to-face interactions. Subjective measures of user satisfaction in the VMC conditions are also included.

5.6.2 Method: Comparison of Two VMC Systems

In this study two groups of subjects completed the same collaborative task (The Map Task) using VMC. One group of participants using an open-channel audio link, the other group using a half-duplex audio conferencing tool. The data from this study is then compared to data taken from the face-to-face condition of the HCRC Map Task corpus.

5.6.3 Design

A between subjects design was used, with type of audio tool as the between groups factor. The two types of audio tool were 'open channel' (OC) and 'click to speak' (CTS).

5.6.4 Subjects

40 students (20 pairs) from Glasgow University volunteered to take part in this study. The participants were aged between 17 and 31 years, mean age of 20.27 years. A small financial prize was offered to the pair in each VMC condition who completed the Map Task most accurately.

Familiarity

All the participants were familiar with their partner, having volunteered to take part in this study with a friend. Partners had known each other for at least 2 months (mean length of familiarity was 3.16 years, range 2 months - 10 years).

5.6.5 Task and Materials

The task used in this study was the Map Task. The maps were presented in the same format as in Study 1. They are reproduced on A3 paper (29.7 cm by 42.0 cm), and the landmarks on the maps are depicted by simple line drawings. The maps were fastened onto hardboard and fixed over the right-hand side of the monitor screen for right-handed subjects, and over left-hand side of monitor screen for left-handed subjects.

Materials

3 pairs of maps from HCRC Map Task corpus, with three different pathways. Maps 1 and 2 were used alternately in each audio tool condition; map 3 was kept in reserve, to be used if participants had to restart a trial. Examples of these maps are shown in appendix (C).

5.6.6 Apparatus

2 x SunSPARC 20 stations operating over a dedicated local area network, running the operating system Solaris and OpenWindows 3.1 Graphical User Interface; this environment enables the running of video (VIC) and audio (VAT) conferencing tools, both of which are publicly available. VIC provides full colour jpeg encoded video at 5 -6 frames per second, and was used in both VMC conditions. The visual input was provided via 2 JVC Videomovie GR.AX60 compact VHS recorders, which were centrally placed over the monitor. A very small lag between the visual and audio signals occurred in both VMC systems. The audio tools used are described below.

Open channel audio.

An audio link was run between the 2 rooms in which the SunSPARC stations were situated, providing full duplex sound. Participants communicated via SHURE SM2 dual receiver head and boom mounted microphones. This provided the subjects with full duplex (open channel) audio. The audio output was combined (using a MACKIE micro Series 1202 12-Channel Mic/line Mixer), and an analogue recording made via a JVC KD A33 stereo Cassette recorder.

Click to speak audio

In this VMC set-up the audio tool was VAT 3.4 Van Jacobson (1994), which was set for a 'click to speak' method of channel activation. The audio was captured on SunSPARC microphones, and relayed to each participant on the computer’s loud speakers. The spoken dialogues were recorded using a Sony Professional Walkman Stereo Cassette-recorder WM-D6C.

Monitor Configuration.

For a right-handed subject the A3 paper version of the map was located over the right-hand side of the monitor screen, and the video window and audio tool (when in use) was position on the left hand side. This was reversed for left-handed participants. A line illustration of the monitor configuration is given in Appendix D.

5.6.7 Procedure

The two types of audio-tools were used on alternative days, with the 'OC' being in use on the first day of the experiment, and 'click to open' audio tool being used on the second day, etc. Participants were divided into two groups, depending on which days they were available to take part in the study. As the participants did not know in advance that different VMC environments were being used on different days, the allocation of subjects to the conditions was quasi-random. On arrival the pairs of subjects were randomly allocated the role of Instruction Giver or Instruction Follower, as in Study 1. Participants were then taken to the two separate rooms in which the computers were set-up. They were given written instructions on the Map Task; these instructions are identical to those used in Study 1 and are therefore not repeated here. Written instructions concerning the method of communication were also available, as follows:

Instructions for communicating: Open channel.

"Put on the headphones, and adjust the microphone so that it is in front of your mouth. You will now be able to hear and talk to your partner, and you will be able to see your partner on the computer monitor.

Once you have both read the instructions, and are happy that you can communicate to each other, let the experimenter know that you are ready to begin."

Instructions for communicating: click to speak

"In front of you on the computer monitor, you will see a video picture of your partner. Under the video picture of your partner there is an audio tool box.

To talk to your partner, move the 'mouse' so that the pointer is in the large grey square of the audio tool box. Hold down the right-hand button on the 'mouse' whilst you talk to your partner. You need to keep the right hand button down all the time you are talking, but release the button to hear what you partner is saying. Please have a short talk with your partner, and make sure that you both understand the instructions before you tell the experimenter that you are ready to start the Map Task."

On completion of the task the subjects were asked to fill in a short questionnaire, which elicited their views on the task and the communication environment they had just used. An example of the questionnaire is given in Appendix E.

Once the Map Task had been completed participants were thanked for their assistance in the study, and the over-all aims of the study were explained to them. Permission to use the audio and video recordings for transcription and coding reasons was sought at a later date, when the winners of the study were announced to each participant. No objections were raised to the use of the recordings for the stated purpose.

5.6.8 Transcriptions of Dialogues

Full orthographic transcriptions of the 10 dialogues in each VMC group were made, using the audio recordings taken during the study.

5.7 Comparison Data taken from HCRC Map Task Database

A sample of ten dialogues were taken from the HCRC Map Task data base. These dialogues represent the first trial of the Map Task with a familiar partner, in the face-to-face condition of the Corpus. Since the sample is comprised of the dialogues where participants attempt the task for the first time, this samples forms an appropriate comparison group for the VMC dialogues.

The conditions under which the dialogues from the HCRC corpus were collected were described in Chapter 3, but further details for the sample used in this Study are described below.

5.7.1 Subjects

Ten pairs of subjects from the University of Glasgow student population. This sample consisted of eleven male and 9 female students who were aged between 17 - 30 years (mean age 19.5 years)

Familiarity

All participants were familiar with their partner, having known each other for about 2 years.

5.7.2 Map Task and Materials

In the face-to-face version of the HCRC Map Task, pairs of subjects are seated one on each side of a double-sided easel, on which the maps are displayed. The easel prevents participants from seeing each other's maps whilst allowing a view of the face and upper half of their partner's body.

5.7.3 Apparatus

The dialogues were recorded on a DAT (Sony DTC100EC) using Shure SNIOA microphones. Separate DAT channels were allocated to each speaker. Split-screen video recordings were also taken of the proceedings..

5.7.4 Procedure and Transcription of Face-to-face Dialogues

The procedure for the collection and transcription of the face-to-face Map Task dialogues was the same as for the 'audio-only' Map Tasks, as described in the Study 1.

5.8 Results

The results begin by comparing task performance in the VMC conditions performance in the face-to-face context. This is followed by analysis of various dialogue measures, such as length of dialogues and turn-taking management, to see if the different audio tools affected collaboration in VMC. Data from the questionnaires will also be considered, as this gives an indication of participants' subjective ratings of the two VMC systems.

Note: Use of parametric and non-parametric statistics.

Each set of measures in the results section were first checked for homogeneity of Variance. This initial analysis showed that parametric statistics could be applied to two of the measures; the route accuracy scores and the proportion of turns that end with a question. Non-parametric statistics were applied to the remaining data.

5.8.1 Task Performance

The routes drawn in the two VMC environments were scored for accuracy, using the method described in Study 1. The route accuracy scores were then put through a square root transformation, to give a standardised route accuracy score for each map in centimetres. The mean raw and standardised route accuracy scores for the VMC and face-to-face Map Tasks are presented in Table 5.1. The standard deviations are given in brackets.

Table 5.1 Mean Route Accuracy Scores for VMC and Face-to-face Map Tasks

	Face-to-face	**Open channel**	**Click to speak**
Raw data	79.3 (55.35)	142.7 (62.41)	106.85 (45.37)
Standardised data	8.37 (3.19)	11.69 (2.6)	10.16 (2.14)

The data presented above show that the route accuracy scores for the VMC context appear to be larger than the scores obtained in the face-to-face context; indicating that the users of the VMC systems may have drawn their routes less accurately than participants in the face-to-face context. An Analysis of Variance on the standardised route accuracy scores was computed, to determine if any of these differences were significant. A One-way ANOVA, with communicative context as a between group factor (3 levels: open channel VMC, click to speak VMC and face-to-face) produced a significant main effect of communicative context [$F(2,27) = 3.84$ $p<0.05$]. Post hoc analysis (by Tukey HSD) showed that the only significant difference occurred between the route accuracy scores for the open channel VMC and Face-to-face context ($p<0.05$). Participants in the open channel VMC context performed the Map Task less accurately than participants who communicated face-to-face; the routes drawn by open channel VMC users were approximately 28% less accurate than routes drawn in the face-to-face context. All other comparisons were non-significant ($p>0.10$).

In summary, users of the open channel VMC system produced routes on their maps that deviated more from the intended route than the routes drawn in the face-to-face context. The route accuracy scores for participants in the click to speak context were also larger than those of face-to-face participants; but, this difference was non-significant. Why did users of the open channel VMC system complete the task less accurately than participants who interacted face-to-face? In order to answer this question the process of communication in each of the contexts will be explored in the following section.

5.8.2 Measures of the Process of Communication

To determine if communicative context effected the way in which participants interacted a set of measures were taken. The measures of process of communication include the length of the dialogues (number of words and turns), average length of turns and time taken to complete the task. The data for the dialogue measures are presented in Table 5.2 below, the standard deviations are given in brackets.

Table 5.2 Measures of the Process of Communication. Mean number of words, turns, words per turn, and time per dialogue.

	face-to-face	Open channel	Click to speak
words	1543.20 (386.33)	1908.40 (683.94)	2910.50 (1476.78)
turns	202.60 (63.60)	235.30 (86.19)	218.70 (114.24)
turn length	7.81 (1.14)	8.11 (3.55)	14.14 (3.45)
time (minutes)	8.14 (2.46)	9.94 (4.13)	22.23 (11.50)

The Means for the various dialogue measures suggest that the different audio tools may have had some impact on the process of communication. For example, dialogues in the CTS condition appear to contain more words, and take longer to complete than in OC context or face-to-face communication. The measures of the process of communication (number of words and turns, turn length and time to completion) were analysed using non-parametric statistics (Kruskal Wallis Analysis of Variance by Ranks), with communicative context as a between group factor (3 levels as above) to determine which comparisons were significantly different. The results are displayed in the table overleaf.

Table 5.3. Table of results for Kruskal Wallis ANOVA on dialogue measures

	H Statistic	df	probability
words	10.129	2	p<0.01
turns	0.895	2	non significant
turn length	16.385	2	p<0.01
time	16.301	2	p<0.01

The critical value of H (for 3 groups, with 10 observations in each group) for probability of $p<0.05$ is 5.99, and for $p<0.01$ is 9.21. As the table above shows, these levels are exceeded for three of the dialogue measures. The results revealed a main effect of communicative context for measures of number of words, length of turns and the time taken to complete the tasks varied significantly between at least two of the communication contexts. Mann Whitney U (two-tailed) tests were applied to determine which of the group means differed significantly. The table below shows which comparisons between the means reached significance.

Table 5.4. Results of post hoc tests. Mann Whitney U

	CTS vs. OC	CTS vs. face	OC vs. face
words	non significant	p<0.01	non significant
turn length	p<0.01	p<0.01	non significant
time	p<0.01	p<0.01	non significant

These tests confirmed that two of the dialogue measures varied significantly between the two VMC contexts. The click to speak VMC dialogues had longer turns than open channel dialogues [$U(df\ 10,10) = 8, p<0.01$]. On average, turns in the CTS contained approximately 42% more words than turns in the OC context. The click to speak dialogues also took more time to complete than open channel VMC dialogues [$U(df\ 10,10) = 8, p<0.01$]; participants in the CTS context took more than twice as long to complete the tasks than participants using the open channel VMC system; the mean times were 22.23 minutes versus 9.94 minutes

Comparisons between the dialogue measures for CTS and face-to-face contexts gave significant results on all three dialogue measures. Dialogues in the CTS context contained more words than face-to-face interactions [$U(df\ 10,10) = 11\ p<0.01$]. Turns in the CTS context were longer than turns in the face-to-face dialogues [$U\ (df\ 10,10) = 0, p<0.01$]. Finally, CTS interactions took longer to complete than face-to-face interactions took [$U\ (df\ 10,10) = 2, p<0.01$]. To summarise, dialogues in the click to speak VMC context took nearly twice as many words, turns were almost twice as long, and required more than twice the amount of time to complete as face-to-face interactions.

In contrast, all of the comparisons between open channel and face-to-face dialogues were non-significant. The open channel and face-to-face dialogues were very similar in terms of length and structure of turns. Participants using an open channel VMC system produce dialogues which were structurally very similar to face-to-face interactions.

The analysis of the process of communication in the three contexts has shown that users of the click to speak VMC system produced dialogues which were structured quite differently from both face-to-face and open channel VMC. The dialogues were longer, both in terms of the amount of linguistic output and time taken to complete the task. However, the number of turns in a dialogues was not effected by context, but users of the CTS system used longer turns (more words per turn) which accounts for the increase in overall dialogue length. This may have been a result of using a half-duplex audio link, which can only transmit one voice at a time. Users of a CTS audio link cannot talk simultaneously, as the one-way audio channel prevents this occurring. If people attempt to interrupt each other the audio signals cannot be transmitted, and all that is heard is a nasty buzz (or 'system noise'). Managing smooth transition of turn-taking becomes more problematic in a CTS context. One way of overcoming this problem has been commented upon in previous research; speakers use longer turns than normally occur in open channel communication. O'Conaill et al. (1993; 1997) report that this pattern of communication occurred in ISDN video-conferences, and similar patterns of speech have been found in telephone conversations (Rutter, 1987).

The next step in the analysis of the impact of the different audio tools in VMC is to examine how turn-taking was achieved in these communicative contexts.

5.8.3 Management of Process of Communication

Several methods of assessing the process of communication can be chosen; in this study the process is analysed in terms of:

1) The number of interruptions, and episodes of overlapping speech that occur in eachdialogue.
2) The use of explicit turn-taking procedures.

Analysis of Interruptions and episodes of overlapping speech

Interruptions and areas of overlapping speech are included in the transcriptions of the dialogues, the transcriber marking the dialogues to indicate where the interruption occurred as precisely as possible on the basis of the audio recordings. In this study interruptions and episodes of overlapping speech in the face-to-face and open channel contexts are operationally defined using the following definitions (which were described earlier in Chapter 3):

Interruptions occur when one person starts to speak whilst another is already talking. Overlaps occur if one or more words of the second speaker's contribution are perceived to overlap the first speaker's contribution.

Attempts to interrupt, or talk simultaneously, were impossible in the click to speak context, due to the nature of the half-duplex audio tool. Any attempts to talk simultaneously were noticable, due to the noise made by the system, and the whereabouts of these attempts to talk simultaneously were also noted after the dialogues had been transcribed.

The frequency of interruptions and episodes of overlapping speech were obtained for each of the dialogues in the Open channel, click to speak and face-to-face contexts. The process of turn-taking in this study is based upon the number of episodes of overlapping speech, rather than upon the frequency of interruptions. The former is a more conservative way of estimating how much simultaneous speech occurs, but it may provide a more accurate idea of frequency of overlapping speech, as it is very difficult to decipher who is interrupting whom when a sequence of overlaps occurs.

Examples are shown below to clarify the distinction been made between interruptions and episodes of overlapping speech.

Example 1 An Interruption

<IG: come up so you're coming up/

IF :on the left hand side?

IG: yeah back up towards the caravan park>

In this example, the Instruction follower interrupts the Instruction Giver; the forward slash marks the point where the interruption begins, and the two turns are bracketed together (< >) to show that some overlapping speech occurred. On some occasions there are a series of interruptions and areas of overlapping speech, as in Example 2

Example 2 An episode of overlapping speech

<IG: You want to be going horizontal towards/

IF: You mean towards

IG: towards West Lake but/

IF: West Lake?>

IG: Yes just for 2 centimetres.

Example 2 has two points of interruption, but would be counted as one area of overlapping speech, as the overlaps run into each other.

In this study a standardised rate of overlapping speech was also calculated; the standardised rate takes into account the significant differences in length of dialogues in the VMC and face-to-face contexts. The standardised rate of overlapping speech used here is, the number of episodes of overlapping speech that occur in every 100 turns of dialogue.

The transcripts dialogues were searched for episodes of overlapping speech, and the frequency per dialogue was calculated. The raw number of incidences of overlapping speech, standardised rate of interruption are shown in table 5.5 (with standard deviations in brackets).

Table 5.5. Mean Number of Overlapping Turns of Speech.

	face-to-face	Open channel	Click to speak
overlaps (raw)	17.00 (7.50)	37.80 (13.37)	10.00 (6.75)
overlaps (std)	8.25 (1.14)	17.09 (7.10)	4.26 (3.45)

The Means in table 5.4 suggest that a greater number of stretches of overlapping speech occurred in the open audio channel context than in either the CTS or face-to-face interactions. Analysis of the data was computed using non-parametric statistics, because of heterogeneity of variance in the data.

A Kruskal Wallis One way ANOVA was applied to the frequency of overlapping speech, first to the raw scores and then to the standardised scores. In both analyses, the between group factor was Communicative context (three levels; OC, CTS and Face-to-face). In the first analysis, the main effect of context was confirmed [H (df 2) = 17.31, $p<0.01$]. Mann Whitney U comparisons revealed that more overlaps occurred in the open audio channel than either the CTS or face-to-face contexts ($p<0.01$); more than twice as many episodes of overlapping speech occurred in open channel VMC context. The number of overlaps in the CTS and face-to-face context did not vary significantly ($p>0.10$).

Analysis of the standardised data, which indicates the rate of overlapping speech per 100 turns of dialogue, produced a similar pattern of results. However, in this analysis a significant difference occurred between all three group means; the highest rate of overlapping speech occurred in OC context, less overlaps occurred in face-to-face interactions, and the lowest rate of overlap occurred when the CTS audio tool was used. Again, analysis was by Kruskal Wallis one way ANOVA, with communicative context as the between group factor (with three levels) and the standardised rate of overlapping speech as the dependent variable. The main effect of context was confirmed [H (df 2) = 21.231, $p<0.01$]. Mann Whitney U tests were applied to the pair-wise comparison of the three means, and confirmed the order of the effect at the $p<0.01$ level. Open channel VMC contained a higher rate of overlapping speech than face-to-face dialogues, which had a significantly higher rate of overlaps than dialogues in the click to speak VMC context. The percentage of turns with episodes of overlapping speech were 17% in open channel, 8% in face-to-face and 4% in click to speak contexts.

In summary, analyses of the frequency and rate of overlapping speech show that the switching between speakers was handled more smoothly in the click to speak VMC context. In contrast, this process appears to have been disrupted in the open channel VMC system, where episodes of overlapping speech occurred twice as frequently as in the face-to-face dialogues. The major difference between the dialogues in the Open channel and click to speak VMC contexts seems to be that they varied in how they were *structured*.

The way in which this difference affected the process of communication could be explored in several ways. For instance, Conversational Games Analysis could be applied. This form of analysis would, however, illuminate differences in the function and content of utterances, rather than how the utterances were structured into a dialogue. Conversational Games Analysis is also extremely time consuming; since it normally takes approximately 3 hours to code a 10 minute spoken dialogue, the longer CTS dialogues could take at least 6 hours to code. For these reasons, Conversational Games Analysis will be excluded from Study 2. Instead, this study will follow the example set by O'Conaill et al. (1993; 1996) and Sellen (1992; 1994), who explored the impact of VMC by examining how turns were allocated in these different contexts.

Explicit turn-taking procedures

O'Conaill et al (1993; 1997) reported that turn taking in VMC was more formal than in face-to-face interactions. This was observed in both the ISDN and LIVE-NET video-conferences. Formality was measured by looking at the occurrence of explicit handovers; such as naming of next intended speaker, use of 'tag questions' (that is, adding redundant information to the end of a turn, such as 'you know') and the use of question-ending turns to indicate that the previous speaker had finished their turn.

In Study 2, participants very rarely used names to assist turn-taking, this might have been because the interactions took place between pairs of subjects whereas O'Conaill et al.(1993; 1997) examined group discussions where competition for the floor was greater. A few incidences of using names to address the next speaker did occur in the OC and CTS contexts, usually when a speaker asked their partner if they could be heard. So this form of explicit turn-taking was not explored in this study. However, the dialogues appeared to have a large number of turns that ended with a question, so the frequency of this phenomena was calculated from the transcribed dialogues. A few examples of question-ending turns are shown below.

Examples:

1) I am now just below the left hand corner of East Lake, Yeah?

2) Okay, the start is right next to the Caravan Park, okay?

3) What happens with the route at this point, l presume the path goes round the Forest, does it?

4) Have you got the Pelicans marked on your map?

The frequency of question-ending turns was standardised (as the number of question-ending turns per 100 turns of dialogue), so that variations in length of dialogues in the different communicative contexts would not have a confounding effect. The group means are given in table 5.6, with standard deviations given in brackets.

Table 5.6 Means for Question-ending turns.

Face-to-face	Open channel	Click to speak
22.85 (7.73)	26.57 (10.69)	34.36 (6.18)

The data in table 5.6 suggest that participants in the CTS context made greater use of question-ending turns than speakers in OC or face-to-face interactions. To see if the group means differed significantly, a One way ANOVA was applied to the standardised scores, with communication context as a between group factor, with the three levels as before. The main effect of communicative context was significant [F (2,27) = 4.87, p<0.01]. Post hoc tests revealed that the only significant difference was between the group means for CTS and face-to-face contexts (Scheffe F = 4.68, p <0.05), the means being 34.36 and 22.85 respectively. Therefore, speakers in the CTS context completed their turns with a question more frequently than speakers in either the OC or face-to-face dialogues. Over one third of all the turns in the CTS dialogues were completed with a question, compared to 27% and 23% of turns in the OC and face-to-face interactions.

The use of question-ending turns in the click to speak VMC context would be a successful way of indicating that a speaker had completed their turn, which in the CTS interactions were unusually long. This communicative strategy would assist in the smooth transition of speaker turns, clearly signalling when a speaker was about to release control of the half-duplex audio channel. This could account for the surprisingly small number of episodes of overlapping speech that occurred in the click to speak VMC interactions.

The results of analysis of process of communication in the VMC and face-to-face contexts have highlighted some interesting points. The different audio tools used by the VMC systems appear to have effected how people interact and collaborate in these contexts. Users of an open channel VMC system structured their dialogues in a similar manner to face-to-face interactions; they said approximately the same number of words, whilst using turns of similar lengths. However, the rate of overlapping speech in open channel conversations was considerably higher than in face-to-face dialogues. Participants in the click to speak context said much more, and hence took longer to complete the task. They also adopted different strategies for dealing with the problem of turn-taking in a half-duplex communicative context; using longer turns to reduce the number of handovers, and more question-ending turns to indicate that the next speaker could take the floor.

Before discussing the possible implications of this analysis of the process of communication, and how it might account for the observed differences in task performance, the subjects' opinions and ratings of the VMC contexts will reported in the following section.

5.8.4 Subjective Data from Post Trial Questionnaires

Participants in Study 2 were asked to complete a short questionnaire, which elicited their opinions on four aspects of communication in the VMC contexts. The four questions were as follows:

Q1) How easy was it to communicate with your partner?
Q2) How easy was it to take turns at speaking?
Q3) How often do you think you interrupted your partner?
Q4) How often do you think you looked at the video of your partner?

Participants were asked to indicate their answers on a five point Likert-type scale. For example, the scales for questions 1 and 2 were: very easy, fairly easy, neither easy or difficult, fairly difficult, and very difficult. A copy of this questionnaire is shown in Appendix (E), and summary tables of the distributions of responses to this questionnaire are presented in Appendix F.

To establish if the distribution of responses varied with communicative context Analysis of Variance was carried out on the data for each question, with type of audio tool (CTS or OC) and role of participant (Instruction Giver or Instruction Follower) as between group factors. The use of parametric analysis can be supported by the fact that the ANOVA is said to be a "very robust statistical procedure, and the assumptions (normality and homogeneity of variance) can be violated with relatively minor effects" (Howell, 1997, p. 321). This is especially true if the sample sizes are of equal number.

In this analysis of the distribution of the questionnaire answers, each of the 5 possible answers was allocated a number; so that responding "very easy" was given a rank of One, "easy" was given a rank of two, etceteras. The results of these analyses produced only one significant result, subjective ratings of turn-taking were affected by communicative context and the role of participants. Distribution of the ratings for Question 1 ("ease of communication"), Question 3 ('frequency of interrupting partner) and Question 4 (frequency of gazing at partner) resulted in non significant main effects of communicative context and role of the participants, as well as non-significant interaction effects ($p>0.10$).

Only responses to the question concerning how easy it was to take turns whilst using the two VMC systems showed any effect of communicative context. The ratings from CTS users showed that they considered turn-taking to be more difficult (given a higher rating) than participants who used an open channel audio link; [$F(1,36) = 10.075$, $p<0.01$]. The mean ratings were 2.1 and 1.35 for the CTS and OC contexts respectively. The interaction between audio tool and role of participant approached significance [$F(1,36) = 3.63$, $p = 0.065$]. Further analysis of the means involved in this interaction (by Simple Effects analysis) showed that the Instruction Followers rated turn-taking differently in the two VMC contexts; [$F(1,36) = 12.89$, $p<0.01$]. Instruction Followers who used the open channel audio tool thought that turn-taking was easier to accomplish than Instruction Followers in the click to speak context. All other differences between the means in the interaction were non-significant ($p>0.01$). The main effect of participant's role was also non-significant ($p>0.1$).

It is interesting to note that the subjective ratings of the two VMC systems appear to be at odds with some of the objective measures of the process of communication. In particular, the subjective ratings of how frequently participants interrupted their partner appear to differ from the objective measures of episodes of overlapping speech in the VMC contexts. The majority of participants (78 % of the subjects) thought that they had only interrupted their partners on a moderate number of occasions (defined on the questionnaire as "approximately every other time you spoke"). However, the objective data shows that participants in the open channel context interrupted each other very frequently, whilst interruptions by users of the CTS system occurred very rarely.

One possible reason for the differences between the objective and subjective measures is that interruptions and episodes of overlapping have different impacts in the two VMC contexts. In the open channel context overlapping areas of speech did not have a disruptive impact on the flow of communication; participants in the open channel VMC could talk over the top of each other for short periods of time without this really being noticed. This happens in everyday conversations fairly regularly, people are not always aware of the overlaps, and therefore under-estimate how frequently they occur. However, in the click to speak interactions the effects of interrupting a partner were more disruptive. The CTS audio only allows the transmission of one voice at a time; if participants try to talk at the same time the audio outputs mask or block each other; the result is that it is very difficult to hear what either person is saying. Therefore, episodes of overlapping speech are disruptive, and very obvious. This could lead to people over-estimate how frequently the interrupt each other. It is possible, however, that participants are simply unaware how frequently they interrupt each other, which could account for discrepancies between objective and subjective data.

5.8.5 Summary of the Questionnaire data

At a general level, the questionnaire data has shown that participants reported that they found it easy to communicate with each other in both VMC systems. None of the participants thought that the VMC contexts made it difficult to communicate. Analysis of the responses to each of the questions showed only one significant difference in distribution of responses; participants felt that turn-taking was slightly more difficult in the half-duplex VMC context. This was particularly noticeable in the responses from Instruction Followers. These participants tended to say less than the Instruction Givers, and may have experienced some problems in timing their turns to provide feedback to their partner. The responses to the other questions did not show any significant effects of communicative context.

Overall, the subjective data has confirmed the general feeling obtained informally during de-briefing sessions; participants had enjoyed using this new form of technology and had not found it difficult to communicate with each other in this novel context.

5.9 General Discussion of the results

This study has investigated the effects of two types of audio tool used in VMC; the provision of half-duplex (CTS) and full-duplex (OC) audio channels. These VMC contexts were compared with face-to-face interactions. The study explored the effects these different audio links have upon task performance and the process of communication, whilst holding the quality of the visual signals constant in the two VMC contexts.

The results show that one of the VMC contexts (open channel VMC) had an impact upon task performance; tasks were completed less accurately in the open channel VMC context than in the face-to-face interactions. The task performance for users of the click to speak VMC system did not differ significantly from the performance of participants in the face-to-face context.

Why did users of the open channel VMC system perform less well in the Map Task? The results from analysis of the process of communication suggest some possible explanations. The structure of the dialogues in the OC context were similar to face-to-face interactions in many ways. For instance, the length of the dialogues in the OC and face-to-face were very similar, and turn length (words per turn) did not vary with context. However, the open channel dialogues contained many more episodes of overlapping speech than face-to-face interactions. In contrast, participants interacted quite differently in the CTS context; they said more, used longer turns and interrupted each other very infrequently. Since the quality of the visual signals was identical in both of the VMC contexts, it is unlikely that variations in task performance and the process of communication can be attributed to the video images. It seems more probable that these effects were due to the different types of audio links used in the VMC systems.

One of the outstanding differences between the way people interacted in the OC VMC context was that they interrupted each other very frequently, more frequently than participants in either face-to-face or CTS VMC contexts.; 17% of turns contained episodes of overlapping speech in the open channel context, compared to 8% and 4% of turns in the face-to-face and CTS dialogues. How do these findings compare with the rate of overlapping speech or interruptions reported by other researchers in this area? Studies which provide data on the rate of overlapping speech include O'Conaill et al. (1993;1997) and Anderson et al. (1996).

Comparing rates of overlapping speech with data from other studies is not a straight-forward process, because researchers tend to measure and define overlaps and interruptions in various ways. As noted previously, O'Conaill et al. analysed interruptions and overlapping speech separately, examining how frequently turns were interruptions or instances of overlapping speech. The overall rate of overlapping speech did not differ between the two VMC contexts, nor between the VMC and face-to-face contexts. O'Conaill et al. did not observe higher rates of interruption or overlapping speech in the LIVE-NET context, which provided an open channel audio link with high quality visual signals. They did, however, find that users of an ISDN system (which provided a half-duplex, voice activated audio link and poor quality visual signals) interrupted each other very rarely; only 2 % of turns were interruptions in this context. This latter effect is similar to the results obtained in Study 2, where the percent of turns that contained simultaneous speech was very low in the CTS context.

The research by Anderson et al. (1996) provides a more straight-forward comparison to some of the data collected in Study 2. Anderson et al. analysed the rate of interruption (a measure which included all forms of simultaneous speech) which occurred in a variety of communicative contexts, including a VMC context very similar to the open channel system employed in Study 2. Anderson et al. found that 10.9 % of turns were interrupted in this VMC context compared to 13.8% of turns in face-to-face interactions. This difference was, however, non-significant. In a latter study Anderson et al. (1997) report on the use of VMC system for remote collaboration over the Internet, using public-domain network video and audio tools. This VMC system provided a half-duplex (CTS) audio link, and low quality video images. The rate of interruption in this VMC context was again found to be very low, though exact rates are not reported.

These comparisons with data from other studies would suggest that the rate of simultaneous speech is likely to be higher when participants use an open channel audio link, and lower in a VMC context that provides half-duplex audio signals. This fits the pattern of results obtained in analysis of overlapping speech in Study 2, but does not explain why the rate of overlapping speech was unusually high in the open channel VMC context used in the present study.

The question that remains to be answered is whether these differences in rate of simultaneous speech, especially the very high rate of overlapping speech observed in the OC VMC system, could account for the observed differences in task performance in this study?

One possible explanation for the decline in task performance by users of the open channel VMC context is that they had gained the impression that they could communicate in this context **as if** they were in a face-to-face setting. The transparency of the audio link could have been the cause of this illusion. However, the quality of the visual signals provided by the VMC systems was moderately poor; the temporal resolution was approximately 5 frames per second. The visual information required to achieve smooth transitions of speakers turns was probably not available in these VMC contexts due to the low frame rate and size of the video images. In the open channel context this resulted in participants interrupting each other very frequently, which in turn disrupted the process of grounding. Since establishing mutual understanding is essential if participants are to complete the Map Task accurately, the effects of assuming that they could communicate in a style similar to face-to-face interactions resulted in poorer task performance.

If users of the open channel VMC system had been more aware of the restrictions of the VMC system then they might have made greater allowances for the communicative restraints they were working under. Participants in the click to speak context were more conscious of the restraints imposed by the technology, as they had to manually activate the audio channel each time they wanted to speak. They appear to have made some allowances for the VMC context, which could explain why they achieved levels of task performance similar to those obtained in the face-to-face interactions.

Participants in the CTS context appear to have made two major adjustments to the way they communicated and collaborated. Firstly, they said more to each other used longer turns. Secondly, they appear to have taken greater care in managing the turn-taking procedures which greatly reduced instances of simultaneous speech. These two ways of adjusting to the CTS context will be discussed in the following sections.

The fact that dialogues were longer in the click to speak context was one of the unexpected findings of Study 2. The dialogues in the CTS context were 35% longer (in words and minutes) than either the OC or face-to-face interactions. Since dialogues in the Map Task tend to be very task-oriented, there are few episodes of off-task chat, the increase in linguistic output could suggest that the CTS dialogues contained more task-oriented information. The extracts below illustrate the length differences of turns in the CTS and OC contexts. In these extracts participants are talking about one part of the route on map 1 (see appendix B); they are drawing the route from the West Lake to the Monument. The following symbols have been used in the extracts: IG refers to the Instruction Giver; IF refers to the Instruction Follower; pointed brackets show areas of overlapping speech, and a forward slant (/) indicates the start of an episode of overlapping speech.

Extract 1. Click to speak.

IG: OK, so right. The trig point is a little bit ehm further down from where you are an a little bit ehm east. And have you got you've got the monument do you?

IF: I've got the monument yeah

IG: OK from where you are just ignore the trig point actually it doesn't matter. Now head down in just a straight line really down to right directly underneath the monument just I mean just underneath it.

Extract 2. Click to speak.

IG: Now do you have a monument on your right had side of your page?

IF: Ehm I do it's just up from mid centre to the right, yeah?

IG: That's right. What I want you to do is from that point, ehm the point that you're at on the shoulder of the west lake draw a line to the bottom of the monument going right, at about 40 degrees to the bottom of the monument but underneath it, as if you were going to go round it again. So straight from the shoulder of the west lake, straight down at a 45 degree angle to underneath the monument. Is that clear?

IF: Yep it's done.

Extract 3. Open channel VMC

IG: So have you got a monument down there?

<IF: I've got a monument/

IG: right good

IF: down there>

IG: So you're going down and round the monument

<IF: Right, round the monument and then/

IG: Uh huh, and then sort of up again>

Extract 4. Open channel VMC

IG: You do a 'U" turn away from west lake

<IF: yeah/

IG: heading back down towards a monument. Have you got a monument?>

IF: yeah

<IG: Right, well it's a straight diagonal line towards the monument. Do you/

IF: okay

IG: get me?>

These extracts show that participants in the CTS context used longer turns which possibly conveyed more information than occurred in the Open channel context; this may have assisted participants in the CTS context to complete the task more successfully. Whether the longer turns did actually contain more information is difficult to ascertain, it may be that the CTS participants were being more cautious and repeating information rather than providing additional information. It is possible to speculate, however, that the provision of longer, more detailed instructions could have reduced the amount of clarification required to establish mutual understanding in this context; just as the more precise instructions required less checking and clarification in the CMC interactions described in Study 1. In addition, the use of longer turns, instead of a greater number of short turns, could be beneficial in this communicative context; it would reduce the number of speaker switches, and thus assist the management of turn-taking.

The extracts of dialogues from the two VMC contexts also demonstrate another distinctive difference between the CTS and OC dialogues; users of the CTS system appear to have taken great care not to interrupt each other. This has already been commented upon, but how did these participants manage to switch between speakers without interrupting each other? Why was it so important to avoid episodes of simultaneous speech?

The quality of the video images were identical in both of the VMC systems, so the reduction in the number of interruptions was not due to variations in the quality of the video signals. It seems more likely that the consequences of interrupting each other in the CTS context where sufficient to induce greater care in turn-taking procedures. As mentioned earlier, if both participants tried to talk at the same time, the audio signals masked each other and it was impossible to make out what either person was saying. In some of these situations the participants had to negotiate who was going to speak next; a process which sometimes required several speaker turns. The effects of interrupting were, therefore, quite disruptive in the CTS context, and this probably explains why these participants took great care to avoid doing so.

Several methods of reducing the amount of overlapping speech have been observed in this study. First of all, the speakers used longer turns which reduced the number of times turn-taking had to be negotiated. Secondly, the CTS participants made greater use of question-ending turns than participants in the OC or face-to-face interactions; over a third of turns in the CTS dialogues ended with a question. This is illustrated in the extract 2 above. Putting a question at the end of a turn is a clear signal that you have finished speaking, and that the other person can now take a turn at talking.

The significantly greater use of longer turns and question-ending turns in the CTS dialogues suggests that these participants made greater allowances for the restraints imposed upon them by a half-duplex audio channel. The CTS users appear to have adapted well to the novel communicative environment. It is possible that they did so because the audio channel had to be manually activated, which made them more aware of the technology they were using. By making greater allowances for the communicative context, the CTS users achieved a high level of task performance, and altered the way in which they communicated to a style appropriate to the communicative context. This suggests that CTS users took greater care in planning what they said to each other, and how they interacted.

5.10 Conclusions

Study 2 has examined the impact of two forms of VMC upon collaborative interactions. The results have shown that task performance was detrimentally effected in open channel context, though this did not occur in the click to speak VMC context. The findings suggest that the poorer task performance achieved by users of the open channel VMC context could have been due to these participants making in-sufficient allowance for the VMC technology. When the open channel dialogues were examined in detail, it was observed that they were very similar in structure to the face-to-face interactions. There was one important difference; the OC dialogues contained a large number of episodes of overlapping speech. These could have occurred because the open channel audio link made it easy for the participants to talk to each other, but the relatively low quality of the video images meant that it was difficult to make use of non verbal communication which assist in the smooth transition of speaker turns. The high incidence of overlapping speech disrupted the process of communication, and may have disrupted the process of establishing mutual understanding which is a crucial factor in many collaborative tasks.

In contrast, participants in the CTS context achieved a level of task performance similar to face-to-face interactions. The findings from analysis of the process of communication suggest that they do so by adapting the way in which they interacted in this VMC context. Although the adaptations to the novel CTS environment inflicted some penalties - in terms of linguistic output, time to complete the task, the need for taking greater care in handling turn-taking and a more formal style of communication - these participants achieved a reasonable level of task performance. Analysis of the process of communication revealed that these participants took great care not to interrupt each other, they may also have altered the way in which information was exchanged; using longer turns to reduce the number of speaker turns. These were sensible adaptations to the restraints imposed by a half-duplex audio channel.

The study has highlighted an important point. If the means of communication are made too transparent (non-intrusive) then participants may not make sufficient allowance for the technology they are using. Sensible adaptations to communicative strategies were made by users of the CTS system, as it clear that the VMC context was different from previously encountered contexts, and the need to adapt was apparent. This did not appear to happen in the open channel VMC context; the audio channel made it easy to contribute verbal interactions, but the quality of the visual channel meant that it was hard to time the onset of these contributions.

If the quality of the visual channel had been higher, then the interactions in the open channel context might have been easier to time. Our claim is that effective timing yields a more effective style of communication and collaboration. This aspect of Video-Mediated Communication will be considered in more detail in the next study.

Chapter 6. Study 3. Impact of Video-Mediated Communication. Effects of quality of the visual channel of communication upon collaboration and communication

6.1 Introduction

The research reported so far has examined the ways in which computer mediated communication restricts the use of a variety of channels of communication. Study 1 looked at the effects of text-based CMC upon collaborative interaction, contrasting this written medium with a co-present audio-only communicative context. Study 2 explored what happened when the spoken channel was restricted in VMC, by comparing full-duplex and half-duplex audio channels. The third study also explores the restrictions imposed by mediated communication, this time focusing on the visual channel of communication. In this final study a comparison is drawn between the quality of the visual channels available in face-to-face and VMC contexts, and how this could effect communication and collaboration. The VMC system used in this study is similar to that reported in Study 2. The quality of the visual signal provided by this system is impoverished, when compared to the quality of the visual signals available in face-to-face interactions. For instance, the size of the video image is relatively small and the frame rate (how frequently the video picture is refreshed) of the image is low. However, the quality of the video image was similar to that provided by many publicly available video-conferencing systems, at the time of this study (May 1994).

The third study differs slightly from the experimental research reported previously in this thesis, as it is more exploratory in nature. The analysis utilises a corpus of dialogues, which was collected with colleagues from the Human Communications Research Centre. The purpose of the HCRC study was to explore the impact of VMC upon a collaborative problem-solving task, the Travel Game task. This task simulates service encounters, and can be thought of as a more 'real-world' version of the Map Task. Participants took part in the Travel Game in a range of communicative contexts; face-to-face, remote audio-only (spoken) interactions, and two different forms of computer-mediated communication (VMC and audio conferencing). Some of the results from the Travel Game experiments have already been published (Anderson et al., 1996; Newlands et al., 1996). The results of the effect of communicative context upon task performance and dialogue measures - the length of dialogues and rate of interruption - will be presented in this chapter to set the scene for Study 3.

The aim of this study is to test the hypothesis that the quality of the visual signals in VMC will affect the processes of communication and collaboration. This hypothesis can be tested by comparing the structure and content of dialogues from VMC and face-to-face environments, using Conversational Games Analysis (Kowtko et al., 1992). This form of discourse analysis was introduced in Chapter 4, where it was applied to the text-based CMC interactions. The analysis illustrated the ways in which effective communication was achieved in CMC, how this differed from spoken interactions, and how novice users adapted to the CMC context.

To date, Conversational Games Analysis has been applied mostly to dialogues from the HCRC Map Task corpus. The coding scheme has been applied to a small number of dialogues from another problem solving task, the Maze Game (Garrod and Anderson, 1987); undertaken by Kowtko et al. (1992) as part of a reliability study of Conversational Games Analysis. It should, therefore, be possible to apply Conversational Games Analysis to dialogues from the Travel Game task. If this can be achieved with relatively few changes to the original coding scheme, then this form of analysis could be a useful tool to include in future assessments of different communication environments.

6.2 Previous Research and Evaluation of the Effect of Video Mediated Communication

A brief over-view of the impact of VMC upon the process of communication was presented in Chapter 2, and more literature was introduced in Chapter 5. These reviews concluded that it is difficult to obtain an overall impression of the impact of VMC upon communication and collaboration, because of the range of VMC systems being evaluated and the breadth of tasks used in the studies. This chapter focuses mainly on one of the channels of communication provided by a VMC system, the visual channel; specifically exploring the effects that the quality of visual signals may have upon the processes of communication and collaboration.

6.3 Factors Effecting the Quality of Visual Signals Provided by VMC Systems

The available literature shows that a range of video technologies have been used to provide the visual signals in VMC, and that the quality of the video image varies with the technology that provides it. Only a few studies have systematically varied the quality of video images provided in VMC, whilst holding other aspects of the conferencing system constant. For example, Barber and Laws (1994) explored a range of factors that have an impact upon the quality of video images. They concluded that the temporal resolution of a video image (the frame rate) was the factor that had the largest impact upon speech based tasks, performance on these tasks deteriorated as the frame rate was reduced from 25 to 8.3 Hz.

Monk and Watts (1995) and Anderson et al. (1996) explored the effect of just one screen parameter, the size of the video image. In these two experiments the size of the on-screen video image was varied between a small and a large image, the effect this had upon the way users interacted was examined. Anderson et al. report that changing the size of the image (from 3.5 by 4.5 inches to 6.5 by 8 inches) had no significant impact upon task performance or the length of the interactions, and concluded that increasing the size of the video window did not increase the sense of social presence (Anderson et al., 1996). Monk and Watts (1995) reached a different conclusion, they found that the size of the video image did change the verbal behaviour of VMC users; conversation was less fluent when the image was reduced in size from approximately 4.5 by 5.5 inches to 1.5 by 2.5 inches. However, the amount of gaze directed at the video image did not vary with the size of the image.

Apart from the handful of studies which have systematically explored the impact of the quality of the video image, there is a growing literature which reports the impact of various VMC systems upon communication and collaboration. Much of this research is based upon comparisons of different VMC systems, or by comparing VMC with face-to-face or audio-only contexts. In these studies the quality of the video image varies greatly, depending upon the technology used to provide the VMC context. This type of research has produced some useful insights into the effects that the quality of the video image has upon communication, but the findings are very mixed and difficult to interpret. Examples of these studies will be reviewed in the following section of this chapter. Before commencing the review, the experiments by Barber and Laws (1994) will be briefly outlined; this will introduce some of the factors that can affect the quality of video image, and set the scene for the following review.

Barber and Laws (1994) used a series of experiments to determine the effects that the quality of video images had upon communication. These experiments systematically varied several factors that could effect the quality of monochrome video images; the factors included frame rate (temporal resolution), the number of grey scales (contrast resolution), image size and spatial resolution (number of pixels). The research focused on some of the perceptual and cognitive tasks that occur during communication, examining how the quality of the image effected face recognition, emotion recognition, message comprehension and the use of visual speech cues (such as lip-reading). Four different communicative contexts were included in the research: audio-only, audio with normal analogue video, and two audio contexts where the digitised displays provided images with frame rates of either 25 or 8.3 frames per second.

Only the findings which relate to variations in temporal resolution will be described here. There are several reasons for focusing on this particular parameter. Firstly, Barber and Laws found that temporal resolution was the parameter that had the greatest impact on the measures they used for evaluation purposes. Secondly, the frame rate of the video image is the factor most often commented upon in research literature; in the following review the frame rate will be given whenever this is possible, as an indication of the quality of the video images provided by the various VMC systems. Thirdly, Barber and Laws were unable to complete a systematic evaluation of the impact that image size and spatial resolution had upon the quality of the video image. The only other factor they examined was contrast resolution, this factor is rarely commented upon in the following studies where colour images were frequently provided. These reasons explain the emphasis on the temporal resolution of video images in the following review.

The temporal resolution of video images provided in VMC was found to have "the most consistent and powerful effect on human performance measures" (Barber and Laws, 1994, p. 174). The measures of performance used in these experiments were based upon the number of words, or numbers, that participants could correctly shadow (that is, repeat aloud) or lip-read when they were spoken in the different communicative contexts. Performance in the shadowing task was found to be poorest in the audio-only context, to be similar when 25 Hz analogue and 25 Hz digital video images were provided, and to "decline as frame rate fell from 25 to 8.3 Hz." (Barber and Laws, 1994, p. 172). So subjects performed the shadowing task better when a visual signal was provided, and performance decreased as the temporal resolution of the image was reduced. Performance in the lip-reading task was also affected by temporal resolution. Subjects were able to correctly lip-read 70% of the numbers spoken when the frame rate was set at 25 Hz, but accurate identification of numbers declined to 55% when the temporal resolution was reduced to 8.3 frames per second. Reducing the frame rate had a larger impact upon the lip-reading task than upon the shadowing task.

Barber and Laws (1994) summarise their results by suggesting that "for speech-based tasks (which are most likely to be relevant to video conferencing), frame rates should not be below 12.5 Hz and should ideally be in the region of 16.6 Hz and above." (Barber and Laws 1994, p. 174). It should be noted that this suggested frame rate is well above the temporal resolution provided by many publicly available VMC systems, which is often no more than 5 frames per second (Masoodian et al., 1995). Barber and Laws were, however, trying to determine the most appropriate screen parameter settings for a specific application of VMC, its use in a mental health care environment. Whether the results of this research can be generalised to other communicative tasks, or to other video-conferencing situations, still needs to be explored.

6.4 Literature Review: Quality of Video Images Provided by VMC Systems

The following review introduces some of the recent research which has evaluated the impact of a range of VMC systems, which provide video images of different quality. The review highlights the range of factors that can effect the quality of video images, and concludes that there is insufficient research (to date) to obtain an overall impression of how the quality of video images affects human interaction. Experiments, and field studies, which have included the use of high quality video images are reviewed first, followed by research which has examined VMC systems which provide video images of a lower quality.

6.4.1 Effects of High Quality VMC

Sellen (1992; 1995) compared patterns of speech in three multi-party video-conferencing systems; Picture-in-Picture (PIP), Hydra and LiveWire VMC systems. These VMC systems were compared to face-to-face and audio-only contexts, by examining the communicative behaviour which occurred in each context. Each VMC environment displayed the images of participants in a different manner: the PIP system divided the monitor screen into quadrants (one for each participant), the LiveWire system showed just the current speaker, and the Hydra system displayed each participant on a separate monitor. Whilst these three systems varied the way in which the video images were presented, the quality of the visual signal was held constant. Sellen (1995) reports that the video images were of good quality and the signals were not lagged, but gives no further details other than the size of the video windows provided in each system. The video window was approximately four times larger in the LiveWire system then in the PIP or Hydra systems. The audio channel provided full duplex speech, which was synchronous with the visual channel.

Sellen explored the effects of these VMC systems by focusing on two particular aspects of the systems; the amount of selective attention (selective gaze and listening) which each of the systems afforded, and the "amount of visual access to other participants" (Sellen, 1995, p. 412). In the PIP system selective attention is difficult to use, or ascertain, as all of the group members are displayed on a single monitor. This problem also occurs in the LiveWire system, as only the video image of the current speaker is shown to the group (the person holding the floor can see the previous speaker). Selective gaze and listening is easier to use in the Hydra systems, as the spatial array of the individual monitors enables participants to judge which person is being attended to by the members of the conference. The positioning of the video cameras (which were placed either centrally above or below the monitor screens) would, however, have made it difficult for users to make eye-contact in any of the VMC contexts.

In both of the experiments, groups of participants took part in informal debates which lasted for approximately a quarter of an hour. The discussions were recorded and the on-off patterns of speech were computed using a Speech Tracking system. This facilitated analysis of the surface structure of the group discussions, in terms of the number and length of turns, occurrences of simultaneous speech, and formality of floor taking procedures. The first experiment compared the Hydra and PIP systems with face-to-face interactions. Analysis of the structure of the group discussions revealed very few effects of communicative context; turn size and frequency were similar in both of the VMC contexts, and did not vary significantly from face-to-face interactions. Similarly, the number of episodes of overlapping speech did not vary with context. The rate of interruption did, however, vary with communicative context, but only in the first experiment; interruptions occurred more frequently in the face-to-face than the Hydra context. Analysis of the patterns of speech also showed that turn-taking was more formal in the Hydra and PIP VMC contexts than in face-to-face discussions.

In the second experiment the PIP and LiveWire systems were compared to audio-conferencing (VMC with the visual channel disabled). Again non significant differences were observed in the surface structure of the conversations; the frequency and length of turns, and episodes of simultaneous speech were similar in the VMC and audio-conferencing context. The absence or presence of the high quality video images did not significantly effect the structure, or formality, of the group discussions (Sellen, 1992; 1995). However, the subjective data obtained from questionnaires and interviews revealed that participants preferred using the Hydra and PIP systems, as these provided video images of all of the participants, and allowed users to carry out parallel conversations.

Olson, Olson and Meader (1997) found a similar preference for the provision of a visual channel in remote interactions. In a series of studies they explored the differences between face-to-face and remote mediated contexts, which included VMC and audio-conferencing. In the first study (Olson et al., 1992; 1993) participants worked in a face-to-face context, supported either by an electronic workspace (ShrEdit) or using conventional tools (a whiteboard, paper and pencils etc.) Participants in the second study worked remotely, using VMC or an audio-conferencing system. All of the participants in the second study were provided with the ShrEdit workspace.

The VMC system in these experiments provided high quality video and audio signals. The video signals were run over analogue links, and the audio link produced full-duplex sound, with directional input and output. The video output from each participant was displayed on separate monitors, which were placed on either side of the shared workspace to represent the spatial arrangement of the group members. The VMC systems were specifically designed to provide a video-conferencing context that was as ideal as possible (Olson et al., 1997).

Each group of participants was given an hour and a half to design an Automatic Post Office (APO). This task is representative of design tasks in the real world (Olson et al., 1997), and had been carefully researched in field studies before it was used in this series of experiments. The task provides three sets of measures for evaluating the impact of communicative contexts; these are measures of the quality of the work product, participant satisfaction with the process of design and method of communication, and characteristics of group processes (Olsen et al., 1997).

The quality of the product is assessed by evaluating the final document written by each group, using a set of criteria proposed by Olsen et al. (1993). Participant satisfaction with the process of design, task performance and group participation was assessed using post-session questionnaires. The characteristics of the group processes were analysed using a range of measures based on the transcriptions of the meeting. The transcripts were coded to show categories of design activities; these included task management (such as, the time taken to plan and write the final document) and process management activities (organising the group activities, clarification of ideas and design issues). In addition, content analysis of a subset of the VMC and audio conferencing dialogues was also performed, to illuminate the degree of participant 'engagement' (how much participants contributed to the group processes) and the amount of 'critical discussion' (positive and negative evaluations of ideas etc., Olsen et al., 1997) generated by the groups.

Olson et al. (1997) report that the quality of the designs in the VMC and audio-conferencing contexts were very similar; the absence of a video link did not significantly affect the quality of the design. However, when the remote contexts were compared to the face-to-face meetings supported by ShrEdit the absence or presence of a video channel had a pronounced effect upon the quality of task performance. Participants in the VMC and face-to-face contexts produced designs of similar quality, but designs constructed by users of the audio-conferencing system were significantly poorer than those from the face-to-face context. Olson et al. (1997) conclude that "remote work without video is not as good as face-to face" (p. 166).

The questionnaire data, relating to satisfaction with task performance and group discussions, revealed an effect of communicative context similar to the results of task performance. Groups working in traditional face-to-face groups displayed the highest levels of satisfaction with the quality of their discussions. VMC users were just as satisfied in this respect as participants in the face-to-face context who had use of ShrEdit. The audio-conferencing groups, however, were significantly less satisfied with the quality of their discussions. These participants felt that the audio-conferencing system had made it difficult to judge the reactions of other group members, to use persuasion, or resolve disagreements (Olson et al., 1997).

The analysis of characteristics of the group processes showed how each group conducted their work. The main differences occurred between the face-to-face and mediated contexts. Participants in the VMC and audio-conferencing contexts spent "more time managing their work and clarifying what they meant than the face-to-face groups" (Olson et al., 1997, p. 170). The measures of engagement, which were only assessed in the VMC and audio-conferencing contexts, did not vary significantly. For instance, VMC and audio-conferencing meetings were of approximately the same length, switching between speakers occurred just as frequently in both contexts, and the amount of critical discussion was similar regardless of whether or not a video link was provided (Olsen et al., 1997).

One of the main difference between the VMC and audio-conferencing conditions was, therefore, the perceived ease of communication; participants felt it was easier to communicate when they could see each other via the video link, and this increased their sense of satisfaction with the task. The analysis also showed that working remotely - with or without video images - required more effort being expended on establishing mutual understanding; users of the audio-conferencing system spent more time clarifying what they said, and a greater amount of time in managing the work itself. Olson et al. suggest that "there is more sense of what others are doing and what they mean when we are face-to-face than can be present via even very good video channels" (Olson et al., 1997, p. 170). Certainly the VMC system in this provided video images (and audio signals) of high quality, but eye contact was still difficult to achieve due to the placement of the video cameras. It could be that the inability to make eye contact reduced the feeling of co-presence in this VMC context.

Some researchers have examined the effect that facilitation of eye contact may have upon collaboration in VMC. For instance, research by O'Malley et al. 1996; Doherty-Sneddon et al. 1997; and Anderson et al. 1997. These studies used 'video-tunnels' (Smith et al., 1991) to provide an analogue video-link between two remotely situated work-stations. Video tunnels afford high quality visual images, with a frame rate of 25 - 30 frames per second and high spatial resolution (O'Malley et al., 1996). Video images of the head and shoulders of participants were displayed on the video monitors, which measured 7 inches by 5 inches. The laboratory set-up ensured that there were no transmission delays, and that the video and audio signals were synchronised. Full-duplex analogue links were also provided in these studies. In effect this VMC system provided a context that was "as close as possible to face-to-face interaction, without the participants actually being co-present" (O'Malley et al., 1996, p. 180).

Comparisons between two versions of the VMC context were compared with a remote audio-only context (audio-conferencing) are reported by O'Malley et al., (1996) and Doherty-Sneddon et al. (1997). In one of the VMC contexts eye contact between participants was possible, but in the other condition the camera within the video tunnel was re-positioned so that participants could not make direct eye contact (Doherty-Sneddon et al., 1997). To study the effects that these different VMC contexts had upon task performance and the structure and content of the dialogues, thirty-four pairs of subjects completed computer-supported versions of the Map Task; each subject pair completing one task in each of the three communicative contexts.

Task performance, measured in terms of route accuracy scores, was found to be un-effected by communicative context, or ability to make eye contact. However, the way in which the dialogues were structured was significantly effected by communication context. Dialogues in the VMC contexts contained significantly more interruptions than in the audio-conferencing interactions. Rather surprisingly, the dialogues in the VMC plus eye contact context were found to be longer than interactions in either of the other contexts; that is, they "contained significantly more words and turns than the dialogues in either of the other two contexts" (Doherty-Sneddon et al., 1997, p.

28). This lengthening effect is attributed by Doherty-Sneddon et al. (1997) to three factors; attenuation of visual cues, unfamiliarity with the medium and effects of remoteness. For example, being able to make eye contact in a VMC context could reduce the formality that has been found to occur in remote interactions (Sellen, 1992; O'Conaill et al. 1993), which could result in longer conversations in a VMC context that enabled users to make eye contact.

The differences in the structure of the VMC dialogues reported by Doherty-Sneddon et al. (1997) and O'Malley et al. (1996) are interesting, as they vary from the results reported by Sellen (1992; 1996); even though both sets of data were obtained using high quality visual signals. The reason for the conflict between the results is difficult to determine, but might be due to a range of factors. For instance, Sellen and Doherty-Sneddon et al. used different types of task, the number of people who participated in each conference also varied (group discussions or dyadic interactions), and (potentially more interesting) the VMC systems varied in their ability to facilitate eye contact between participants.

Doherty-Sneddon et al. (1997) report one other significant effect of communicative context. The content of the dialogues from the VMC and audio-conferencing contexts was also explored, using Conversational Games Analysis. This analysis examines the pragmatic function of utterances, and illuminates the way in which Conversational Goals are achieved during collaborative problem solving. Conversational Games Analysis has been described in full earlier in this thesis, along with the results from the study by Doherty-Sneddon et al. (1997). The analysis showed that dialogues in the audio-conferencing context contained a significantly higher frequency of ALIGN Games; speakers in this context attempted to elicit a greater amount of verbal feedback from their partners than in either of the two VMC contexts. In the VMC contexts participants can use the visual channel (which was of a very high quality) to convey listener feedback through non-verbal communication; this channel of communication was not available in the audio-conferencing context, so participants made greater use of verbal means of establishing mutual understanding in this context. It would be interesting to see if similar differences occurred when the visual signals afforded by a VMC system were of lower quality; a point that will be returned to later in this chapter.

6.4.2 The Effects of Low Quality Visual Signals in VMC

One study which did compare the impact of high and low quality video images is reported by Masoodian, Apperley and Frederickson (1995). In their experiment twelve pairs of subjects attempted a series of collaborative tasks, solving different jigsaw puzzles, in a range of contexts. The tasks were computer-supported, using a shared work-space system (the 'Aspects' conferencing system) which was run over linked Macintosh computers. Each pair of subjects collaborated in all four communicative contexts; face-to-face, remote audio-only, full motion video and slow motion video[1]. A video link provided the full and slow motion video images, the slow motion video was produced by reducing the frame rate of the full motion video from 25 to 5 frames per second (5 frames per second is the frame rate afforded by many networked VMC systems, Masoodian et al., 1995). The video images, which showed the upper half of participants bodies, were displayed on separate monitors rather than being incorporated into the 'Aspects' conferencing system. In all the remote conditions an audio link was provided between the two rooms, but details of the quality of the audio signal are not given.

Video recordings were made of the verbal and non-verbal interactions. The on-off patterns of speech were generated by a computer program, using a method of analysis similar to the one applied by Sellen (1995). This analysis revealed that the style of verbal interaction was very similar in all four communication contexts. Turn-

[1] Order effects were partially controlled by counter-balancing the order in which participants encountered the first two contexts.

management was not affected by the absence or presence of a visual channel, and changes in the temporal resolution of the video image (from 25 Hz to 5 Hz) appeared to have little effect upon the structure of the dialogues. The results obtained by Masoodian et al. (1995) are a little surprising. Especially in the light of the experiments by Barber and Laws (1994), which illustrated the impact that a low frame rate can have upon speech based tasks. Some of the results reported by Masoodian et al. are in agreement with findings from the second experiment run by Sellen (1992; 1995), where comparisons are made between audio-conferencing and VMC. For instance, both Sellen and Masoodian et al. report that the number of turns and episodes of simultaneous speech did not differ significantly with communicative context. But again, the results conflict with the findings presented by O'Malley et al. (1996) and Doherty-Sneddon et al. (1997).

In contrast to the experimental approaches reported so far, O'Conaill et al. (1993; 1997) adopted a naturalistic approach to the evaluation of VMC systems. As mentioned earlier, these studies evaluated the impact of two VMC systems when they were introduced into the workplace. O'Conaill et al. (1993) explored the impact that these VMC technologies (ISDN and LIVE-NET) had upon real group meetings. The two VMC systems provided very different qualities of video and audio images. The precise quality of the video image is not detailed by O'Conaill et al. (1993); that is, the frame rate, spatial resolution etc. of the video signals are not given. However, the LIVE-NET system was said to provide 'broadcast' quality images, indicating a temporal resolution of approximately 25 frames per second. The audio channel in this system provided full duplex signals. The ISDN system produced video images of a 'poor quality', which were displayed on 26 inch monitors. The audio channel in this system was half-duplex, the switching device was voice activated. In both of these VMC contexts mutual gaze was probably difficult to achieve, due to the central placement of the video cameras.

The findings from these studies have already been commented upon; but its is worth recalling that the speech patterns of dialogues taken from the two VMC conditions showed several significant differences. Participants working in the ISDN system used longer turns, and interrupted each other infrequently in comparison to participants using the LIVE-NET system (O'Conaill et al., 1993). The group discussions in the LIVE-NET context bore a close resemblance to face-to-face interactions, as they were made up of sequences of shorter interactions. The ISDN system, therefore, appeared to have a greater impact on the patterns of speech than the LIVE-NET system. However, the effects that these differences in speech patterns had upon the group discussions could not be judged, as objective measures of task performance were difficult to attain in these naturalistic studies (O'Conaill et al., 1993). The need for laboratory based experiments, which can examine the impact of changes in dialogue structure upon task performance, was acknowledged by O'Conaill et al. (1993).

The research by Anderson et al. (1996) and Anderson et al. (1997) demonstrates a laboratory based approach to evaluating the impact of VMC upon task performance. This research is based upon an experimental paradigm, which allows the researchers to systematically vary the communicative context and observe the effects that this has upon task performance and the structure of the interactions. The task used in these experiments was carefully chosen to be representative of real-world applications of VMC. The task was a simulated service encounter, called The Travel Game (Anderson et al., 1996; 1997). The Travel Game is a collaborative problem-solving task, in which participants (a travel agent and a series of clients) plan an itinerary around the USA. Two separate experiments from the research by Anderson et al. (1996) are outlined here. The first experiment compared collaboration in four communicative contexts, and the second examined the possible benefits of providing a visual channel in long distance video-conferencing.

The first experiment examined the impact of four communicative contexts upon collaboration in the Travel Game. The four contexts were face-to-face, remote spoken, VMC and audio-conferencing. The VMC conditions for the Travel Game were presented on two 20 inch colour monitors on SunSPARC 10 workstations, running NV 3.2 (Frederick, 1994), operating over a dedicated local area network. The VMC system in this study provided moderately low quality video images. The temporal resolution was approximately 4 - 5 frames per second, which is approximately equivalent to the slow motion video used by Masoodian et al. (1995), but the size of the image was much smaller. The video image measured 4.5 inches by 3.5 inches. An audio-link provided full-duplex sound via microphones and headphones for the two mediated contexts.

Anderson et al. (1996) drew comparisons between the face-to-face and remote spoken interactions, and between VMC and audio-conferencing contexts. The effects of context were explored by comparing task performance, measures of the process of communication and decision-making, and user satisfaction. Full details of these experiments will be given later in this chapter, as Study 3 - which examines the impact of impoverished video signal - is based upon the corpus of dialogues collected in these two experiments by Anderson et al.

The results from experiment One (face-to-face versus remote spoken contexts) revealed significant differences in the way the dialogues were structured. In the face-to-face dialogues the Travel Agent said less than in the remote spoken context, but a similar level of task performance was achieved in both contexts. The decision-making process also varied with context. In the face-to-face context the clients asked the travel agent to make more searches for details of flights connections, and made more optional changes to their travel plans than occurred in the remote spoken context. When a similar comparisons were made between the VMC and audio-conferencing contexts, the results were found to be non-significant; task performance, length of dialogues, rate of interruptions, and the decision-making process were very similar in the VMC and audio-conferencing contexts. These results indicate that being able to see who one was talking to had a beneficial effect in the face-to-face context, but this benefit was not replicated in the VMC context.

In a later experiment, Anderson et al. (1996) used a similar procedure to examine the impact of VMC on long distance collaborations. One of the aims of this study was to ascertain whether provision of a visual channel, albeit an impoverished one, was beneficial when participants knew that they were collaborating over a long distance. This study made use of a full range of publicly available video-conferencing network tools (nv, nt and vat). The VMC system was run over the European Internet, between the United Kingdom and the Netherlands. The video and audio signals provided by this VMC system were of low quality; the frame rate of the video image was approximately 5 frames per second, and the audio signals (provided by a Click to Speak audio tool) afforded only a half-duplex channel. During the experiment twelve subjects took part in two Travel Game tasks, one task being completed in the VMC context and the other in an audio-conferencing context (that is, the same VMC system with the visual channel disabled).

Anderson et al. (1996) report that there were non significant effects of communicative context upon task performance and the length of the dialogues. These results indicate that in remote collaborations, the provision of a visual channel was not advantageous; the video images did not facilitate shorter interactions, or alter the level of task performance. However, the subjective feelings of the participants (from post- session questionnaires) indicated that this was not the whole story; the questionnaire data showed that users felt that it was easy to communicate with each other in both contexts, but the provision of a visual channel was considered helpful during long-distance collaborations, as it assisted in "establishing and maintaining social presence" (Anderson et al., 1996, p. 203.)

The research reviewed so far has been shown that task performance appears to be un-effected by communicative context. The use of VMC and audio-conferencing has not been associated with a reduction in task performance. This may not very be surprising, as these studies have asked participants to take part in collaborative tasks (for example, Anderson et al., 1996; Doherty-Sneddon et al., 1997; O'Malley et al., 1996). A long line of research has shown that collaborative tasks are not normally adversely effected by communicative context (for example, Short et al., 1976; Williams 1977).

However, one study reported by O'Malley et al. (1996), and Anderson et al. (1997), has shown that collaborative task performance can be adversely effected if the process of communication is considerably disrupted. This effect was observed when comparisons were made between high and low bandwidth VMC systems, with and without delay in the audio and video signals. These studies again made use of 'video-tunnels' to provide high quality video images, and comparatively low-quality video context was produce using 'video-phones'. Video-phones are designed to run over ordinary telephone lines, but the bandwidth available is limited. This causes delay in the transmission of the video and audio signals of approximately 500 milliseconds. At the same time the two sets of signals get out of synch with each other, due to "the sheer amount of information in the video signal and the limitations of current compression technology. Synchronisation with the delayed video signal leads to a delay in the audio channel" (O'Malley et al., 1996, p. 184).

The quality of the video images provided by video-tunnels and video-phones varied on several parameters, which were not all due to transmission delay. For instance, visual images provided by the video-tunnels system had a frame rate of approximately 25 Hz, high spatial resolution and were shown on a monitor that measured 7 by 5 inches (O'Malley et al., 1996; Doherty-Sneddon et al., 1997). In contrast, the video-phones (BT Relate 2000) provided images with a frame rate is approximately 8 -10

Hz, low spatial resolution, and the size of the image was only 2.2 by 1.6 inches, (private communication O'Malley, 1997). The video-phones provided lower quality visual and audio signals than the video-tunnels, and the signals were subject to delay and a lack of synchronicity. The researchers were interested in assessing the impact of these factors upon collaborative communication.

These factors were explored by examining task performance and the structure of the interactions in four communicative contexts: Video with delay (video-phone), Audio with delay (video-phone with video enabled, but covered-up), Video with no delay (video-tunnel, with eye contact) and Audio with no delay (video-phone working as a normal phone). Pairs of subject were assigned to either the delay conditions, or no-delay condition. Each pair tackling two versions of the Map Task; once with the video link, and once with just an audio link.

Task performance was found to differ significantly between the delay and no-delay conditions. The routes drawn by subjects in the delay conditions were significantly less accurate (approximately 36 % less accurate) than the routes drawn in the no-delay conditions (O'Malley et al., 1996). The provision of video images did not improve task accuracy, in either the delay or no-delay conditions. So in this experiment, being able to see who one was talking to did not enable participants to overcome the effects of signal delay (O'Malley et al., 1996).

Analysis of the structure of the dialogues showed that the lengths of the dialogues - measured in number of turns, words or length of turns- were not effected by delay, or visibility of partner. However, there were significantly more interruptions in the delay than in the no-delay contexts; a similar pattern of results was found when the rate of interruption was examined (O'Malley et al., 1996). Provision of a video-link also affected the number of interruptions, which occurred more frequently when a video signal was available. Anderson et al. (1997) provide a clear indication of the impact that delay and visibility had upon turn-taking; 50% of turns in the video-phone context were interrupted, 40% of turns were interrupted in the Audio with delay context, but only 15% of turns were interrupted in the two no-delay (video-tunnel) conditions. The very high rate of interruptions in the video-phone context "suggests that delay in the audio-visual signal leads to difficulties in managing turn-taking which severely disrupt performance" (O'Malley et al., 1996, p. 187).

The previous study has demonstrated the striking effect that transmission delay can have upon communication and collaboration. The effect of a small delay in transmission of the audio signals has also been commented upon by other researchers (for example, Tang and Isaacs. 1993; Isaacs and Tang, 1994), who report that "small delays in the audio can disrupt the participants' ability to reach mutual understanding and reduce their satisfaction with the conversation" (Isaacs and Tang, 1994, p. 68). The effects of delay, and lack of synchronity between audio and video signals, are obviously important factors which need careful consideration in the design of VMC systems. Which of these two effects should be avoided may depend on the type of task being undertaken, and the number of participants using the VMC system.

6.4.3 Summary of Review: Quality of Video Images in VMC

The review has demonstrated the wide range of VMC systems, and methods of evaluation, employed by researchers as they explore the impact of VMC upon communication and collaboration. These factors probably explain why there appears to be so little consensus of opinion concerning the effects that the quality of the video image has upon the process of communication. As Finn (1997) points out, it is difficult to make direct comparisons across studies when the VMC systems and methods of evaluation vary so greatly.

The review of literature has also illustrated the tremendous variation in quality of the video images provided by different VMC systems, though the exact quality of the video images was not always easy to assess. Some of the studies gave exact measures (for example, frame rate and size of the video image, etc.) whilst other researchers simply stated that the quality of the image was of 'low' or 'high' quality. For example, Sellen (1992; 1995) stated that the quality of the video images was 'good', with no transmission lag, and gave the sizes of the images provided by each of the VMC systems. O'Malley et al. (1996) offered information on temporal resolution, size of the video images, stated that the spatial resolution was 'high', and that the signals synchronised. Whilst it was usually possible to categorise the quality of the video image as either 'high' or 'low', the different dimensions used to describe the quality of the video image means that it is difficult to make comparisons between individual studies. Interpreting and comparing results across studies is made even more complicated by the variety of factors that can effect the quality of the video images; such as, temporal resolution, available bandwidth, or the size of the video image.

6.5 Restricted Scope of Video-Mediated Images

Even when the VMC systems being evaluated provide high quality video images, the amount of information in the visual signals is restricted when compared to face-to-face interactions. The visual images afforded by VMC systems provide very restricted views of participants. In most of the studies reviewed, the video image portrays the head and shoulders, or head and upper body, of each participant; very little of the surrounding environment is included in the picture. This is very different from face-to-face interactions, where we can see each other and the surrounding environment. The visual information available in VMC is also restricted in another sense; in most VMC systems it is difficult for participants to establish eye contact. This is due to the placement of the cameras, which are usually mounted over the centre of the monitor (Fussell and Benimoff, 1995).

One of the consequences of the reduced quality and scope of video images in VMC, is that it can be difficult to use and interpret non-verbal communication in this context. A naturalistic study by Heath and Luff (1991) illustrates this point. Heath and Luff observed the way that people interacted (verbally and non-verbally) when using audio-visual technologies for informal and groupwork purposes. Their analysis of over 100 hours of recordings of video mediated interactions, showed that some forms of visual conduct (for example, gestures) "appear to lose their communicative impact when performed through video" (Heath and Luff, 1991, p. 100). For instance, the communicative content of iconic gestures was lost in the VMC context, even if the gestures were exaggerated (Heath and Luff 1991).

The last section has shown that there are a number of factors which can effect the impact that video-mediated images may have upon communication and collaboration. The quality of the video images will be determined by the VMC technology being used, but even when high quality images are provided the visual information available to users is greatly restricted when compared with face-to-face interactions.

6.5.1 Motivation for Study 3

The review of the literature available on the quality of video images in VMC systems has shown that the video image can be affected by many factors, and that the findings of research to date are not conclusive. There does seem to be some evidence to support the view that the process of communication in VMC differs from face-to-face or audio-conferencing interactions, and that this may be partly due to the quality of the video images. However, most of the studies which have examined the process of communication in VMC have done so by looking at the **structure** of interactions in VMC; for example, Sellen, 1992; 1995, O'Conaill et al., 1993). There has been very little in-depth analysis of the **content** of contributions in this communicative context. On the few occasions that this level of analysis has been carried out, the VMC systems provided high quality visual signals (for instance, Doherty-Sneddon et al., 1997). The effects of low quality video images (low temporal resolution) upon the structure and content of interactions are still relatively unknown.

6.5.2 Aims of Study 3

Study 3 attempts to add to the literature, by exploring the effects of a VMC system which provides low quality visual images (frame rate approximately 5 HZ), but high quality (full duplex) audio signals. The effects of this VMC context will be compared with face-to-face communication, to see what impact the impoverished video channel has upon communication and collaboration. This can be assessed using the multi-dimensional approach introduced in Study 1. This involves examining the effects of communicative context upon task performance, and exploring how this relates to the structure and content of VMC and face-to-face interactions. The structure and content of the VMC dialogues will be explored by Conversational Games Analysis. The aim of this study is to see if the quality of the visual signals in the VMC context (which mirrors the quality of video images provided by many commercially available VMC systems) effect the processes of communication and collaboration.

6.6 Study 3. Exploration of the Impact of Low Quality Video Signals Upon Collaboration and Communication

Study 3 is based upon the corpus of dialogues collected by Anderson et al. (1996). The design of the Travel Game studies will be explained first, followed by some of the results reported in Anderson et al. (1996) to show the general pattern of communicative behaviour in this task. Details of the application and results of Conversational Games Analysis on the face-to-face and VMC protocols will then be presented.

The Travel Game studies were first reported by Anderson et al. (1996). The Travel Game consists of an Air Travel Consultancy task, in which participants act as ‘clients’ who collaborate with a ‘Travel Agent’ whilst they plan a trip around the

USA. Various restrictions are imposed on the planned itinerary; such as the length of time clients must stay in any State they plan to visit, and financial penalties (surcharges) are imposed if they choose to backtrack to a previous destination or use another Airline to make a connection.

The Travel Game is therefore a collaborative task, similar to the Map Task, but it is closer to tasks undertaken in the real world. It can be viewed as a service encounter task, where information from an 'expert' is sought for a purpose. Because of the nature of the task, The Travel Game imposes less control on the progress of communication than the Map Task, where participants follow a series of communication goals which are determined by the route drawn on the map. The Travel Game still requires participants to complete a series of collaborative goals, but the order in which the goals are encountered is not predetermined. For example, the traveller is free, within the rules of the Game, to fly to any city marked on the map, and in any order he/she desires.

It should be noted that the role of the Travel Agent in each of the two Travel Game Experiments was played by a research associate, who maintained that role throughout each experiment. This is in contrast to the Map Task, where pairs of subjects interact with each other and (in the original corpus) take turns acting as either the Instruction Giver or Instruction Follower. In Study 1 the way in which participants adapted to the restrictions of CMC were explored by observing the changes made, by *both* participants, to the processes of communication and collaboration over a series of tasks. In the Travel Game experiments adaptation to a context (such as VMC) could be assessed by noting the changes made by the Travel Agent, as they interact with a series of clients. The adaptations made by the Travel Agent over time, however, will not be discussed in detail in the chapter, as this is currently outside the scope of the present research.

The first Travel Game study (reported in Anderson et al., 1996; Newlands et al., 1996) compared four modes of communication: face-to-face, remote 'audio-only', VMC and audio-conferencing (that is, VMC with the visual channel disarmed).

6.7 Method: Experiment 1. Face-to-face and Remote Audio-only Contexts

6.7.1 Design

A between groups design was used, with communicative context as the between group factor.

6.7.2 Subjects

20 subjects volunteered to take part in the first experiment of this study, which offered a cash prize of £10 to the participants who constructed the best itinerary. Ten subjects participated in each mode of communication. The participants were all students at the University of Glasgow. The mean age of the subjects was 20.4 years (range 18 to 25 years) and the sample consisted of 4 male and 16 female students.

Familiarity

All the participants were unfamiliar with the 'Travel Agent', who was a third year Psychology undergraduate student taking part in the study as part of her practical course work. The Travel Agent was trained in her role during trial runs of the Travel Game. She was instructed to work within a loose framework of prompts during her interactions with the Clients, to ensure that each client was offered a similar amount of assistance throughout the trials.

6.7.3 The Task and Materials

The client's goal is to plan a 2 week tour of the United States of America; visiting as many cities and States as they can, using an 'Atlantic Airline Travel Pass'.

Both the consultant and the client have paper maps of the USA in front of them, showing the different States, and the airports they may visit using the Air Travel Pass. The Consultant also has the flight timetable for the airline being used (fictitiously called 'Atlantic Airline'), which gives details of connections between airports, times of flight departures and arrivals.

The maps of the USA are printed onto A3 paper (29.7 cm by 42.0 cm), the airports used by 'Atlantic Airline' are labelled and represented by coloured dots. The outlines of each State are also represented on the map, and the States are clearly labelled. An illustration of this Map can be seen in Appendix G.

6.7.4 Procedure

On arrival the subjects were randomly allocated to either the face-to-face or the remote 'audio-only' condition. In the face-to-face condition the client and the Travel Agent were seated at either end of a long table (approximately 10 feet apart), and wore head-set micro-phones to ensure that the audio records of the transactions were of good quality. In the remote audio-only (henceforth remote spoken) condition, the clients were shown into an adjoining room, and seated at a desk. In the remote spoken condition the Travel Agent and client talked to each other over an audio-link, again using microphones and headsets to obtain good quality full duplex speech. A floor plan depicting the experimental set-up is provided in Appendix L.

Before starting the task the clients filled in a short questionnaire, which elicited demographic and personal details. The Questionnaire also ascertained whether the clients had any relevant travel experience, such as extensive travel around the USA. None of the volunteer clients had to be excluded from the study on this basis.

The clients in both the face-to-face and remote spoken conditions were given the following instructions:

Instructions for the Travel Game

"You have map of the USA in front of you which you will use to plan your 14 day trip. You have $600 to spend on travel, but $499 has already been spent on an 'Atlantic Air Pass', which enables you to travel on any Atlantic Air path. The Travel Consultant has details of these air paths and the timetable.

The object of the game is to plan to visit as many cities as you can during the trip. The game will last in the time 20 minutes. Both the Consultant and yourself will keep an itinerary of your progress on the sheets provided.

Rules of the Travel Game

Your trip starts when you fly from Glasgow to any airport shown on the Map. You must stay for at least 24 hours in each State you visit, but you can travel within State in that time if you wish, provided you stay a minimum of 4 hours in each city, and there are connections available.

There will be a surcharge of $50 if you backtrack to an airport you have already visited, or if you choose to make a flight with another airline.

You may re-start your itinerary again at any point in the time allowed. The final itinerary which covers the most cities will win the prize money.

GoodLuck.:"

A set of instructions were also made for the Travel Consultant. These consisted of a set of prompts which were only to be used if a long pause occurred in the proceedings. The prompts were as follows:

Where would you like to go next?

Are you happy with that arrangement?

Do you wish to know what other destinations are available from that airport?

Do you wish to start you itinerary again?

Do you wish to know what other flight times are available?

6.7.5 Apparatus

An audio-link was used between the two rooms in the remote spoken condition, and these participants used microphones/audio headsets (SHURE SM2). The face-to-face participants used microphone only headsets to enable good quality recordings of the transactions. In both conditions the spoken dialogues were recorded on DAT (Sony Digital Audio Tape Deck DTC-1000ES) and analogue tape (DENON Precision Audio Component/Stereo Cassette tape deck DRM-500)

6.7.6 Transcriptions

Full orthographic transcriptions of the dialogues were made from the audio tape recordings.

6.8 Experiment 2. Method for Computer Mediated Versions of the Travel Game

6.8.1 Design

A between subjects design was used. Communication context was the between subject factor, with two levels: Video Mediated Communication with a video-link, and VMC without a video-link (audio-conferencing)

6.8.2 Subjects

20 subjects from Glasgow University population volunteered to take part in this Study. The Sample consisted of 15 females and 5 males, whose average age was 21.25 (range 18 to 28 years).

Computer experience of Participants

A short questionnaire was used to ascertain the level of computing experience of the participants. A copy of the questionnaire is shown in Appendix H. All of the subjects were familiar with working on computers, and had previous experience of using a 'mouse'. Sixteen of the 20 subjects had used some form of electronic-mail, but none of the subjects had any previous experience of VMC.

Familiarity

None of the subjects were familiar with the Travel Agent, who was a member of the Research Team but naive to the purpose of this study. The Travel Agent was instructed to work within a loose framework of prompts during her interactions with the Clients, to ensure that each of the clients received approximately the same amount of assistance through out the trials. The set of prompts were identical to those used by the Travel Agent in experiment 1.

6.8.3 The Task and Materials

The task was the Travel Game. Minor changes were made to the task so that it could be accommodated by the VMC technology. The number of airports shown on the VMC maps was reduced to 77, distributed as evenly as possible over the USA. Where-ever possible 2 airports in each State were represented on the map, but in some States (e.g. South Dakota) 'Atlantic Airline' had access to only one airport. The rules of the Travel Game were changed slightly to allow for the reduction in possible flight connections; this was achieved by increasing the number of days (from one day to three days) that the clients had to stay in each State.

The Travel Game was presented via a Video Mediated Communication system run between two Workstations, in two separate rooms. A floor-plan of the experimental set-up is provided in Appendix N. A shared screen facility was used to present the map of the USA in a window on each monitor, and a whiteboard was provided so that the Travel Consultant could keep an update of the planned itinerary.

The client's version of the map of the USA had the States clearly outlined and labelled. The airports were represented as un-labelled dots. Clients could determine the identity of an airport by moving the mouse pointer near to the dot, where upon the name of the airport appeared at the bottom of the map window. The Travel Agent's map was similar to the client's, but had an additional feature; when the Agent clicked on a dot representing an airport, a file containing the details for flights to other 'Atlantic Airline' airports was displayed on the monitor.

In the VMC condition a third window (11.5 by 9 cm) was used to display the video image of each participant. The video cameras were positioned so that a 'heads and shoulders' view of the participants was available. In audio-only conferences the visual window was disabled, and only the spoken channel was made available to the participants; the video window was closed and moved to the bottom of the monitor screen, so that the VMC systems were kept as constant as possible in both contexts. In both VMC conditions full duplex audio was provided via an audio-link between the two rooms. An illustration of the Monitor set-up for the VMC context is shown in Appendix J.

6.8.4 Apparatus

Two SunSPARC 10 workstations were used to run both VMC conditions, these had 20 inch colour monitors. Publicly available software was run over a dedicated local network, the video-conferencing software was nv 3.2 Frederick, 1994, with a refreshment rate of 4 -5 frames per second. JVC Videomovie GR.AX60 compact VHS recorders were positioned above the monitors, to enable a video of the participants head and shoulders to be transmitted in the VMC context.

Lightweight SHURE headsets and SunSPARC microphones were used to transmit the spoken output via the audio-link between the two rooms. This provided the subjects with full duplex (open channel) audio. The spoken output was also combined (using a MACKIE micro Series 1202 12-Channel Mic/line Mixer), and then recorded directly onto DAT (Sony DAT DTC-1000ES) and analogue (using JVC KD A33 stereo Cassette Deck).

6.8.5 Procedure

On arrival, participants were alternately allocated to the VMC condition or the audio-conferencing condition. The subjects then filled in brief questionnaires on their personal details and their previous experience of computers. The monitor set-up was explained to each participant, and they were presented with written instructions:

Instructions for VMC task and communication

"You have a map of the USA in front of you which you will use to plan your 28 day trip. You have $600 to spend on travel, but $499 has already been spent on an 'Atlantic Air Pass', which enables you to travel on any Atlantic Air path. The Travel Consultant has details of these air paths and the timetable. The object of the game is to plan to visit as many cities as you can in the time allowed, which is 20 minutes. The Consultant will keep an itinerary of your progress in the Itinerary box on the monitor screen

Rules of the Travel Game

Your trip starts when you fly from Glasgow to any airport shown on the Map. You must stay for at least 3 days in each State you visit, but you can travel within State in that time if you wish, provided there are connections available.

There will be a surcharge of $50 if you backtrack to an air port you have already visited, or if you choose to fly with another Airline.

You may re-start you itinerary at any point in the time allowed. The final itinerary which covers the most cities will win the prize money.

You can find out the names of each airport (dot) by moving the pointer near to the dot, the name of the airport then appears at the bottom of the map window.

Good Luck."

On completion of the task the subjects were thanked for their participation, and the aim of the experiment was explained to them. Details of the winners of the competition were notified when the study was completed.

6.8.6 Transcriptions

Full orthographic transcriptions were made of the 20 dialogues from the audio tape recordings made during the study.

6.9 Results

The main results from Anderson et al. (1996) will be given first, to set the scene for the analysis of the dialogues by Conversational Games Analysis. The Results reported by Anderson et al. were made by comparing the communication contexts investigated in experiment 1 - face-to-face and remote spoken - with the two VMC contexts from experiment 2. Comparisons between the two experiments were not made because of the changes made to the Travel Game for the VMC conditions.

6.9.1 Task Performance

In the Travel Game studies the number of cities shown on the completed itinerary was taken as a measure of task performance. The means for experiments 1 and 2 are given in the table below, the standard deviations are given in brackets.

Table 6.1 Mean number of cities visited in each communicative context.

Experiment 1	face-to-face	remote spoken
	16.60 (2.87)	19.3 (2.06)
Experiment 2	**VMC**	**audio-conference**
	13.9 (2.42)	14.5 (2.37)

Anderson et al. (1996) report non significant differences in task performance (p>0.10). The itineraries contained a similar number of cities in the face-to-face and spoken contexts, and this pattern was replicated in the second experiment. These results follow the pattern predicted by previous research, communicative context has little effect on task performance in collaborative tasks.

6.9.2 Dialogue Measures: Number of Words per Dialogue

To see if communicative context made any difference to length of the dialogues, the number of words spoken in each dialogue by the Travel Agent and client was calculated. The group means for each communicative context are given below, split by the role of the participant. The standard deviations are given in brackets.

Table 6.2 Group mean words per dialogue, split by role of participant.

context	Travel Agent	client
face-to-face	1356 (66.12)	621 (263.24)
remote spoken	1635 (231.03)	572 (190.03)
VMC	800 (269.83)	337 (180.30)
audio-conference	707 (319.72)	422 (273.07)

Anderson et al. analyse the word length of dialogues in the two experiments separately. The findings from experiment 1 showed that Travel Agents said more than clients, and that there was a significant interaction between communicative context and role of participants ($p<0.005$). Travel Agents used fewer words (22% less) in the face-to-face context than when they were using an audio link between two rooms. Travel Agents also said significantly more than clients in the two computer-mediated contexts ($p<0.001$). The interaction between context and role was again significant ($p<0.05$). Simple Effects analysis of this interaction showed that the effect was not due to communicative context, but to the Travel Agent saying slightly more in the VMC context. Whether the Travel Agent said less, or more, as she became accustomed to the working in this VMC context, and more experienced with the role she was playing, has not as yet been determined. This is an interesting point that deserves further research and consideration.

In summary, the Travel Agent said substantially less in face-to-face than in remote spoken interactions (a decrease of 22%). This result is interesting, as it replicates the length advantage for face-to-face dialogues first reported by Boyle et al. 1994, who compared the amount of linguistic output between face-to-face and spoken (co-present) trials of the Map Task. Comparisons of word length of dialogues between the two mediated contexts again showed that Travel Agent spoke more than the client in both experiments, however the length advantage for the Travel Agent in face-to-face tasks was not replicated in the VMC context.

6.9.3 Structure of the dialogues: Rate of interruption

Anderson et al. also explored the structure of the interactions, by looking at the turn-taking procedures in each communicative context. This analysis examined the frequency of interruption and rate of interruption. The rate of interruptions (that is, the number of interruptions per 100 turns of dialogue) in each context will be reported here. This form of analysis makes allowance for the different lengths of the interactions in all four communicative contexts, allowing comparisons to be made between the data collected in the first and second experiment.

Table 6.3 Rate of interruption in each communicative context.

face-to-face	remote spoken	VMC	audioconference
13.8%	14.5%	10.9%	12%

The differences between the mean rate of interruption in each context was computed, using Analysis of Variance. This produced a significant main effect of context ($p<0.01$). Further analysis (by post hoc tests) showed that the only significant differences occurred between the two extreme ends of the distribution of the means, between the rate of interruption in the remote spoken context and VMC context ($p<0.05$). The effect of visibility of partners in each experiment was non-significant; that is, non significant differences were found between the rate of interruption in the face-to-face and remote spoken contexts, and between the two VMC contexts (VMC and audio-conferencing). Anderson et al. (1996) suggest that the decreased rate of interruption in VMC could be due to two factors; the benefit of being able to see one's partner in this context, and the more formal style of interaction that appears to occur in VMC contexts.

The results reported above are the starting point for Study 3, which is a detailed exploration of the impact that an impoverished visual signal has upon the processes of collaboration and communication.

To examine whether the quality of the visual signals available in VMC affects the content and structure of dialogues, Conversational Games analysis was applied to the VMC and face-to-face Travel Game dialogues. As each twenty minute dialogue takes approximately 6 hours to code, time constraints made sampling essential. Therefore, only the dialogues most relevant to the aims of this Study (to determine the effects of an impoverished visual channel upon collaboration) were included in the

Conversational Games Analysis; these were the dialogues from the VMC and face-to-face contexts.

6.10 Conversational Games Analysis of the Travel Game Dialogues

The main objective in the following analysis is to determine whether there are significant differences in the structure and content of the dialogues in the face-to-face and VMC versions of the Travel Game, and to what extent any variations are caused by the different qualities of the visual signal in these two communicative contexts.

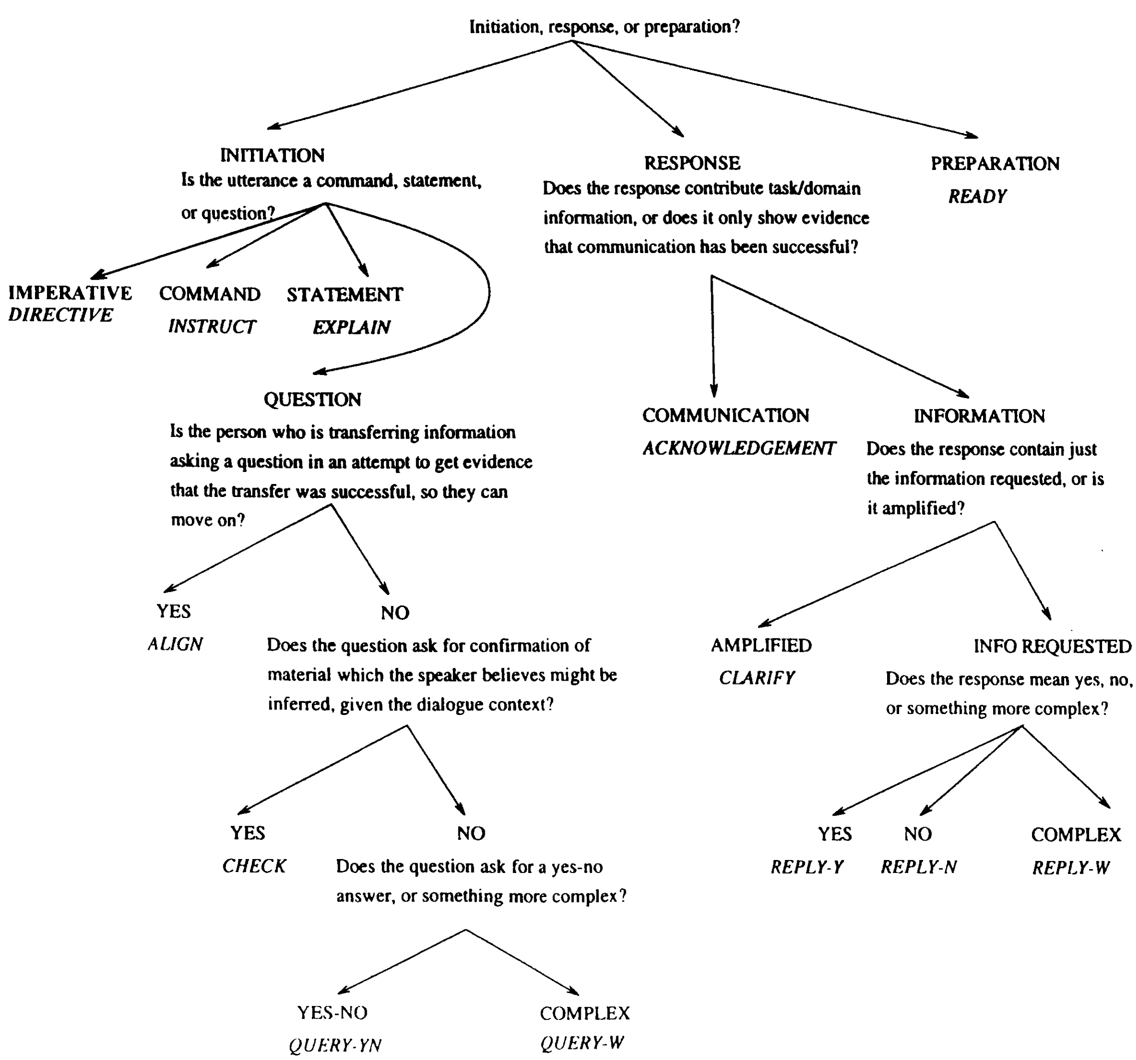
Conversational Games Analysis was applied to the VMC and face-to-face dialogues by two coders. Coder 1 was more experienced in the application of Conversational Games Analysis than Coder 2. The second coder was first trained in the application of Conversational Games analysis, by coding dialogues from the Map Task Corpus, before commencing work on the Travel Game dialogues.

As this was one of the few times that Conversational Games Analysis had been applied to a task other than the Map Task, the coders were interested in seeing whether the same functional categories of Conversational Games could be applied to another task. Only one new Conversational Game was required to enable coding of every utterance in the Travel Game dialogues. This new Game was called a 'DIRECTIVE'. In the Travel Game dialogues, the DIRECTIVE was normally initiated by the client, without any elicitation from the Travel Agent. The function of this new Game is to inform the Travel Agent of a decision made by the client. The working definition was as follows:

DIRECTIVE Game: an imperative type of Game, where the initiator tells the other participant what the initiator has decided to do. The Game is not a response to an earlier question, and may occur when the initiator has changed his/her mind about an earlierdecision.

The DIRECTIVE is similar to an INSTRUCT Game. However, whereas the objective of an INSTRUCT Game is to get one's partner to carry out some action (such as draw a line round an object), the function of a DIRECTIVE is to inform the Travel Agent of the client's next decision; this may implicitly require the Agent to take some action. The DIRECTIVE Game should, therefore, be positioned in Conversational Games Analysis decision tree (Carletta et al. 1997, p. 15) adjacent to the INSTRUCT Game, this is illustrated in Figure 6.1 below.

Figure 6.1 Decision tree for Conversational Games Analysis of Travel Game dialogues.



The following examples of DIRECTIVE Games are taken from different Travel Game dialogues, and demonstrate how this new type of Game was initiated during the Travel Game. The examples were all spoken by participants acting as the Client in the Travel Game.

Extracts from face-to-face dialogues illustrating initiation of DIRECTIVE Games.

1) Right, I'll fly to Memphis to start with

2) and then I'll move to Nashville

3) okay and then just straight from Dallas to San Antonio

4) okay. Then I'll go from San Antonio to Waco

Extracts from VMC dialogues illustrating initiation of DIRECTIVE Games.

5) Ehmm, I'd like to go to ehmm Tennessee, Nashville, Tennessee.

6) Right um I'm going to fly into Maine. In Portland

7) and I'll move on to ... this one here, Tulsa.

8) And then to Louisville.

The extracts illustrate the point that the DIRECTIVE Game functioned in a very similar manner in both communication contexts. The DIRECTIVE was used very sparingly, it accounted for 1.45% of all the Games in face-to-face interactions, and 1.64% of Games in VMC.

After one of the dialogues from the Travel Game corpus had been coded with Conversational Games Analysis, the coding was checked by one of the originators of the scheme (Doherty-Sneddon, University of Stirling). This ensured that the functional definitions had been applied appropriately to the Travel Game dialogues. The two coders then applied Conversational Games Analysis to the rest of the dialogues in the face-to-face and VMC conditions; each coder analysing half of the face-to-face and VMC dialogues.

The following extract gives an example of a dialogue from the face-to-face context of the Travel Game, and shows application of Conversational Games Analysis to this new task. In this extract 'TA' refers to the Travel Agent, and 'C' to the client. Conversational Games are indicated above the text of the dialogue in upper case, and Conversational Moves are shown underneath the text in italics. The start and end of each Game is shown.

Example of coded dialogue. Face-to-face.

Game 1: QUERY-W

TA: where would you like to go from Syracuse?

Move: Query-w

Game 2: QUERY-YN embedded

C: I can still go, I have to be still in New York

Move: Query-yn

Game 3: EXPLAIN embedded

TA: you have just to stay in New York until 5.30 that day/

Move: Explain

C: 5.30 >

Move: Acknowledge

End Game 3, End Game 2.

Game 4: QUERY-W embedded

<TA: you could just stay in Syracuse until 5.30 and choose to st.. fly out of state then if you wish/

Move: Query-w

C: yeah>

Move: Acknowledge

TA: or you could go to another city in the meantime?

Move: Query-w cont

C: no l think l will stay

Move: Reply-w

TA: right

Move: Acknowledge

End Game 4.

TA: so where would you like to fly then

Move: Query-w (continuation of Game 1)

C: l would like to fly to ehmm ... let me see ... Detroit

Move: Reply-w

TA: to Detroit ... uh huh ...

Move: Acknowledge

Game 5: EXPLAIN em

C: its in Micshigan, Michigan

Move: Explain

TA: to Detroit in Michigan,

Move: Acknowledge

End 5.

Game 6: EXPLAIN embedded

TA: I am sorry there isn't actually a connection between those two airports.

Move: Explain

C: right, ehmm

move: Acknowledge

End Game 6

6.10.1 Reliability of Coding

An inter-coder reliability test was performed on one dialogue. Out of a possible 177 coded utterances, only 15 were found to have been coded differently by the two coders; giving an inter-judge agreement of 91.5%. The Kappa coefficient of agreement for nominal data (see Siegel and Castellan, 1988) was calculated, to check the overall consensus between the two coders, and to ensure that agreement between coders was not due to chance factors alone. Agreement on the classification of each Conversational Move (all Moves, including responses were include in this analysis) was calculated, giving a kappa of 0.94 ($N=177, k=2$). Testing the significance level of the obtained kappa showed that it differed significantly from any agreement due to chance factors alone, for example by randomly allocating functional categories to each utterance ($p<0.001$).

It was noted during this process that there were a few occasions (4 in total) where there was disagreement over the classification of an Initiating Move, that is the coders disagreed whether the Move was a continuation of a previous Move, or marked the beginning of a new Initiating Move. Each of these 4 occasions involved *Explain*

Moves. Further analysis was carried out, taking both classification and boundaries of the Moves into account, which produced a percentage agreement of 89.4%, and a kappa statistic of 0.864 (N= 180, k=2). The observed value of the Kappa statistic was then tested to see if it was greater than a value obtainable by chance, and was found to be significant ($p<0.001$).

The consensus of agreement on classification, and boundary, of units was sufficient to allow the dialogues of both coders to be combined. Perfect agreement would be indicated by a kappa value of one, and a Kappa greater than 0.8 is generally agreed by researches (for instance, Krippendorff 1980) to be sufficient for the analysis of data subjectively classified in this manner (Carletta et al., 1997, p. 25). Therefore, the data from both coders could be collapsed, and comparisons made between the two contexts from the VMC and face-to-face contexts.

6.11 Results of Conversational Games Analysis

Conversational Games Analysis can either be carried out at the level of the Conversational Games, or at the more detailed level of the Conversational Moves. In this thesis the more micro level of analysis is used, examining how frequently each type of Initiating Move occurred in a dialogue. The unit of analysis is the dialogue, rather than each participants separate contributions.

6.11.1 Standardisation of Frequency of Initiating Move per Dialogue

Due to the differences in length of dialogues in the face-to-face and VMC contexts, the results of Conversational Games Analysis can not be directly compared between the two contexts unless the data is standardised in some way. Since the analysis is being carried out in this study at the level of Conversational Moves, which frequently correspond to speaker turns, the data from the Conversational Games Analysis was standardised to represent the frequency of Initiating Moves in every 100 turns of dialogue.

The frequency with which each type of Initiating Move occurred in the VMC and face-to-face contexts was calculated, and the standardised scores (per 100 turns of dialogue) were obtained. The raw data is given in Appendix K, and the group mean standardised scores for the VMC and face-to-face dialogues are given in Table 6.4.

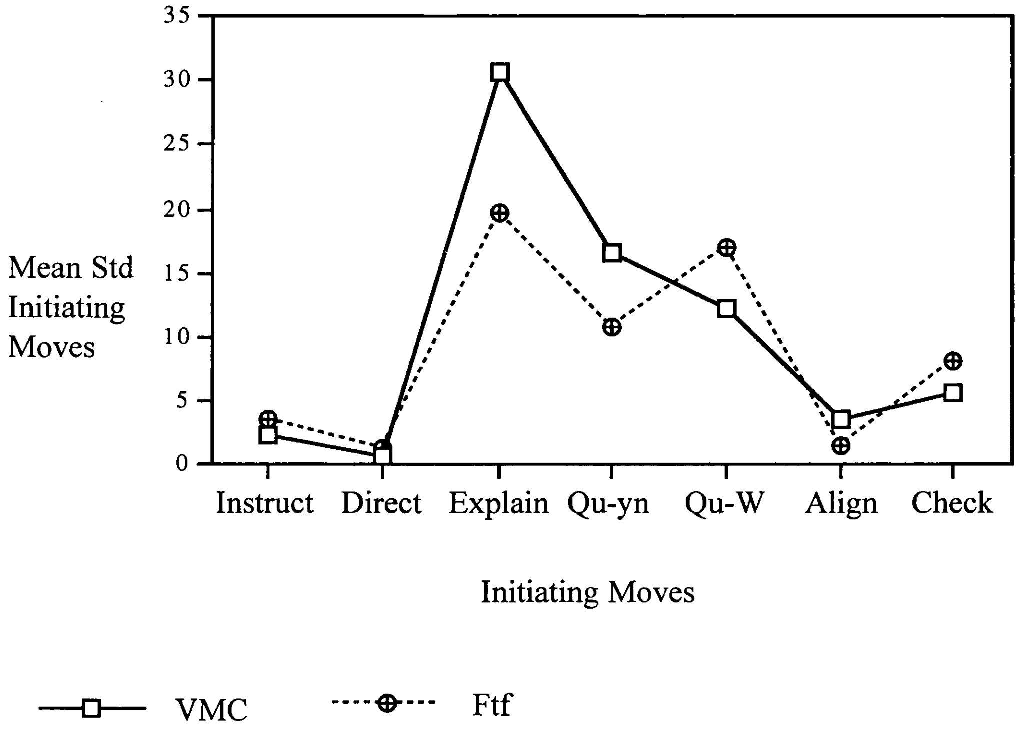
Table 6.4 Mean standardised frequency of Initiating Moves

Initiating Moves	**VMC**	**face-to-face**
Explain	30.62 (11.87)	19.82 (7.00)
Query-yn	16.55 (4.57)	10.92 (4.12)
Query-w	12.18 (5.37)	17.05 (7.11)
Check	5.56 (3.36)	8.20 (3.27)
Align	3.68 (2.54)	1.57 (0.95)
Instruct	2.38 (1.74)	3.63 (2.42)
Directive	0.68 (0.63)	1.23 (0.98)

The Group Means in Table 6.4 show that some Initiating Moves were used more frequently in the Travel Game than others. The distribution of Initiating Moves is different from the typical pattern of Initiating Moves found in the Map Task. The proportional usage of Initiating Moves shows clearly the nature of the Travel Game task. The task is primarily one of seeking and giving information, demonstrated by the large number of *Explain* Initiating Moves (giving information) and frequent use of open-ended and yes-no type questions (*Query-w* and *Query-yn*).

Table 6.4 also highlights some potentially interesting differences in the use of Initiating Moves between the two contexts. For example, the VMC dialogues appear to contain proportionally more *Explain* Initiating moves than the face-to-face dialogues, suggesting that speakers offered a greater amount of unsolicited information in VMC than in face-to-face collaborations. The means from Table 6.4 are presented graphically in Figure 6.2 below, to illustrate the effect of communicative context upon the distribution of Initiating Moves.

Figure 6.2 Graph of Interaction of Context and Initiating Moves



In order to investigate these differences in detail analysis of variance was computed for each category of Initiating Move. The role of the participant (Travel Agent or client) was included in these analyses as the effect of speaker role as this had not previously been determined in this new task, and the Travel Agent has already been shown to say considerably more than the client, and was therefore likely to initiate more Conversational Games than the client. The means for the standardised data (Initiating Moves per 100 turns of dialogue) involved in these ANOVAs are shown in Table 6.5 below, and the raw data is presented in Appendix M.

Table 6.5 Mean Initiating Moves by Travel Agent and Client in the VMC and Face-to-face Contexts (Standardised data).

Context	**Face-to-face**		**VMC**	
Role	**Travel Agent**	**Client**	**Travel Agent**	**Client**
Instruct	0.17 (.033)	3.46 (1.22)	0.38 (0.55)	1.99 (1.34)
Directive	0.04 (0.1)	1.19 (1.99)	0.00 (0.00)	0.66 (0.63)
Explain	17.01 (5.99)	2.80 (4.05)	27.64 (9.22)	4.50 (3.99)
Query-yn	5.29 (2.33)	5.63 (3.44)	3.40 (1.88)	12.55 (4.42)
Query-w	13.71 (8.25)	3.72 (2.08)	7.78 (2.3)	4.40 (2.77)
Align	1.05 (0.75)	0.52 (0.66)	3.29 (2.58))	0.39 (0.54)
Check	2.13 (1.24)	6.06 (3.53)	2.75 (2.35)	2.80 (1.75)

6.11.2 Results of separate two-way ANOVA on each category of Initiating Move

In each of the seven ANOVAs communicative context (VMC vs face-to-face) was treated as a between group factor, with the role of the participant (Travel Agent vs client) as a within dialogue repeated measure. The results from these ANOVAs are givenbelow.

***Instruct*, *Directive*, *Query-w* and *Check* Initiating Moves**

The ANOVAs based upon the data for Instruct, Directive, Query-w and Check Initiating moves showed only non-significant main effects and interaction effects ($p>0.1$).

***Explain* Initiating Moves**

The analysis showed that there was a significant main effect of context [$F(1,18) = 6.33$, $p<0.05$]. A greater number of *Explains* were initiated in the VMC context than in face-to-face interactions (means being 15.32 vs 9.91 respectively). The main effect of role of participant was also significant [$F(1,18) = 119.91$, $p<0.001$]; the Travel

Agent initiated a greater number of *Explain* Moves than the client (means 22.17 vs 3.05). The interaction between communicative context and role of participant was also significant [$F(1,18) = 7.94$, $p<0.05$]. Further analysis by Simple Effects showed that the Travel Agent initiated more *Explain* Moves in the VMC context than in the face-to-face context [$F(1,18) = 8.82$, $p<0.01$]. The proportion of *Explain* Moves initiated by the client did not differ significantly with communicative context ($p<0.1$)

Query-yn Initiating Moves

The analysis showed that there was a significant main effect of context [$F(1,18) = 8.37$, $p<0.01$]. A greater number of Query-yn were initiated in the VMC context than in face-to-face interactions (means being 8.28 vs 5.46 respectively). The main effect of role of participant was also significant [$F(1,18) = 18.48$, $p<0.001$]; the client initiated a greater number of *Query-yn* Moves than the Travel Agent (means 9.09 vs 4.64). The interaction between communicative context and role of participant was also significant [$F(1,18) = 15.79$, $p<0.001$]. Further analysis by Simple Effects showed that the client initiated more *Query-yn* Moves in the VMC context than in the face-to-face context [$F(1,18) = 15.69$, $p<0.001$]. The proportion of *Query-yn* Moves initiated by the Travel Agent did not differ significantly with communicative context ($p<0.1$)

Align Initiating Moves.

The analysis showed that there was a significant main effect of context [$F(1,18) = 6.03$, $p<0.05$]. A greater number of *Aligns* were initiated in the VMC context than in face-to-face interactions (means being 3.68 vs 1.57 respectively). The main effect of role of participant was also significant [$F(1,18) = 13.68$, $p<0.01$]; the Travel Agent initiated a greater number of *Align* Moves than the client (means 2.17 vs 0.45 *Aligns* per 100 turns of dialogue). The interaction between communicative context and role of participant was also significant [$F(1,18) = 6.57$, $p<0.05$]. Further analysis by Simple Effects showed that the Travel Agent initiated more *Align* Moves in the VMC context than in the face-to-face context [$F(1,18) = 6.94$, $p<0.05$]. The proportion of *Align* Moves initiated by the client did not differ significantly with communicative context ($p<0.1$)

These analyses highlight the effect of communicative context and role of participants in the Travel Game. In the VMC dialogues the Travel Agent initiates a greater proportion of *Explain* and *Align* Moves, and the Client increases the use of *Query-yn* Moves in VMC compared to face-to-face context.

6.12 Discussion of Conversational Games Analysis

The results of the Conversational Games Analysis show that participants in the VMC and face-to-face contexts interacted and collaborated in different ways. This is reflected in the structure and content of the dialogues, which vary with communicative context. In the VMC dialogues the Travel Agent initiated more *Align* and *Explain* Moves, and the Client initiated more *yes-no questions* than occurred in the face-to-face dialogues. In other words, in the VMC travel Games the Travel Agent used proportionally more Initiating Moves to elicit feedback from the listener (*Align* Moves), or to offer information about the task and her activities *(Explain* Moves) to the client. At the same time, the client sought more information by asking a greater number of yes-no questions (*Query-yn*).

The hypothesis being tested in this study was that the impoverished visual channel provided by the VMC system would effect the structure and content of the VMC dialogues, as compared to face-to-face interactions. The results of the Conversational Games Analysis support the hypothesis that the structure and content of dialogues in the VMC context differed significantly from the dialogues in face-to-face interactions. We now need to ascertain whether these changes can be attributed to the impoverished nature of the visual channel provided in the VMC context, rather than alternative explanations.

There are several factors that could have contributed to changes in the way participants interacted in the VMC and face-to-face trials. The first factor concerns the technology used to support communication in the VMC context, which was novel for all the participants in this study. Secondly, several minor changes were made to the task to make it operable in the VMC context, such as small changes to the rules of the Travel Game and the number of airports displayed on the map of the USA. These factors could offer alternative explanations for the results obtained from the Conversational Games Analysis, and will be discussed in the next sections.

6.12.1 Novelty of VMC Context

All of the participants in VMC context had previous experience of working on computers, and had prior experience of electronic mail. However, they were all novice users of a VMC environment, and may have commented upon the context whilst taking part in the Travel Game. If the users of the VMC system had commented frequently upon the environment in which they were communicating, then this could account for some of the differences in the structure and content of the dialogues in face-to-face and VMC contexts. To ensure that the results from the Conversational Games Analysis were not simply due to references about the VMC context per se., the Game coded dialogues were surveyed, and the number of Initiating Moves in which participants commented on the communicative context was calculated. A total of 53 Initiating Moves, concerning context only, were found in the VMC dialogues. Most of the comments about the context were made by the Travel Agent (89%). The majority of these comments occurred in *Explain* Moves (68%), when the Travel Agent informed the client what she was doing. For example, that she was about to 'update the itinerary', which was displayed in a text window on the monitor screen. An example of such an incidence is shown in the Extract 1. In both extracts the Travel Agent (TA) is speaking.

Extract 1.

TA: Right. I'll just update your itinerary. You should see that in a minute.

Other references to the context occurred when the Travel Agent made an error whilst she was typing; as shown in the extract below.

Extract 2.

TA: Whoops, sorry I hit the wrong key on the keyboard.

As these extracts demonstrate, reference to the context were made when the Travel Agent informed the Client what she was doing, but these would not have occurred in the face-to-face context. For example, in the VMC context only the Travel Agent could update the itinerary, and the revised itinerary was then displayed on the Client's monitor. Mentioning that the itinerary was about to be updated occurred regularly in the VMC dialogues, but not in the face-to-face dialogues.

Explain Moves which commented on the context accounted for approximately 10% of all *Explain* Moves initiated by the Travel Agent. It is possible that these Moves could have accounted for some of the differences observed in the structure and content of VMC and face-to-face dialogues. These *Explain* Moves were excluded from the data obtained in the Conversational Games analysis, and the ANOVA on the revised standardised frequency of *Explain* Initiating Moves in the face-to-face and VMC contexts was re-run. The results of this analysis confirmed the original results; in the VMC dialogues the Travel Agent offers more un-elicited information to the client than occurred in face-to-face interactions [$F(1,18) = 6.17$, $p<0.05$]. The mean number of *Explain* Initiating Moves made by the Travel Agent in the face-to-face context was 17.01, compared to 25.00 in the VMC context.

The other seventeen references to the context were scattered across most of the other types of Initiating Moves initiated by the Travel Agent and the Client. These comments accounted for less than 2% of the overall number of Initiating Moves in the VMC dialogues, and are therefore unlikely to have had any significant effect upon the structure and content of VMC interact. The analysis on the revised standardised frequency of *Explain* Moves indicates that the changes in structure and content of the VMC dialogues, compared to the face-to-face interactions, was not simply due to participants commenting upon the VMC technology. Therefore, this alternative explanation for the differences in structure and content between VMC and face-to-face interactions can be dismissed.

6.12.2 Changes made to the Task in the VMC Context

The second alternative explanation, for the observed differences in the structure of the VMC and face-to-face interactions, arises from the minor changes that were made to the Travel Game for the VMC trials; these included small changes in the rules of the Game, and simplification of the flight information available to the Travel Agent. It is possible that these minor changes could account for the differences obtained in the Conversational Games Analysis. For example, reducing the number of airports available in the VMC context could have affected the amount of collaboration required to complete the task, and this might have changed the way VMC users interacted during the task. However, the changes made should have simplified the task, and reduced the amount of information being offered or sought by Agent and the Client in the VMC Travel Games.

To test this alternative theory, the dialogues from the face-to-face and VMC contexts were studied carefully. Any *Explain, Query-yn* or *Align* Initiating Moves which were involved with offering or requesting information about the rules, or flight details, were noted and their frequency calculated (the other types of Initiating Moves were not investigated, as they did not differ significantly with communicative context). *Align*

Moves were not associated with requests for information about rules or flight details, so this type of Initiating Move need not be considered further in respect to this alternative theory.

The question remains, were there more *Explain* and *Query-yn* Initiating Moves in the VMC dialogues because these participants in this context asked for more information about the rules of the Game, or about flight details? The percentage of *Explain* Moves which offered unsolicited information about flight details and the rules of the Travel Game accounted for 75% of *Explain* Moves initiated by the Travel Agent in the face-to-face context, but only 51% of *Explain* Moves by the Travel Agent in VMC dialogues. Looking at the client's contributions, the percentage of *Query-yn* Initiating Moves which sought information about the rules and flight details was 42% in face-to-face dialogues, but only 12% in VMC interactions.

These comparisons indicate that the changes in the rules and flight details, which were made to accommodate the Travel Game in the VMC context, could not account for the increased use of *Explain* Moves by the Travel Agent or the increased use of yes-no questions by the Client. Indeed, the comparisons show that the opposite effect occurred; in the face-to-face dialogues the Travel Agent tended to offer more information about the rules of the Game, and Client tended to ask more yes-questions about the flight details. We can conclude, therefore, that the changes made to the rules and simplification of the flight details, were not the cause of the increased use of *Explain* Moves by the Travel Agent, or the greater number of *Query-yn* Initiating Moves by the Client, in the VMC dialogues.

It would appear that neither the VMC context itself, nor the minor changes made to the task can account for the different ways that the participants interacted in the VMC and face-to-face contexts. So the differences observed in the face-to-face and VMC dialogues are probably a result of differences in the communicative contexts.

Since the quality of the spoken channel was excellent in both contexts, it would seem likely that the changes in structure and content of the interactions could be a result of the impoverished visual signals provided by the VMC system. The reasons for this suggestion will be explored in the following sections, by examining possible reasons for the increased frequency of the *Explain, Query-yn* and *Align* Initiating Moves in the VMC dialogues.

6.12.3 *Explain* Moves by the Travel Agent

Offering un-elicited information (*Explain* Moves) accounted for a greater proportion of a dialogue than any other type of Initiating Moves, in both communicative contexts. The majority of *Explain* Moves were initiated by the Travel Agent, with significantly more unelicited information being offered by the Travel Agent in the video-conferencing context. It has already been established that the increased use of *Explain* Moves in the VMC context was not due to the Agent offering more details about the task (either the rules or the flight details), so what other type of information was being offered during the Travel Game task?

As stated earlier, information about flights and the rules accounted for approximately 75% of all *Explain* Moves by the Travel Agent in the face-to-face context, but only 51% in the VMC context. The remaining *Explain* Moves in the face-to-face dialogues were chiefly concerned with offering information about the lack of flight connections between airports (approximately 15%). However, this type of information was offered less often by the Travel Agent in the VMC context than in face-to-face interactions; accounting for approximately 6% of the *Explain* Moves in VMC interactions. So what other type of information was the Travel Agent offering in the VMC interactions?

From careful examination of the VMC dialogues, it appears that the Travel Agent was offering information about what she was doing, or what she was currently attending to. This also occurred in the face-to-face dialogues, but very rarely (5 times in the whole face-to-face corpus). For example, the Travel Agent would tell the client that she was looking up the flight details, or inform the client that he/she had now landed at the next destination. The following extracts demonstrate this usage of *Explain* Initiating Moves in VMC dialogues. The relevant *Explain* Initiating Moves are emphasised in bold print in the following extracts. In these extracts 'TA' stand for the Travel Agent, 'C' for the client:

Extract 3. VMC dialogue

C: Can l get a connecting flight to Casper?

Move: Query-yn

TA: Right, I'll just check that for you

Move: Explain

Yes, there are flights between Dallas and Casper.

Move: Reply-y

Extract 4. VMC dialogue

C: Right, can l get a connection to Jacksonville?

Move: Query-yn

TA: I'll just check that for you

Move: Explain

Yes you can

Move: Reply-y

Do you want to go there?

Move: Query-yn

C: Yes.

Move: Reply-y

TA: (pause) Right, you're in Jacksonville

Move: Explain

C: Okay.

Move: Acknowledge

These types of explanations accounted for nearly 42% of *Explain* Moves initiated by the Travel Agent in the VMC context. This would indicate that, in the VMC dialogues, the Travel Agent spent a considerably amount of time and effort in keeping the Client informed of her activities, or their position in the task. In a typical VMC dialogue, offering this sort of information involved the Travel Agent in initiating an additional 27 EXPLAIN Games. Considering that the average VMC dialogue consisted of 104 Initiating Moves, offering information about her activities took up approximately 25% of the Travel Agent's time in VMC dialogues. Whether the proportion of Explain Moves initiated by the Travel Agent varied over time would be worth further investigation in the future, and would demonstrate how she adapted to the context as she gained experience.

Why did the Travel Agent offer this sort of information in the VMC context, but not during face-to-face interactions? One possible reason is that the video images provided by the VMC system were impoverished, when compared to the visual signals available in face-to-face communication. In the face-to-face Travel Games both participants were in the same room and could easily see what their partner was doing, and when they were both ready to proceed to the next stage of the Travel Game. However, in the VMC context participants had a very restricted view of each other, as the video image was small and showed just the head and shoulders of both participants. These factors made it difficult for VMC users to judge what the other person was doing, or when they were ready to proceed with the task. So in the VMC context the Travel Agent verbally offered the Client information that would have been available visually in face-to-face interactions.

If the video image had shown more of the communication environment, then the Travel Agent might not have felt the need to explicitly state what she was doing, or what was going on at her end of the VMC system. Another factor could be the quality of the video image. The size of the image was relatively small, had a low frame rate, and did not facilitate the use of mutual gaze. The quality of the video image, as noted by Heath and Luff (1991), would have made it difficult for participants to indicate in a non-verbal manner that they were ready to progress to the next stage of the task. The Travel Agent in the VMC context made use of the spoken channel to keep the Client informed of her activities, and to maintain the process of communication and collaboration.

6.12.4 Initiating Query-yn Moves by the Client

The use of yes-no questions (*Query-yn* Initiating Moves) was the only category of Initiating Moves used by the Client to vary between the VMC and face-to-face communicative contexts; the Client initiated more yes-no questions in the VMC dialogues. Although the difference is quite modest, in terms of the number of Moves initiated in each dialogue, the difference was significant. In an average VMC dialogue the client asked 6 times as many yes-no questions than in face-to-face interactions (see Table 6.8).

Examining the functions of *Query-yn* Moves initiated by the Client in the face-to-face and VMC dialogues, revealed that these questions were used to gain a wide range of information. For example, yes-no questions were asked to gain information about the rules of the Game, or the possibilities of changing the itinerary. However, the majority of the Initiating Moves occurred when the Client asked the Travel Agent if there were connecting flights between two Airports. These questions accounted for more than 78% of all of the *Query-yn* Moves initiated by the Client in VMC dialogues, but only 40% of *Query-yn* Moves initiated by the Client in face-to-face interactions.

Why does the Client ask more yes-no questions about connecting flights in the VMC context? To answer this question, the coded dialogues in each context were examined. The following extracts show incidences of yes-no questions (Moves are emphasised in bold print) in which the Client asks about the possibility of connecting flights between airports. The extracts are taken from both communicative contexts.

Extract 5. Face-to-face dialogue

GAME 1. QUERY-W

TA: and where would you like to go from Salt lake City

Move: Query-w

C: I will stay there for three days, and then I will fly out of State

Move: reply-wh

TA: Okay

Move: Acknowledge

GAME 2. QUERY-YN embedded

C: can I, is there, are there flights to Seattle from there?

Move: Query-yn

TA: I'll check that.

Move: Explain

Example 6. VMC dialogue

TA: Right, you've now arrived in Grand Rapids, in Michigan

Move : Explain (Ends previous Game)

GAME 1. QUERY-YN

C: Can l move on to Detroit?

Move: Query-yn

TA: I'll just check.

Move : Explain

No, you cant fly to Detroit from Grand Rapids.

Move: Reply-no

C: Right, okay.

Move: Acknowledge. Ends Game 1

GAME 2. QUERY-YN

Uhmm, can l fly to Great Falls

Move: Query-yn

TA: Ehmm lets see. No, I'm afraid you can't fly to Great Falls either.

Move: Reply-no. Ends Game 2

The examples show typical use of *Query-yn* Initiating Moves in the two communication contexts. In face-to-face interactions, the Client is often prompted by an open-ended question (*Query-w* Initiating Move) from the Travel Agent to say where he would like to go next, and many of the Client's *Query-yn* Moves concerning flight connections then occur within the context of an already initiated *Query-w* Move. These demonstrate cases of an Initiating Move being 'embedded' within an already existing Initiating Move. However, in the VMC context, the Travel Agent tends to state explicitly that the Client has reached his next destination (using an *Explain* Move), and then the Client asks if he can get a flight to somewhere else, thereby initiating a new Conversational Game. So in the VMC dialogues, yes-no questions by the Client frequently started off the next stage of the Travel Game; they were opening Initiating Moves made by the client, rather than being embedded within an already existing Initiating Move made by the Travel Agent.

Whilst it is difficult to tell if the variation in use *Query-yn* Initiating Moves is due to different people taking the role of the Travel Agent, or a result of the communicative contexts themselves, there does appear to be some evidence to suggest that the effect is linked to the communicative context. In the VMC dialogues, the Travel Agent is careful to explain what she is doing, using the spoken channel to keep in touch with the Client and informing him of her activities. This seems to be a result of the restricted nature of visual channel provided in the VMC system. The explicitness of this style of communication by the Travel Agent appears to determine how the Client responds during the task. As the Travel Agent tends to round off each set of Games, by offering explicit information about the Client's progress in the task, this puts the Client in the position of starting off the next stage of the Travel Game; which can be achieved most simply (and most explicitly) by asking the Agent if there are connections to a particular airport. Thus this behaviour does seem to result from the communicative context of VMC, but probably depends in part on the Travel Agent's response to the VMC condition.

6.12.5 *Align* Initiating Moves used by Travel Agent

During VMC dialogues the Travel Agent initiated significantly more *Align* Moves than during face-to-face interactions; that is, she elicited a greater amount of listener feedback than in the face-to-face interactions. The previous sections have already ruled-out the possibility that this increase was due to the minor changes made to the task for the VMC trials. Why then did the Travel Agent use proportionally more *Aligns* in the VMC context? Could this be due to the quality of the visual signals provided in this context?

The VMC system used in this study provided an impoverished video image. The quality of the video image was low, as the video image had a frame rate of only 4 - 5 Hz, the size of the image was small (11.5 by 9 cm), and the view that participants were given of each other was also greatly restricted in the VMC context when compared to face-to-face interactions. The quality of the video image in the VMC context could have reduced the possibility of using non-verbal communication (such as gestures or gaze) to elicit feedback, and establish mutual understanding. The audio link provided in this video-mediated context was of high quality; it afforded a full-duplex audio channel which made spoken interactions relatively easy. Therefore, in this VMC context it might have been easier (take less collaborative effort) for the Travel Agent to use the spoken channel to elicit listener feedback and establish mutual agreement, whereas in the face-to-face dialogues this could be achieved non-verbally due to the richness of the visual cues available when participants are co-present.

Apart from the reported increase in use of *Align* Games to verbally elicit listener feedback, is there any other evidence to support the claim that the Travel Agent in the VMC context used verbal alignments to replace the use of gaze? Searching through the coded dialogues revealed that Align Moves are used in a variety of ways. The following extracts demonstrate some of the ways in which Aligns were used by the Travel Agent to elicit feedback in the face-to-face and VMC contexts. In these extracts the following symbols are used: TA and C indicate the Travel Agent and Client respectively; a short pause is represented by three dots (...).

Extract 7. Examples of Align in Face-to-face context.

TA: Can you make a note of your decisions as we go along

**M Instruct*

TB: Uhmmm

**MReply-y*

TA: I think it said that in the instructions

****M Align***

TB: Sure, yeah it did.

**MReply-y.*

Extract 8. Example of Align in VMC context.

TA: Morning or afternoon?

*M Query-w

TB: Morning.

*M reply-w

TA: Okay so that's day 18? Or day 19?

****M Align***

TB: yeah, nineteen

*M reply-y

Extract 9. Example of Align in Face-to-face context

TA: So you are departing on day 1 at 7.55 am. And that flight arrives in Syracuse at 9.15 am.

**M Explain*

TA: ... Okay?

M Align

TB: Okay, I've got that.

**M: Acknowledge*

Extract 10. Example of Align in VMC context

TA: Its actually going to be day 28 before you can actually leave Arizona

**M Explain*

TA: Okay?

M Align

As these extracts show, Align Initiating Moves can be quite lengthy (as in extracts 7 and 8), but sometimes they can be initiated with just a single word (extracts 9 and 10). Careful examination of the dialogues showed that these shorter, one word, Aligns occurred more frequently in the VMC dialogues than the face-to-face interactions: 60% of the VMC Travel Agent's Aligns consisted of single words, such as 'okay', 'right', compared to 46% of Aligns in the face-to-face interactions. It is possible that these one word Aligns were being used more frequently in the VMC context to assist the process of grounding; short Align Moves were used instead of gaze to ascertain that the Client had understood the previous contribution. In the face-to-face context participants could see each other clearly, non-verbal forms of establishing mutual understanding were easily accessible, so the need to use verbal alignments was reduced in this context. Here again it would be interesting to see if the Travel Agent varied the proportional use of Align Moves as she became more adapt at working in the VMC context. Casual inspection of the raw data suggests that the Travel Agent initiated fewer ALIGN Games as she progressed from one client to another in the VMC context, but little difference occurred over time in the face-to-face interactions. This effect requires further, detailed analysis but the results would indicate one way in which the Travel Agent adapted to working in the VMC context.

6.13 Summary of Discussion of Study 3

Study 3 set out to examine the effects of low quality visual signals upon communication and collaboration, using an in-depth analysis of the process and content of communication. The discussion of these results suggests that in a VMC context where the visual channel provides restricted, impoverished visual signals, participants may attempt to achieve a greater amount of collaboration verbally; as the impoverished nature of the visual channel limits the use of non-verbal communication. This appears to be the common theme, linking the differences in proportional use of the *Align, Explain* and *Query-yn* Moves in the VMC context. The increased use of these Initiating Moves can be seen as a response to the restrictions imposed on people by VMC, especially when the visual channel is impoverished but the quality of the audio link is high.

Is there any support for these suggests from previous literature? The most relevant paper is by Doherty-Sneddon et al (1997), who examined the structure and content of dialogues from face-to-face and remote spoken contexts as well as several VMC contexts. The findings from Conversational Games Analysis of these contexts showed that people communicate in a more cautious manner when they use an audio-only context; they adopt what Shadbolt (1984, in Doherty-Sneddon 1997) calls a 'low risk' style of communication. This was apparent in the greater use of ALIGN and CHECK Games in remote spoken interactions, and an increased use of ALIGN Games in a remote computer-mediated (audio-conferencing) context. Doherty-Sneddon et al. concluded that participants interacting in a spoken only context use a greater number of verbal alignments, and a more cautious style of communication.

The results from Study 3 also show an increased use of verbal alignments. In this case, the effect may be due to the poor quality of the visual signals, rather then the total absence of video images. The quality of the visual signals may have been low enough to engender a more cautious style of communication than occurred in the face-to-face context. So the findings reported by Doherty-Sneddon et al. do support the view taken here, that the increased use of verbal alignments could have been due to the quality of the visual signals provided by the VMC system in Study 3. The other differences in structure and content of the VMC dialogues (increased use of *Explain* and *Query-yn* Initiating Moves) receive no support from previous literature. However, as suggested in the previous discussion these changes may demonstrate some of the different ways in which users adapted to working in a VMC context which affords low quality video images. Travel Agents requested a greater amount of listener feedback (*Align* Moves), and they spent proportionally more time informing the Client of their activities. The Client appears to have responded to this style of interaction by making greater use of simple yes-no questions.

6.14 Conclusions

This study has explored the effects of the quality of visual signals upon communication and collaboration. The findings show that VMC interactions are structured differently from face-to-face dialogues, and it has been suggested that this could be a result of the impoverished quality of the video image provided by the VMC system used in this study. When the visual signals are impoverished VMC, but the audio channel is of high quality, users appear to spend a greater proportion of their time (and proportionally more of their spoken output) in ensuring that their contributions are being interpreted correctly. At the same time, users appear to adopt a more explicit style of interaction, which not only offers more information about the task and about the activities of the participants, but also maintains contact between the collaborators when the quality of the video image is low.

These adaptations to the VMC environment are quite subtle, and may be difficult to observe when only the surface structure of the dialogues are examined; as the length of turns and number of turns may not be greatly affected. This could partly explain the mixed results, and lack of consensus, reported in the literature on the impact of VMC. The finer grain analysis employed in this study, which examines the pragmatic functions of utterances, has shown how effective communication is achieved in a VMC system which provides less than perfect visual signals. The study has illustrated the subtle ways in which people adapt to this communicative context.

Chapter 7. Thesis Conclusions

The aim of this thesis was to investigate the effects of two forms of computer mediated communication upon collaborative problem solving. The chosen contexts, CMC and VMC, are quite dissimilar. CMC is a text-based context, which affords a very restricted range of channels of communication. In contrast, VMC offers its users a richer array of channels, both verbal and non-verbal; VMC more closely resembles the communicative context we are most familiar with, face-to-face communication.

The literature reviewed in chapter 2, and at the beginning of chapters 5 and 6, showed that the effects of mediated communication systems upon interpersonal interactions have attracted a moderate amount of research. Previous studies which explored the impact of CMC and VMC have, however, tended to concentrate either on task outcome or variations in the interactional structure of communication in these contexts. Few researchers have attempted an in-depth analysis of the process and content of mediated communication. Another issue raised by the literature review was the apparent paucity of literature showing how people adapt to these contexts, very little research has examined how people adjust to new communicative contexts or the types of adjustments users make to their style of interaction as they gain experience of a novel context. With these points in mind, this thesis set out to explore the relationship of task outcome and process of communication in mediated contexts, how these factors might effect what people say (the content of the interactions) and how novice users adapt to computer-mediated contexts.

7.1 Key Results from the Three Studies

Study 1 set out to explore the effects of CMC upon effective communication, and how novice users adjusted to this restrictive written context. The results showed that task performance of participants in the CMC context was initially much poorer than the outcome from audio-only interactions. As the CMC users gained experience of the context they communicated more effectively, and task performance increased to a level comparable to the audio-only participants; indeed, CMC users completed later trials of the Map Task just as accurately as participants who interacted face-to-face.

Analysis of the process of communication, and a detailed examination of the content and structure of the interactions, showed that there were significant costs in establishing effective communication in CMC. Participants in this context adapted to the situation in several ways. Firstly, they adjusted to the increased production costs imposed by a written context by reducing the amount they wrote, whilst simultaneously sending longer messages. In other words, they sent fewer messages but attempted to accomplish more 'conversational goals' in each message than normally occurs in spoken interactions. Secondly, CMC users made appropriate adaptations to the way in which they achieved mutual understanding. They wrote clearer, more precise instructions and messages, which required very little clarification. The effort required to produce these concise messages was balanced by the fact that the pace of interaction in CMC was much slower, providing these users with more opportunity to plan and edit messages before they were transmitted.

In **Study 2** the effect of restraining access to the audio channel in VMC was investigated, by exploring the impact of two different audio channels commonly used in VMC systems. One VMC system provided full duplex audio signals ('open channel'), the other afforded half-duplex signals using a 'click to speak' method for controlling use of the audio channel. These VMC systems were compared with a face-to-face context. Since participants in this study could speak to each other it was anticipated that there would be little effect of context upon task performance. However, users of the one of the VMC systems (the 'open channel' VMC context) performed relatively poorly in the Map Task, an un-expected finding as most research concerning the use of VMC has not reported a loss of task performance in this context (see for example, Sellen (1992;1995), Anderson et al. (1997), Doherty-Sneddon et al. 1997).

Examination of the structure of the dialogues in all three contexts revealed several interesting features which could explain the differences in task performance. Dialogues from the open channel VMC and face-to-face contexts were very similar, with one crucial difference; a significantly greater number of episodes of overlapping speech occurred in the open channel VMC context. This extensive rate of interrupting could have been due to problems in timing turn-taking in the VMC context, as the quality of the video signals was not high (5 frames per second). Encountering problems in turn-taking could explain why these participants performed less well during the Map Task, which requires close collaboration between participants if they are to reproduce accurately drawn routes. In contrast, dialogues from the 'click to speak' VMC context were significantly longer, and contained very few interruptions. Episodes of overlapping speech often go unnoticed in face-to-face conversations, but in the 'click to speak' VMC context interruptions were very obvious and disrupted the process of communication. Users of the CTS system appeared to avoid interrupting each other, and they expended a considerable amount of collaborative effort on ensuring that the process of communication went smoothly.

Overall, it appears that users of the 'click to speak' VMC context adopted a more cautious style of interaction, they said much more to each other and ensured that they did not disrupt the process of communication. They achieved an effective style of communication by making some adaptations to the way in which they communicated. This finding confirms the results reported by Cohen (1982), who found that users of PicturePhone Meeting Service (PMS) regulated their interactions in a more orderly and polite manner than occurred in face-to-face interactions. Users of the open channel VMC system appear to have made few concessions to the novel context; they tried to interact as if they were collaborating in face-to-face. This could have been an effect of this VMC context, which gave the appearance of affording an effortless means of communication, similar to face-to-face interactions. The technology used in this VMC system appears to have provided these users with an enhanced feeling of 'presence', which Lombard and Ditton (1997) define as "the perceptual illusion of nonmeditation" (p. 8). This illusion occurs when people fail to perceive or make allowance for the mediated nature of a communication context, so they attempt to interact as if the mediating technology was not present. The results of Study 2 suggest that users of the open channel VMC system seem to have been taken-in by the feelings of 'presence' provided by the system, and therefore failed to adapt appropriately to this novel context.

Attention was shifted away from the audio channel towards the visual channel in **Study 3**. In this study the VMC system provided high quality, full duplex audio signals, but low quality video images. Dialogues from this VMC context were compared to face-to-face interactions whilst participants took part in a simulated service encounter. The focus of this study was an in-depth analysis of the content and structure of the VMC dialogues, to see if this could illuminate subtle differences in the process of communication attributable to the impoverished nature of visual signals afforded by this VMC system. The results from this analysis showed that the process and content of communication did vary with context, and this was probably related to small changes in the way that participants interacted in the two contexts.

The results from the Conversational Games Analysis suggest that users of the VMC system made greater use of the verbal channel of communication to compensate for the low quality of the visual signals. A finding also reported by Doherty-Sneddon et al. (1997). The observed differences between face-to-face and VMC interactions imply that participants in the VMC context were not confident that they could rely on the visual channel to convey some of the information normally available in face-to-face interactions. The Travel Agent therefore verbally informed the Client of her activities, and requested feedback on the Client's state of comprehension and readiness to progress with the task; both of these tasks could have been accomplished by participants just looking at each other in face-to-face interactions. The subtle changes in interactional style highlighted by this analysis again illustrate the point that humans are capable of adjusting the way they communicate and collaborate to suit the communicative setting.

The participants in Study 3 appeared to adjust quickly to the VMC context. Whether they would have adjusted the way in which they communicated if they had gained more experience of the context has not been explored in this Study, and there is very little literature on this topic to date. Studies by Rudman and Dykstra-Erickson (1994) and Dykstra-Erickson et al. (1995) have, however, shown that as people gain experience of VMC they become more adept at using the visual channel effectively. Rudman and Dykstra-Erickson found users became more sensitive to the information available in the video images over time; gradually interpreting facial expressions, gestures and turn-taking cues more accurately. Users also learnt how to control the video-camera so that they could change the camera field of view (changing their own appearance and using the camera to emphasis their own non-verbal behaviour), and becoming more adept at using the video to assist coordinate task activities. These changes in behaviour were, however, dependent on a range of factors (such as, team size, familiarity of participants, vested interest in task outcome, control over own image); adaptive effects were smaller if participants worked in pairs, or had little interest in the outcome of the task. So under some conditions, experienced users of VMC will be able to use the visual channel more effectively than novices.

The way that people adapt to VMC over time could be worth exploring, and one way forward would be to carry out further analysis of the changes in communicative style made by the Travel Agent as she adjusted to this novel context. Changes over time could be examined using Conversational Games Analysis, as this would highlight variations in structure and content of the Travel Agents contributions to the dialogues.

7.2 Contributions to the Literature on Computer-Mediated Communication

The series of studies discussed in this thesis can be seen as evolving along two different dimensions. Firstly, the effects of communicative context; moving from the restricted confines of a single channel, text-based means of communication (CMC) towards a multi-channel context (VMC). Secondly, the line of research made use of two different types of collaborative tasks, which represent examples of laboratory experimental tasks and more naturalistic, real-world types of tasks. The results have added to the literature in a variety of ways. For instance, providing insight into the types of adaptations made by users over time; offering evidence in support of the Collaborative Model of communication, and the framework of grounding constraints proposed by Clark and Brennan (1991); methodological issues. These points will be briefly outlined in the following sections.

7.2.1 Adaptation to Novel Contexts

All three studies have explored the kinds of adaptations made by users of novel communicative contexts. Users of VMC systems appear to adopt appropriate ways of collaborating very rapidly, but appear not to feel the need to do so if the restricted nature of the context is not made conspicuous. Fish, Kraut and Chalfont (1990) observed similar effects during informal interactions over the VideoWindow system. They suggest that people using this system did not make sufficient allowances for the context, they assumed that the system afforded full 'reciprocity'; that is, they thought that if they could see and hear who they were talking to then they themselves could be seen and heard. Participants did not always take account of the placement of microphones and cameras, and tried to interact with their remote colleagues when they were out of range of the recording equipment. The VideoWindow system appeared to provide a communicative context very similar to face-to-face interactions, it to provided an enhanced sense of 'presence' as the technology was non-intrusive.

The Studies presented in this thesis show that making appropriate adjustments to communicative style takes longer in communicative contexts that offer a very restricted range of channels of communication, such as CMC. Study 1 showed that when users are given more experience with this sort of communication technology they eventually find an appropriate way of collaborating. If the results had just been based upon the initial trials of the Study, then this CMC system would have been considered to have very dire effects upon collaborative problem-solving. As Hollingshead et al. (1993) and Isaacs and Tang (1994b) point out, it is important to study the effects of communicative systems over a period of time in order to learn how people adjust to the context.

7.2.2 Evidence in Support of a Collaborative Model of Communication and Process of Grounding

The results of all three studies produced evidence in support of the theoretical framework devised by Clark and Brennan (1991). This framework suggests that the process of grounding would change with context, and that participants can reduce the cost of grounding by trading-off one grounding cost against another. In the CMC interactions the increased costs of production were balanced by reducing the cost of making repairs, which require a greater amount of collaborative effort in this context. In Study 2 it was noted that users of the 'click to speak' VMC system appear to have taken great care not to interrupt each other, they seem to have traded-off the collaborative effort required to achieve smooth turn-taking (or 'speaker change costs') against the costs of making repairs. Participants in the 'open channel' VMC context appear to have made few concessions to the context. Finally, in Study 3 participants traded off an increase in collaborative effort required in using the impoverished visual channel by making greater use of the spoken channel. It required less collaborative effort to align with each other verbally than to attempt to do so visually.

Further research into the different ways that people adapt to a novel communicative environment is required, as users of less restrained contexts (such as VMC with high quality audio and video signals) could find more effective ways of collaborating as they gained in experience. Changes in the process and content of communication due to experience could be slight, and might not be detected if analysis is solely focused on the *structure* of the dialogues. These subtle changes might require a more in-depth analysis, as was the case in Study 3. Even though finer grained analysis is noticeably more tedious and time consuming to apply, the results from Study 1 and 3 indicate that they can illuminate phenomena that would otherwise be ignored.

7.2.3 Methodological Issues

The research carried out for this thesis has also added to two main methodological concerns. Firstly, the studies have expanded the use of Conversational Games analysis. Previously this form of analysis had been applied only to spoken interactions, but Study 1 shows that it can be reliably and profitably be used to examine text-based interactions. Conversational Games Analysis was successfully applied to explore the structure and content of dialogues from a different type of interactive task in Study 3. These dialogues were still task-oriented, but were produced under more naturalistic conditions; a simulated service encounter. Although this sort of in-depth analysis may be considered to be too time-consuming for inclusion in evaluations of computer-mediated communication systems, this thesis has shown that its application can provide insights into the way in which effective communication is achieved in a range of contexts. The findings from Conversational Games Analysis have shown mediated contexts do not just alter the structure of interactions, they can also change the content and process of communication.

Secondly, the two collaborative tasks used in this thesis can be seen to vary along a dimension of tasks. At one end of the dimension are laboratory based tasks, which can be tightly controlled under experimental conditions, thought it is difficult to generalise results to other, more realistic settings. Tasks at the other end of the dimension are those carried out in real work environments, under naturalistic conditions. The problem with naturalistic studies is that it is often difficult to make direct comparisons with the effects of other tasks and settings. If the full impact of computer-mediated communications is to be examined, then the dimension of tasks needs to be extended still further to include the type of tasks that people will encounter in the real world.

The Map Task, used in Study 1 and Study 2 is an example of a laboratory based task which allows the researchers to control the content of the dialogues and the order in which people encounter communication problems. The Travel Game is a more real-world version of this task, and captures many of the aspects of service encounters which VMC could be applied to in the near future. The Travel Game is, however, still a simulated task. Work has begun on a real world task which could be of use in the future. This task is based upon information retrieval from a large database, and requires collaboration between pairs of participants; a 'librarian' and a student. The students require information held in the data base to complete part of their course work, and the librarian can be a researcher who has been trained in the use of the data base. It is hoped that this new real-world task will increase our knowledge of the impact of various forms of mediated communication.

The line of research presented in this thesis appears to have taken a different route from other researchers; work which originated in experimental laboratory settings is being extended outwards towards more realistic settings. Other researchers have tried the opposite tactic. For instance, Olson et al. (1992; 1993; 1997) made extensive field trials of group task before bringing the task into the controlled environment of a laboratory setting. Both approaches will have benefits. For the type of research and methodological approach taken in this thesis, which makes use of time consuming fine grained analysis, it is probably more desirable to test the tasks and the system of analysis out in small controlled settings, establishing which phenomena require further investigation, and then transferring the tasks to a more naturalistic setting.

7.3 Applications of this Research

The findings reported in this work could provide valuable insights for future designs of computer-mediated communication systems, and in training people how to make effective use of these communication systems. Study I showed that effective communication can be achieved even in very restrictive contexts, but the effort required to collaborate could be reduced by incorporating additional features into CMC systems. Suggested ways of improving the design of CMC systems have already been raised by several researchers. For instance, McCarthy et al. (1993) suggest that a system of 'flags' could be made available to users, so that they could indicate whether they were composing a message, or waiting for a reply. It could also be helpful to provide users with a simple and quick way of indicating whether they had understood a message, providing them with a means of non-linguistic back channel markers.

Study 2 raised several issues relating to the design of VMC systems. Users of these systems appear to adapt to this form of communication very quickly, but only if they are in some way aware of the restrictions imposed by the context. Providing VMC systems that appear to closely resemble face-to-face contexts may make it harder for users to make appropriate adjustments to their interactional style. On the other hand, VMC systems which require greater collaboration between participants may appear to be more arduous to use, but have the benefit of encouraging users to adapt their communicative style in a sensible manner.

Study 3 illustrated that even contexts that provide both aural and visual channels of communication can present users with the problem of obtaining sufficient feedback from participants. This appears to be a key concern in collaborative interactions, in both spoken and written contexts. Feedback not only assures the speaker that they have been heard, it is also an essential part of the process of grounding; forming the acceptance phase of this process. When feedback is difficult to acquire, then the process of establishing mutual understanding will be more difficult.

7.4 Future Directions

The research carried out in this thesis has focused mostly upon the linguistic channel of communication. One way of extending the research would be to examine the patterns of non-verbal communication which occur in computer-mediated communication. An analysis of the patterns of gaze during VMC was conducted as part of this research. It was hoped that this analysis would provide an indication of how much use participants made of the various visual displays (video-images of partners, white boards displaying the map of the USA and the Itinerary). However, judging which of the visual display users were looking at proved difficult to judge, and the results were not reliable enough for inclusion in this thesis. This may have been a result of the quality of the video images and recordings that were available at the time, and the restricted view of the participant's faces shown on the recordings. It is hoped that these patterns of gaze could be explored in future, by making use of one of the recently available 'eye tracking' systems. This type of research would increase our knowledge concerning the types of visual presented information which assist task-based collaboration, and which assist in maintaining more social aspects of communication. Kraut and Fish (1997) suggest that provision of a visual channel, such as adding video to telephonic contexts (or 'videotelephony') increases the 'social' aspect of communication. This may be more important in interactional rather than transactional communication.

A second line of research is also proposed, which would move the analysis on to examine the effects of group discussions over VMC. As well as adding to the growing literature on the effects of VMC upon the effectiveness of group interactions, this research would also ascertain whether the system of analysis developed in this thesis could be applied to group interactions.

In summary, this thesis has shown that computer-mediated communications can effect several aspects of collaborative problem solving, and that the extent of these effects depends upon the range and quality of the channels of communication afforded in different contexts. Contexts that impose many restraints upon the processes of communication and collaboration require a greater amount of adaptive behaviour from users. Examining the ways in which people adapt to these contexts has provided some interesting insights into the flexibility of human communicators, who usually adapt in appropriate ways to different communicative settings. The capacity of humans to communicate effectively in such a wide array of contexts has been commented upon in previous literature, as the following quote demonstrates:

"It has become a cliché to refer to man as "the communicating animal." Of all his functions, that of building up systems of communication of infinite variety and purpose is one of the most characteristic. Of all living creatures he has the most complex and adaptable systems of language; he is the most widely observant of his physical environment and the most responsive in his adjustment to it." (Cherry, 1957, p. 29).

This thesis has illustrated some of the ways in which these adjustments to a variety of computer-mediated contexts are achieved.

References

Adrianson, L., and Hjelmquist, E. (1988). Users' experiences of COM - a computer-mediated communication system. Behaviour and Information Technology, 7 (1): 79-99.

Adrianson, L., and Hjelmquist, E. (1991). Group processes in face-to-face and computer-mediated communication. Behaviour and Information Technology, 10 (4): 281-296.

Anderson, A.H. (1995). Negotiating coherence in dialogue, in M.A. Gernsbacker and T. Givón (Eds.) Coherence in Spontaneous Text (pp.1-40). John Benjamins Publishing Company, Amsterdam.

Anderson, A.H., Bader, M., Bard, E., Boyle., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Sotillo, C., and Thompson, H. (1991b). The HCRC Map Task Corpus. Language and Speech, 34 (4): 351-360.

Anderson, A.H., Boyle, E., Newlands, A., Doherty-Sneddon, G., and Mullin, J. (1993). Collaboration in different communicative environments: an analysis of performance and dialogue. In A. Derycke (Ed.) Proceedings of Telepresence 93 (pp. 46-57). Ganymede, France: 46-57.

Anderson, A.H., and Boyle, E.A. (1994). Forms of introduction in dialogues: their discourse contexts and communicative consequences. Language and Cognitive Processes, 9 (1): 101-122.

Anderson, A.H., Clark, A., and Mullin, J. (1991a). Introducing information in dialogues: how young speakers refer and how young listeners respond. Journal of ChildLanguage, 18: 663-687.

Anderson, A.H., Clark, A., and Mullin, J. (1994). Interactive communication between children: learning how to make language work in dialogue. Journal of Child Language, 21: 439-463.

Anderson, A.H., Newlands, A., Mullin, J., Fleming, A.M., Doherty-Sneddon, G., and Van der Velden, J. (1996). Impact of video-mediated communication on simulated service encounters. Interacting With Computers, 8 (2): 193-206.

Anderson, A.H., O'Malley, C., Doherty-Sneddon, G., Langton, S., Newlands, A., Mullin, J., Fleming, A.M., and Van der Velden, J. (1997). The impact of VMC on collaborative problem solving: an analysis of task performance, communicative process, and user satisfaction. In K.E. Finn, A.J. Sellen and S.B. Wilbur (Eds.) Video-Mediated Communication (pp. 133-172). NJ: Lawrence Erlbaum Associates.

Anderson, A.H., Robertson, A., Kilborn, K., Beeke. S., and Dean, E. (1997). Dialogue despite difficulties: A study of communication abetween aphasics. In T. Givón (Ed.) Conversation (pp. 1-41). Amsterdam: John Benjamions.

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 and S.B. Wilbur (Eds.) Video-Mediated Communication (pp. 51-73). NJ: Lawrence Erlbaum Associates.

Argyle, M. (1969). Social Interaction. Methuen & Co Ltd. London.

Argyle, M. (1994). Bodily Communication. (2nd edition) Methuen & Co. Ltd.

Argyle, M., and Dean, J. (1965). Eye Contact, distance and affiliation. Sociometry, 28: 289-304.

Aronson, S.H. (1971). The sociology of the telephone. International Journal of Complementary Sociology, 12: 153-67.

Austin, J.L. 1962. How to do things with words. Oxford. Oxford University Press.

Barber, P.J., and Laws, J.V. (1994). Image quality and video communication. In R. Damper, W. Hall and J. Richards (Eds.) Proceedings of IEEE International Symposium on Multi-Media Technology and Their Future Applications (pp. 163-178). London. Pentech Press.

Beaman, K. (1984). Coordination and subordination revisited: syntactic complexity in spoken and written narrative discourse. In D. Tannen (Ed.) Coherence in Spoken and Written Discourse (pp. 45-80). Norwood, N.J: Ablex.

Beattie, G.W. (1980). The role of language production in the organisation of behaviour in face-to-face interactions. In B. Butterworth (Ed.) LanguageProduction. Academic Press

Beattie, G.W., and Barnard, P.J. (1979). The temporal structure of natural telephone conversations (directory enquiry calls). Linguistics, 17: 213-229.

Bennett, A. (1981) Interruptions and the interpretation of conversations. Discourse Processes, 4: 171-188

Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. Language, 62: 384-414.

Biber, D. (1988). Variation Across Speech and Writing. Cambridge. Cambridge University Press.

Biber, D. (1993). The multi-dimensional approach to linguistic analyses of genre variation: an overview of methodology and findings. Computers and the Humanities, 26: 331-345.

Bowers, J., and Churcher, J. (1989). Local and global structuring of computer mediated communication: developing linguistic perspectives on CSCW in Cosmos. Office: Technology and People, 4 (3): 197-227.

Boyle, E., Anderson, A.H., and Newlands, A. (1994). The effects of visibility on dialogue and performance in a cooperative problem-solving task, Language and Speech 37: 1-20.

Brown, G., Anderson, A.H., Yule, G. and Shillcock, R. (1984). Teaching Talk. Cambridge. Cambridge University Press.

Brown, G., and Yule, G. (1983). Discourse Analysis. Cambridge. Cambridge University Press.

Button, K., and Maggi, R. (1995). Videoconferencing and its implications for transport: an Anglo-Swiss perspective. Transport Reviews, 15 (1): 59-75.

Butterworth, B., Hine, R.R., and Brady, K.D. (1977). Speech and interaction in sound-only communication channels. Semiotica 20: 81-99.

Burton, D. (1980). Towards An analysis of Casual Conversation. Routledge and Kegan Paul.

Caldwell, B.S., Uang, S., and Taha, L.H. (1995). Appropriateness of communications media use in organisations: situation requirements and media characteristics. Behaviour and Information Technology, 14: 199-207.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., and Anderson, A.H. (1995) The Coding of Dialogue Structure in a Corpus. In J.A. Andernach, S.P.van de Buirgt, and G. F. van der Hoeven (Eds.) Proceedings of the Ninth Twente Workshop on Language Technology: Corpus-Based Approaches to Dialogue Modelling (pp. 25-34). Universiteit Twente, Enschede, Netherlands.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Newlands, A., Doherty-Sneddon, G., and Anderson, A.H. (1997). The Reliability of a Dialogue Structure Coding Scheme, Computational Linguistics, 23 (1):13-32.

Chafe, W.L. (1982). Integration and involvement in speaking, writing and oral literature. In D. Tannen (Ed.), Spoken and written: exploring orality and literacy, (pp. 35-54). Norwood, NJ: Ablex.

Chalfont, B.L., Fish, R.S., and Kraut, R.E. (1991). Expressive richness: a comparison of speech and text as media for revision. In S.P. Robertson, G.M. Olson and J.S. Olson (Eds.) Conference Proceedings of CHI'91, Human Factors in Computing Systems: Reaching Through Technology (pp. 21-26), New Orleans, Louisiana.

Chapanis, A. (1971). Preclude to 2001: explorations in human communications. American psychologist, 26: 949-961.

Chapanis, A. (1988). Interactive human communication. In I. Greif (Ed.) Computer-Supported Cooperative Work: A Book of Readings (pp. 127-142). Cambridge Mass: Morgan Kaufmann Publishers, Inc.

Chapanis, A., Ochsman, R.B., Parrish, R.N., and Weeks, G.D. (1972). Studies in interactive communications: I. The effects of four communication modes on the behaviour of teams during co-operative problem-solving. Human Factors, 14: 487-509.

Chapanis, A,. Parrish, R.N., Ochsman, R.B., and Weeks, G.D. (1977). Studies in interactive communication: II. The effects of four communication modes on the linguistic performance of teams during cooperative problem solving. Human Factors, 19 (2): 101-126.

Cherry, C. (1957). On Human Communication. MIT Press Cambridge, Mass.

Cherry, C. (1971). World Communication: Threat or Promise?. Wiley, London.

Clark, H.H. (1985). Language Use and Language Users. In G. Lindsay and E. Aronson (Eds.), Handbook of Social Psychology, Vol. II (pp. 179-231). Random House, New York.

Clark, H.H. (1992). Arenas of language use. University of Chicago Press. Chicago.

Clark, H.H. (1996). UsingLanguage. Cambridge University Press.

Clark, H.H., and Brennan, B.E. (1991). 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., and Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, B.L. Webber and I.A. Sag (Eds.) Elements of discourse understanding (pp. 10-63). Cambridge: Cambridge University Press.

Clark, H.H., and Schaefer, E.F. (1987). Collaborating on contributions to conversations. Language and Cognitive Processes, 2 (1): 19-41.

Clark, H.H., and Schaefer, E.F. (1989). Contributing to discourse. Cognitive Science, 13: 259-294.

Clark, H.H., and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. Cognition, 22: 1-39.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Epidemiology and Community Health, 32: 200-205

Cohen, K. M. (1982). Speaker interaction: Video teleconferencing versus face-to-face meeting. In L.A. Parter and C.H. Olgren (Eds.), Teleconferencing and Electronic Communications: Application, technologies and Human Factors (pp. 189-199). Madison, WI. University of Wisconsin Extension, Centre for Interactive Programs.

Cohen, P.R. (1984). The pragmatics of referring and the modality of communication. Computational Linguistics, 10: 97-146.

Connolly, J. H., and Pemberton, L. (Eds.). (1996). Linguistic Concepts and Methods in CSCS. Springer-Verlag, London.

Cook, M., and Lalljee, M. (1972). Verbal Substitutes for verbal signals in interaction. Semiotica, 6: 212-221.

Coulthard, M. (1985). An Introduction to Discourse Analysis, London: Longman.

Crystal, D. (1995). The Cambridge Encyclopaedia of Language. Cambridge University Press. Reprint.

Daft, R.L., and Lengel, R.H. (1984). Information richness: a new approach to managerial behavior and organization design. In L.L. Cummings and B. Staw (Eds.), Research in Organizational Behavoir (pp. 191-233). Greenwich. CT: JAI Press.

Daft, R.L., and Lengel, R.H. (1986). Organizational information requirements, media richness, and structural design. Management Science, 32: 554-571.

Davies, M.F. (1971). Co-operative problem-solving: a follow-up study. Report No. E/71252/ DVS, Post Office Long Range Intelligence Division, Cambridge UK.

Deuel N R. (1996). Our passionate response to virtual reality. In S.C. Herring (Ed.) Computer-Mediated Communication, Linguistic, Social and Cross-Cultural Perspectives (pp. 129-146). John Benjamines, Amsterdam.

Diener, E. (1979). Deindividuation, self-awareness, and disinhibition. Journal of Personality and Social Psychology, 37: 1177-1187.

Doherty-Sneddon, G. (1995). The Development of Conversational and Communication Skills. Unpublished PhD Thesis. University of Glasgow.

Doherty-Sneddon, G., and Kent, G. (1996). Visual Signals and the Communication Abilities of Children. Journal of Child Psychology, 8: 949-959.

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 (2): 1-21.

Dykstra-Erickson, E.A., Rudman, C., Hertz, R., Mithal, K., Schmidt, J., and Marshall, C. (1995). Supporting adaptation to multimedia desktop conferencing. Proceedings of the 15th International Symposium on Human Factors in Telecommunications (pp. 31-38).

Ekman, P., and Friesen, W.V. (1969). The repertoire of nonverbal behaviour: Categories, origins, usage, and coding. Semiotica, 1: 49-98.

Ellis, A., and Beattie, G.W. (1986). The Psychology of Language and Communication. Weidenfeld and Nicholson.

Exline, R.V., Thibaut, J., Brannon, C., and Gumpter, P. (1961). Visual interaction in relation to Machiavelianism and an unethical act. American Psychologist, 16: 396.

Fafchamps, D., Reynolds, D., and Kuchinsky, A. (1991). The dynamics of small group decision-making using electronic mail. In J.M. Bowers and S.D. Benford (Editors) Studies in Computer Supported Cooperative Work (pp. 211-224). Amsterdam: Elsvier Science.

Feldstein, S., and Welkowitz, J. (1978). A chronography of conversation: In defese of an objective approach. In A.W. Siegman & S. Feldstein (Eds.) Nonverbal behavior and communication. Hillsdale, NJ. Erlbaum.

Finn, K.E. (1997). Introduction: An Overview of Video-Mediated Communication Literature. In K.E. Finn, A.J. Sellen and S.B. Wilbur, (Eds.) Video-Mediated Communication (pp. 3-21). NJ: Lawrence Erlbaum Associates..

Fish, R.S., Kraut, R.E., and Chalfont, B.L. (1990). The VideoWindow system in informal communication. In Proceedings, ACM Conference on Computer-Supported Cooperative Work (pp. 1-11). Los Angeles, CA.

Frolich,D., Drew, P., and Monk, A. (1994). Management of repair in Human-Computer interaction. Human Computer Interaction, 9: 385--425.

Furnace, J., Hamilton, N.M. Helms, P., and Duguid, K. (1996). Medical teaching at a peripheral city by videoconferencing. Medical Education, 30 (3):.215-220

Fussell, S.R., and Benimoff, N.I. (1995). Social and cognitive processes in interpersonal communication: implications for advanced telecommunications technologies. Human Factors, 37 (2): 228-250.

Gale. S. (1990) Human aspects of interactive multimedia communication. Interacting with Computers, 2 (2):0 175-189.

Galegher, J., and Kraut, R.E. (1990). Computer-mediated communication for intellectual teamwork: a field experiment in group writing. In Proceedings, ACM Conference on Computer-Supported Cooperative Work, (pp. 65-78). Los Angeles, CA.

Galegher, J., and Kraut, R.E. (1994). Computer-mediated communication for intellectual teamwork: an experiment in group writing. Information Systems Research, 5 (2): 110-138.

Gallupe, R.B., and McKeen, J.D. (1990). Enhancing computer-mediated communication: an experimental investigation into the use of a group decision support system for face-to-face versus remote meetings. Information and Management, 18: 1-13.

Garrod, S. and Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. Cognition, 27: 181-218.

Greenberg, S. (1991). Computer-supported cooperative work and groupware: an introduction to the special issue. International Journal of Man-Machine Studies, 34: 133-141.

Grice, H.P. (1957). Meaning. Philosophical Review, 64: 377-388.

Grice, H.P. (1968). Utterer's meaning, sentence-meaning, and word meaning. Foundations of Language, 4: 225-242.

Gross, D.., Allen, J., and Traum, D. (1992). The trains 91 dialogues. TRAINS Technical Note 92-1, Computer Science Dept., University of Rochester.

Hall, E.T. (1966). The Hidden dimension. London, Bodley Head.

Halliday, M.A.K. (1989). Spoken and Written Language. Oxford University Press. Second edition.

Harasim, L., (Ed.). (1990). On-Line Education: Perspectives on a New Medium. New York: Prager/Greenwood.

Heath, C., and Luff, P. (1991). Disembodied conduct: communication through video in a multi-media office environment. Proceedings of the ACM Conference on CHI 1991, New Orleans: Louisiana.

Heath, C., and Luff, P. (1992). Media Space and Communicative Asymmetries: Preliminary Observations of Video-Mediated Interaction. Human-Computer Interaction, 7: 315-346.

Hiltz, S.R. (1986). Online communities: A Case study of the Office of the Future. New York: Ablex.

Hiltz, S.R., and Johnson, K. (1990). User Satisfaction with computer mediated communication systems. ManagementScience, 36 (6): 739-751.

Hiltz, S.R., and Turoff, M. (1993). The Network Nation: Human Communication via Computer. Revised edition. MIT Press, Massachusetts.

Hiltz, S.R., Turoff, M., and Johnson, K. (1986). Experiments in group decision making, 1: Communications process and outcome in face-to-face vs. Computerised conferences. Human Communications Research, 13 (2): 225-252.

Hiltz, S.R., Turoff, M., and Johnson, K. (1989). Experiments in group decision making, 3: Disinhibition, deindividuation, and group process in pen name and real name computer conferences. Decision Support Systems, 5: 217-232.

Hollingshead, A.B., and McGrath, J.E. (1995). Computer-assisted groups: a critical review of the empirical research. In R.A. Guzzo and E. Salas (Eds.) Team Effectiveness and Decision Making in Organizations (pp. 46-78). San Francisco: Jossey-Bass.

Hollingshead, A.B., McGrath, J.E., and O'Connor, K.M. (1993). Group task performance and communication technology: a longitudinal study of computer-mediated versus face-to-face work groups. Small Group Research, 24 (3): 307-333.

Houghton, G. (1986). The Production of Language in Dialogue: A Computational Model, Ph.D. thesis, University of Sussex.

Houghton, G., and Isard, S. (1987) Why to speak, what to say and how to say it: Modelling language production in discourse, in P. Morris (Ed.) Modelling Cognition (pp. 249-267). John Wiley.

Howell. D.C. (1997). Statistical Methods for Psychology. Wadsworth Publishing Co. USA.

Isaacs, E.A., and Clark, H.H. (1987). References in conversation between experts and novices. Journal of Experimental Psychology: General, 116: 26-37.

Isaacs, E.A., and Tang, J.C. (1994). What video can and cannot do for collaboration: a case study. Multimedia Systems, 2: 63-73.

Isaacs, E. A., and Tang, J.C. (1994b). Do Video-Mediated communication systems support real conversations? In S. Finn, A. Sellen and S. Wibur (Eds.) CSCW'94 Position Papers from Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implication, (pp. 5-10) Chapel Hill, North Carolina.

Isaacs, E.A., and Tang, J.C. (1997). Studying Video-Based Collaboration in Context: From Small Workgroups to Large Organizatons. In K.E. Finn, A.J. Sellen and S.B. Wilbur (Eds.) Video-Mediated Communication (pp. 173-197). NJ: Lawrence Erlbaum Associates.

Jaffe, J., and Feldstein, S. (1970). Rhythms of Dialogue. Academic Press, New York.

Jefferson, G. (1972). Side Sequences. In D. Sudnow (Ed.) Studies in Social Interaction (pp. 294-338). New York: Free Press..

Kendon, A. (1967). Some functions of gaze direction in social interaction. Acta Psycholgica, 26: 1-47.

Kiesler, S., Siegel, J., and McGuire, T.W. (1984). Social psychological aspects of computer-mediated communication. American Psychologist, 39 (10): 1123-1134.

Kiesler, S., and Sproull, L. (1992). Group decision making and communication technology. Organizational Behavior and Human Decision Processes, 52: 96-123.

Kiesler, S., Zubrow, D., Moses, A.M., and Geller, V. (1985). Affect in computer-mediated communication: an experiment in synchronous terminal-to-terminal discussion. Human-Computer Interaction, 1: 77-104.

Knowles, L.E., and Dillon C.L. (1996). Education and the continuing professional education of Architects: the role of satellite videoconferencing and learning. Journal of Architectural and Planning Research, 13 (2): 140-151.

Kowtko, J. (1992). On the intonation of mono- and di-syllabic words within the discourse framework of conversational games, in Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics (pp. 282-284). Delaware.

Kowtko, J. (1995) The function of intonation in spontaneous and read dialogue, in Proceedings of the XIIIth International Congress of Phonetic Sciences Vol. 2 (pp. 286-289). Stockholm, Sweden.

Kowtko, J. (1997). The Function of Intonation in Task-Oriented Dialogue. Research Paper HCRC/RP-84, Human Communications Research Centre, University of Edinburgh.

Kowtko, J., Isard, S., and Doherty-Sneddon, G. (1992) Conversational Games within Dialogue Research Paper HCRC/RP-31, Human Communications Research Centre, University of Edinburgh.

Kruass, R.M., and Fussell, S.R. (1990). Mutual knowledge and communicative effectiveness. In J. Galegher, R.E. Kraut, and C. Egido (Eds.) Intellectual teamwork: Social and technical bases of collaborative work. NJ: Lawrence Erlbaum Associates.

Krauss, R.M., and Fussell, S.R. (1996). Social psychology models of interpersonal communication. In E.T. Higgins and A.W. Kruglanski (Eds.) Social Psychology, Handbook of Basic Principles (pp. 655-701). New York: Guildford Press.

Krauss, R.M., and Glucksberg, S. (1969). The development of communication: Competence as a function of age. Child Development, 40: 255-256.

Krauss, R.M., and Glucksberg, S. (1977). Social and nonsocial speech. Scientific American, 236: 100-105.

Krippendorff, K. (1980). Content Analysis: An Introduction to Its Methodology (pp. 129-154). Sage: Newbury Park, CA.

Kraut, R.E., and Fish, R.S. (1997). Prospects for videotelephony. In K.E. Finn, A.J. Sellen and S.B. Wilbur (Eds.) Video-Mediated Communication (pp. 541-565). NJ: Lawrence Erlbaum Associates.

Lea, M. (1991). Rationalist assumptions in cross-media comparisons of computer-mediated communication. Behaviour and Information Technology, 10, (2): 153-172.

Lea, M., O'Shea, T., Fung, P., and Spears, R (1992). "Flaming" in computer-mediated communication: Observations, explanations, implication. In M. Lea (Ed.) Contexts of computer mediated communications, (pp. 89-112). London: Harvester-Wheatesheaf.

Lea, M. and Spears, R. (1991). Computer-mediated communication, de-individuation and group decision-making. International Journal Man-Machine Studies, 34: 283-301.

Lea, M. and Spears, R. (1992). Paralanguage and social perception in computer-mediated communication. Journal of Computing, 2 (3&4): 321-341.

Levinson, S.C. (1983). Pragmatics. Cambridge Textbooks in Linguistics series. Cambridge University Press. 1991 reprint.

Lewis, D.K. (1969). Convention: A philosophical study. Cambridge MA: Harvard University Press.

Lombard, M., and Ditton, T. (1997). At the heart of it all: the concept of presence. Journal of Computer Mediated Communication, 3 (2): 1-37.

Maclay, H., and Osgood, C.E. (1959). Hesitation phenomena in spontaneous English speech. Word, 15: 19-44.

Mantovani, G. (1996). New Communication Environments. Taylor and Francis Ltd.

Mason, R., and Kaye, A. (1989). Mindweave: Communication, Computers and Distance Education. Oxford: Pergamon Press.

Masoodian, M., Apperley, M., and Frederickson, F. (1995). Video support for shared work-space interaction: an empirical study. Interacting With Computers, 7 (3): 237-253.

McCarthy, J.C., Miles, V.C., Monk, A.F., Harrison, M.D., Dix, A.J., and Wright, P.C. (1993). Text-based on-line conferencing: a conceptual and empirical analysis using a minimal prototype. Human-Computer Interaction, 8: 147-183.

McCarthy, J.C., and Monk, A.F. (1994a). Channels, conversation, cooperation and relevance: all you wanted to know about communication but were afraid to ask. Cooperative Computing, 1: 35-60.

McCarthy, J.C., and Monk, A.F. (1994b). Measuring the quality of computer mediated communication. Behaviour and Information Technology, 13: 311-319.

McCarthy, J.C., Wright, P.C., and Monk, A.F. (1992). Coherence in text-based electronic conferencing: coupling text and context. Journal of Language and Social Psychology, 11: 267-276.

Macdonald, A.M. (Ed.). (1972). Chambers Twentieth Century Dictionary. Edinburgh. W and R Chambers. New Edition.

McGrath, J.E. (1984). Groups: Interaction and Performance. Englewood Cliffs, NJ: Prentice-Hall.

McGrath, J.E. (1990). Time matters in groups. In J. Galegher, R.E. Kraut, and E. Egido (Eds.) Intellectual Teamwork: Social and technological foundations of cooperative work (pp. 23-61). NJ: Erlbaum.

McGrath, J.E. (1993). Introduction: the JEMCO workshop - description of a longitudinal study. Small Group Research, 24 (3): 285-305.

McGrath, J.E., Arrow, H., Gruenfeld, D.H., Hollingshead, A.B., and O'Connor, K.M. (1993). Groups, tasks and technology: the effects of experience and change. Small Group Research, 24 (3): 406-420.

McGuire, T. W., Kiesler, S., and Siegel, J. (1987). Group and computer-mediated discussion effects in risk decision making. Journal of Personality and Social Psychology, 52 (5): 917-930.

McKinlay, A., Procter, R., Masting, O., Woodburn, R., and Arnott, J. (1994). Studies of turn-taking in computer-mediated communications. Interacting With Computers, 6 (2): 151-171.

McLeod, P.L. (1992). An assessment of the experimental literature on electronic support of group work: results of a meta-analysis. Human-Computer Interaction, 7: 257-280.

Miller, J. (1993). Spoken and written language: language acquisition and literacy (pp. 99-141). In R. Scholes (Ed.) Literacy and Linguistic Analysis. Lawrence Erlbaum, Hillsdale NJ.

Miller, J., and Weinert, R. (1995). The functions of LIKE in discourse, Journal of Pragmatics, 23: 365-393.

Monk, A.F., McCarthy, J.C., Watts, L.A., and Daly-Jones, O. (1996). Measures of process. In P. Thomas (Ed.) CSCS Requirements and Evaluation (pp. 125-139). Berlin:SpringerVerlag.

Monk, A.F., and Watts. L. (1995). A poor quality video link affects speech but not gaze. In Proceedings of CHI'95 Conference on Human Factors in Computing Systems. Conference Companion (pp. 274-275). ACM. Denver, Colorado, USA.

Morley, I.E., and 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., and Stephenson, G.M. (1970). Formality in experimental negotiations: a validation study. British Journal of Psychology, 61: 383.

Newlands, A., Anderson, A.H., and Mullin, J. (1996). Dialog structure and cooperatiave task performance in two CSCW environments. In J.H. Connolly and L Pemberton (Eds.) Linguistic Concepts and Methods in CSCW (pp. 41-60). Springer-Verlag.London.

Ochs. E. (1979). Planned and unplanned discourse, in T. Givón Syntax and Semantics Vol. 12: Discourse and Syntax (pp. 51-80). New York: Academic Press.

Ochsman, R.B., and Chapanis, A. (1974). The effects of ten communication modes on the behaviour of teams during cooperative problem solving. International Journal of Man-Machine Systems, 6: 579-6619.

O'Conaill, B. (1996). Private communication.

O'Conaill, B., Whittaker, S., and Wilbur, S. (1993). Conversations over video conferences: an evaluation of the spoken aspects of video-mediated communication. Human Computer Interaction, 8: 389-428.

O'Conaill, B., and Whittaker, S. (1997). Characterizing, predicting, and measuring video-mediated communication: a conversational approach. In K.E. Finn, A.J. Sellen and S.B. Wilbur (Eds.) Video-Mediated Communication (pp. 107-131). NJ: Lawrence Erlbaum Associates.

O'Connell, D. C., Kowal, S., and Kaltenbacher, E. (1990). Turn-taking: a critical analysis of the research tradition. Journal of Psycholinguistic Research, 19, (6): 345-373).

Olson, G. M., Olson, J.S., Carter, M., and Storrøsten, M. (1992). Small group design meetings: an analysis of collaboration. Human-Computer Interaction, 7: 347-374.

Olson, J.S., Olson, G.M., Storrøsten, M., and Carter, M. (1993). Groupwork close up: a comparison of the group design process with and without a simple group editor. ACM Transactions on Information Systems, 11 (4): 321-348.

Olson, J.S., Olson, G.M., and Meader, D.K. (1994). What mix of video and audio is useful for remote real-time work. In S. Finn, A. Sellen and S. Wibur (Eds.) CSCW'94 Position Papers from Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implication (pp. 33-41) Chapel Hill, North Carolina.

Olson, J.S., Olson, G.M., and Meader, D. (1997). Face-to-face group work compared to remote group work with and without video. In K.E. Finn, A.J. Sellen and S.B. Wilbur (Eds.) Video-Mediated Communication (pp. 157-172). NJ: Lawrence Erlbaum Associates.

O'Malley, C., Langton, S., Anderson, A.H., Doherty-Sneddon, G., and Bruce, V. (1996). Comparison of face-to-face and video-mediated interaction. Interacting with Computers, 8 (2): 177-192.

O'Malley, C. (1997). Private communication.

Oviatt, S.L., and Cohen, P.R. (1991). The contributing influence of speech and interaction on human discourse patterns. In J.W. Sullivan and S.W. Tyler (Eds.) Intelligent User Interfaces (pp. 69-83). ACM: New York.

Parker, C., and Joyner, S. (1995). Facilitating simultaneous engineering in the automotive industry via high speed networks, control passing and face-to-face video: results from the car project. International Journal of Industrial Ergonomics, 16: 427-440.

Perry, T. S. (1992). E-mail at work. IEEE Spectrum, 29 (10): 24-28.

Power, R. (1979). The Organisation of Purposeful Dialogues, Linguistics, 17: 107-152.

Reinking, D. (1992). Differences between electronic and printed texts: An agenda for research. Journal of Educational Multimedia and Hypermedia, 1: 11-24.

Reiss, J., Cameon, R., Matthews, D., and Shenkman, E. (1996). Enhancing the role public health nurses play in serving children with special health needs: an interactive videoconference on Public Law 99-457 part H. Public Health Nursing, 13 (5): 345-352.

Robinson, A. (1993). Communication and contact by videoconferencing in the divided society of Northern Ireland. Journal of Educational Television 19 (3): 127-137.

Rommetveit. R. (1974). On Message Structure. A framework for the study of languageandcommunication. John Wiley & Sons. London.

Rudman, C., and Dykstra-Erickson, E. (1994). Adaptation over time is required for effective use of video for group work. In S. Finn, A. Sellen and S. Wibur (Eds.) CSCW'94 Position Papers from Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implication (pp. 49-53) Chapel Hill, North Carolina.

Rutter, D.R. (1987). Communicating by Telephone. Oxford:Pergamon.

Rutter, D.R., and Stephenson, G.M. (1977). The role of visual communication in synchronising conversation. European Journal of Social Psychology, 7: 29-37.

Rutter, D.R., and Stephenson, G.M. (1979). The role of visual communication in social interaction. Current Anthropology, 20: 124-125.

Rutter, D.R., Stephenson, G.M., Ayling, K., and White, P.A. (1978). The timing of looks in dyadic conversation. British Journal of Social Psychology, 7: 29-37.

Sacks, H. and Schegloff, E.A. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In G. Psathas (Ed.) Everyday language: Studies in ethnomethodology (pp. 15-21). New York: Irvington Publishers.

Sacks, H., Schegloff, E.A, and Jefferson, G.A. (1974). A simplest systematics for the organization of turn-taking in conversation. Language 50: 696-735.

Schegloff E.A. (1972). Notes on conversational practise: formulating place. In D. Sudnow (Ed.) Studies in Social Interaction, (pp. 75-119). New York: Free Press.

Schegolff, E.A., Jefferson, G.A., and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. Language 53: 361-382.
Schegloff, E.A., and Sacks, H. (1973). Opening up closings. Semiotica 8: 289-327.

Schiano, D. J., Colston, H. L., and Pederson, E.R. (1994). Psychological perspectives on the "Presence Paradox". In S. Finn, A. Sellen and S. Wibur (Eds.) CSCW'94 Position Papers from Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implication (pp. 55-57) Chapel Hill, North Carolina.

Schiffer, S. (1972). Meaning. Oxford. Clarendon Press.

Schiffrin. D. (1994). Approaches to Discourse. Oxford: Blackwell.

Schober, M.F., and Clark, H.H. (1989). Understanding by addressees and overhearers. Cognitive Psychology, 21: 211-232.

Schrum. L. (1993). Social interaction through online writing. In R. Mason (Ed.) Computer Conferencing, the Last Word, (pp. 171-196). Victoria BA: Beach Home Publications

Searle, J.R. (1969). Speech Acts: An Essay in the Philosophy of Language, Cambridge University Press.

Sellen, A.J. (1992). Speech patterns in video mediated conversations. In P. Bauersfeld, J. Bennett and G. Lynch (Eds.) Proceedings of CHI'92, Conference on

Human Factors in Computing Systems (pp. 49-59). Monterey, CA. New York: ACM.

Sellen, A.J. (1995). Remote conversations: The effects of mediating talk with technology. Human Computer Interaction, 10: 401-444.

Severinson Eklundh, K. (1986). Dialogue processes in computer-mediated communication. A study of letters in the COM system. Linköping Studies in Arts andSciences, 6.

Short, J. (1974). Effects of medium of communication on experimental negotiation. Human Relations, 27: 225-243.

Short, J., Williams, E., and Christie, B. (1976). The Social Psychology of Telecommunications John Wiley and Sons. London.

Siegel, S., and Castellan, N.J. (1988). Nonparametric Statistics for the Behavioral Sciences. McGraw Hill. Second Edition.

Siegel, J., Dubrovsky, V., Kiesler, S., and McGuire, T.W. (1986). Organizational Behavior and Human Decision Processes, 37: 157-187.

Sinclair, J.H. and Coulthard, R.M. (1975). Towards an Analysis of Discourse: The English Used by Teachers and Pupils, Oxford: Oxford University Press.

Smith, R., O'Shea, T., O'Malley, C., Scanlon, E., and Taylor, J. (1991). Preliminary experiments with a distributed, multi-media, problem-solving environment. In J. Bowers and S. Benford (Eds.) Studies in Computer-Supported Co-operative Work: Theory, practice and design, (pp. 31-48). Amsterdam: Eslevier.

Smolensky, M.W., Carmody, M.A., and Halcomb, C.G. (1990). The influence of task type, group structure and extraversion on uninhibited speech in computer-mediated communication. Computers in Human Behavior, 6: 261-272.

Spears, R., Lea, M., and Lee, S. (1990). De-individuation and group polarization in computer-mediated communication. British Journal of Social Psychology, 29: 121-134.

Sperber,D., and Wilson, D. (1986). Relevance: Communication and Cognition. Cambridge, MA: Harvard University Press.

Spritzer, M. (1986). Writing style in computer conferences. IEE Transactions on Professional communications. Vol. PC 29 (1): 19-22.

Sproull, L., and Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. ManagementScience, 32: 1492-1512.

Sproull, L., and Kiesler, S. (1991). Connections: New Ways of Working in the Networked Organization. MIT: USA.

Straus, S.G. (1996). Getting a Clue: The effects of communication media and information distribution on participation and performance in computer-mediated and face-to-face groups. Small Group Research, 27: 115-142.

Straus, S.G., and McGrath, J.E. (1994). Does the medium matter: The interaction of task and technology on group performance and member reactions. Journal of Applied Psychology, 79: 87-97.

Stephenson, G.M., Ayling, K., and Rutter, D.R. (1976). The role of visual communication in social exchange. British Journal of Social and Clinical Psychology, 15: 113-120.

Sumner. M. (1988). The impact of electronic mail on managerial and organizational communication. In P.B. Allen (Ed.) Proceedings of the IFIP Conference on Office Information System (pp. 96-109). New York: ACM.

Svartvik, J., and Quirk, R. (1980). A Corpus of English Conversation. Lund, Sweden: Gleerup.

Tang, J.C., and Isaacs, E. (1993). Why do users like video?. Computer Supported Cooperative Work (CSCW), 1: 163-196.

Taylor, T.J., and Cameron, D. (1987). Analysing Conversation: Rule and Units in the Structure of Talk. Oxford. Pergamon Press.

Taylor, P.A., Shimodaira, H., Isard, S., King, S., and Kowtko, J. (1996). Using prosodic information to constrain language models for spoken dialogue. In Proceedings of the International Conference on Spoken Language Processing. Philadelphia.

Thompson, L.A., and Ogden, W.C. (1995). Visible speech improves human language understanding: implications for speech processing systems. Integration of Natural Language and Vision Processing, 9 (4-5): 105-116.

Traum, D.R., and Hinkleman, E. H. (1992). Conversation Acts in task-oriented spoken dialogue. Computational Intelligence, 8 (3): 575-599.

Traxler, M. (1993). Collaboration and Coherent Written Communication. Unpublished Doctorial Dissertation. University of Oregon, USA.

Trevino, L.K., Daft, R.L., and Lengel, R.H. (1990). Understanding managers' media choices: asymbolic interactionist perspective. in J Fulk and C.W. Steinfeild (Eds.) Organizations and Communication Technology. Sage, Newbury Park, CA.

Tröster, A.I., Paolo, A.M., Glatt, S.L., Hubble, J.P., and Koller, W.C. (1995). Interactive video conferencing in the provision of neuropsychological services to rural areas. Journal of Community Psychology, 23: 85-88.

Walther, J.B., and Burgoon, J.K. (1992). Relational communication in computer-mediated interaction. Human Communication Research, 19 (1): 50-88.

Walther, J.B. (1993). Impression development in computer-mediated interaction. Western Journal of Communication, 57: 381-398.

Walther, J.B., Anderson, J.F., and Park, D.W. (1994). Interpersonal effects in computer-mediated interaction; A meta-analysis of social and antisocial communication. CommunicationResearch, 21 (4): 460-487.

Walther, J. B. (1995). Relational aspects of computer-mediated communication: experimental observations over time. OrganizationScience, 6 (2): 186-203.

Watson, A., and Sassse, M.A. (1996). Evaluating audio and video quality in low-cost multimedia conferencing systems. Interacting With Computers, 8 (3): 255-275.

Watts, L.A., and Monk, A. (1996). Inter-personal awareness and synchronization: assessing the value of communication technologies. International Journal of Human-Computer Studies, 44: 849-873.

Watzlawick, P., Beavin, J.H., and Jackson, D.D. (1967). *Pragmatics of Human Communication.* New York: W. W. Norton.

Weisband, S. (1992). Group discussion and first advocacy effects in computer-mediated and face-to-face decision making groups. Organizational Behavior and Human Decision Processes, 53: 352-380.

Werry. C.C. (1996). Linguistic and interactional features of Internet Relay Chat. In S.C. Herring (Ed.) Computer-Mediated Communication, Linguistic, Social and Cross-Cultural Perspectives (pp. 47-64). John Benjamines, Amsterdam.

Wichman, H. (1970). Effects of isolation and communication on cooperation in a two-person game. Journal of Personality and Social Psychology, 16: 114-120.

Wiener, M., Devoe, S., Bubinow, S., and Geller, J. (1972). Nonverbal behavior and nonverbal communication. Psychological Review, 79: 185-214.

Wilbur, S. (1994). Characteristics of video-mediated conversation in studio and desktop settings. In S. Finn, A. Sellen and S. Wibur (Eds.) CSCW'94 Position Papers from Workshop on Video-Mediated Communication: Testing, Evaluation, and Design Implication (pp. 73-74) Chapel Hill, North Carolina

Wilkes-Gibbs, D. (1995). Coherence in collaboration: Some examples from conversation. In M. A. Gernsbacher and T. Givón (Eds.) Coherence in Spontaneous Text, (pp. 239-267). John Benjamin, Amsterdam.

Williams. E. (1975). Medium or message: communications medium as a determinant of interpersonal evaluation. Sociometry 38 (1): 119-130.

Williams. E. (1977). Experimental comparisons of face-to-face and mediated communication: A review. Psychological Bulletin, 84 (5): 963-976.

Winograd, T., and Flores, F. (1986). Understanding computers and cognition: a new foundation for design. Norwood, New Jersey: Ablex.

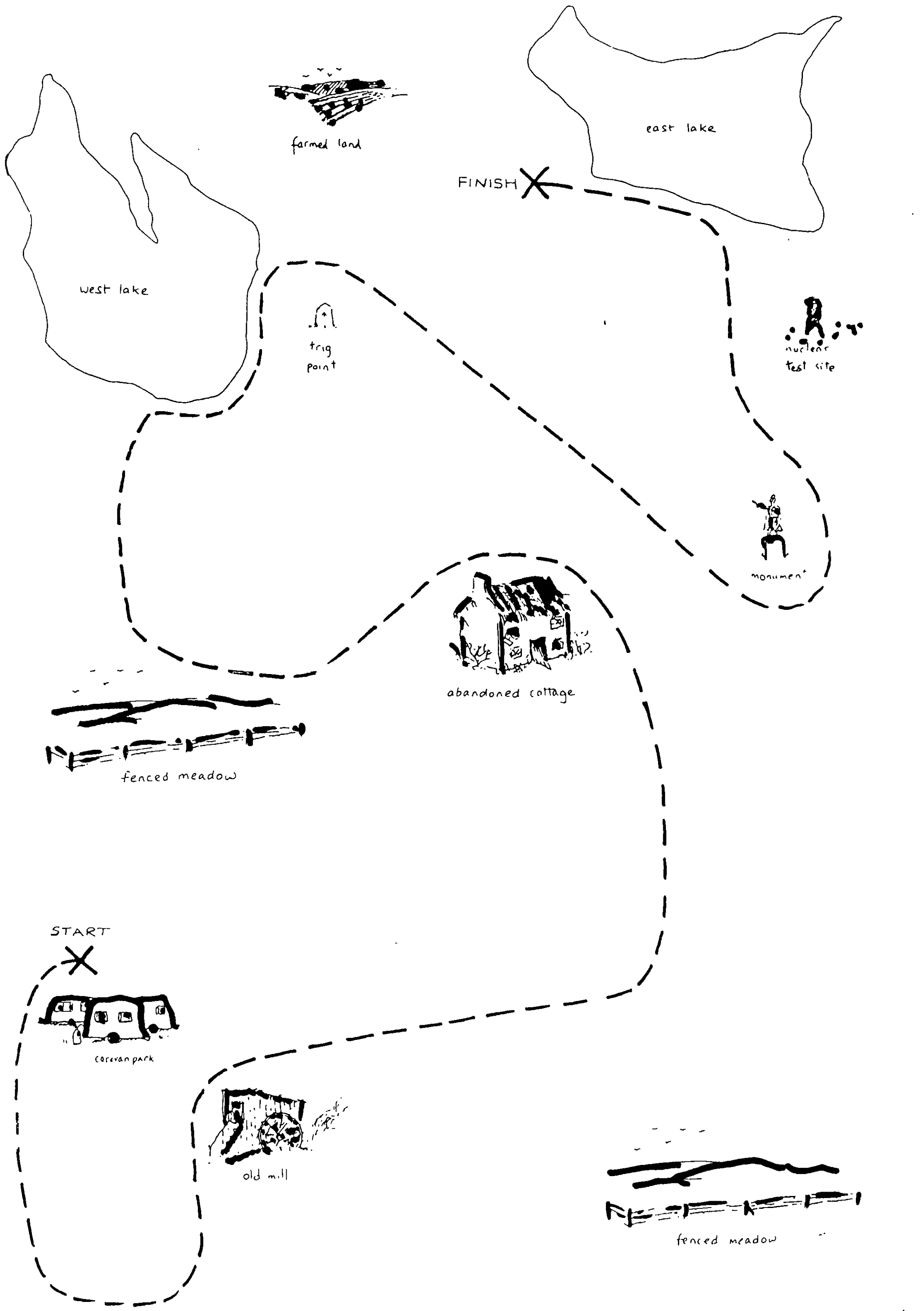
Wood, J.B., and Parham, I.A. (1996). Distance learning - videoconferences as vehicles for faculty-development in gerontology geriatrics. Educational Gerontology, 22 (1): 105-115.

Yates, S. (1993). Speech, writing and computer conferencing: An analysis. In R. Mason (Ed.), Computer Conferencing: the last word (pp. 37-56). Victoria BA: Beach Home Publications.

Yates, S. (1996). Oral and written linguistic aspects of computer conferencing. In S.C. Herring (Ed.) Computer-Mediated Communication, Linguistic, Social and Cross-Cultural Perspectives (pp. 29-46). John Benjamines, Amsterdam.

Appendix A: Study 1. Examples of HCRC Maps used for CMC context

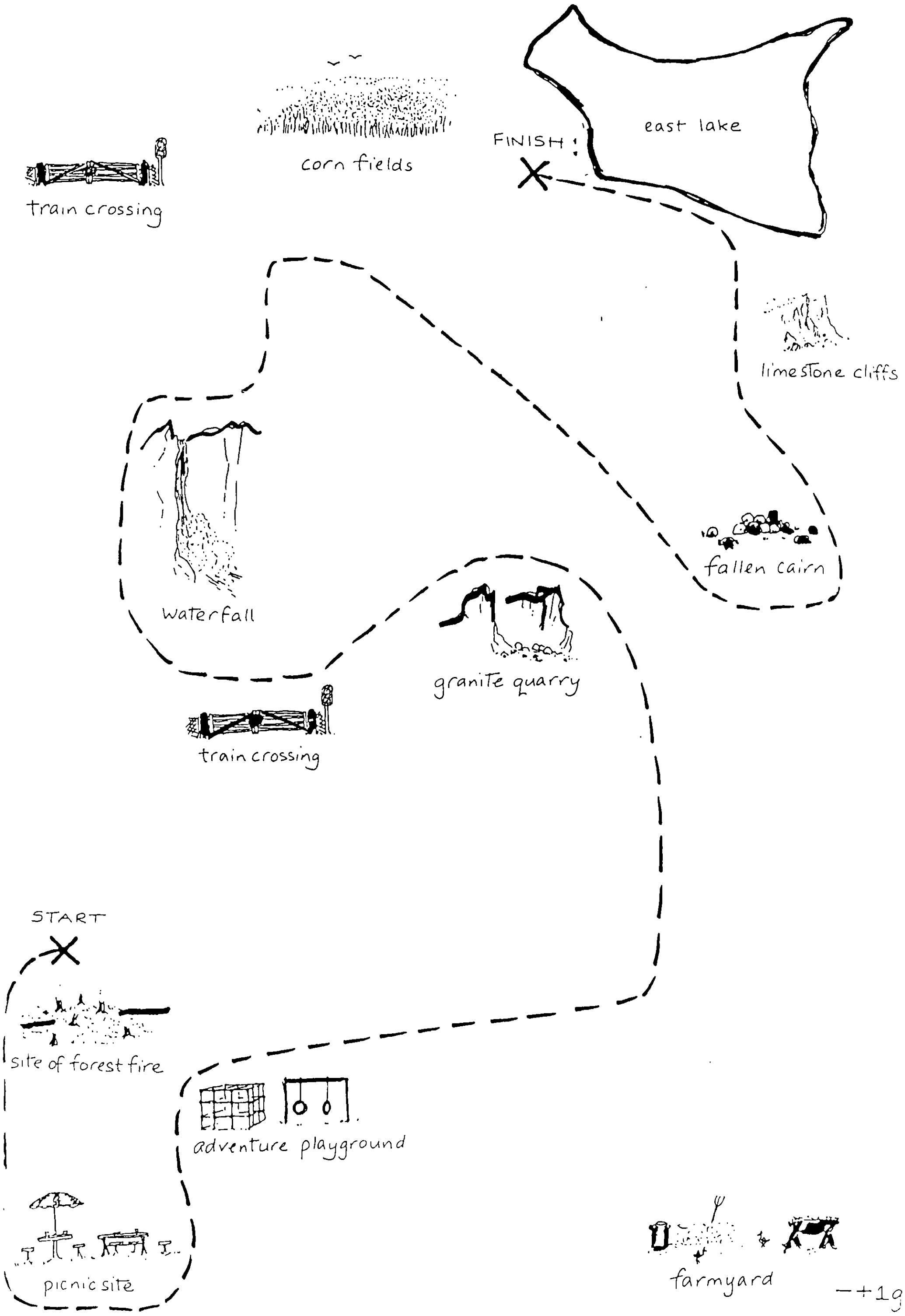
A.1 Map ++1 Instruction Givers' Map



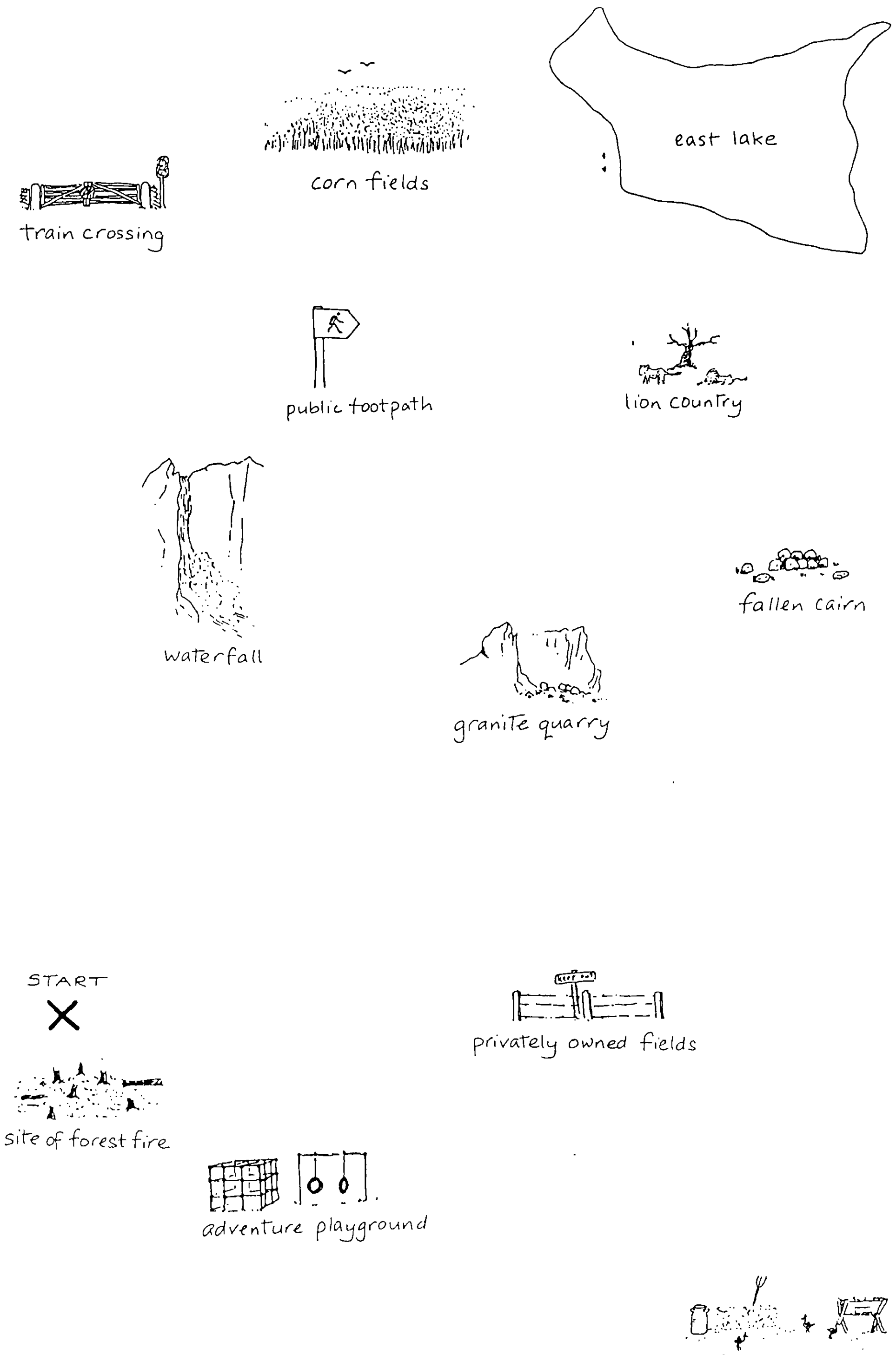
++1g

A.2 Map ++1 Instruction Followers' Map

A.3 Map -+1 Instruction Givers' Map

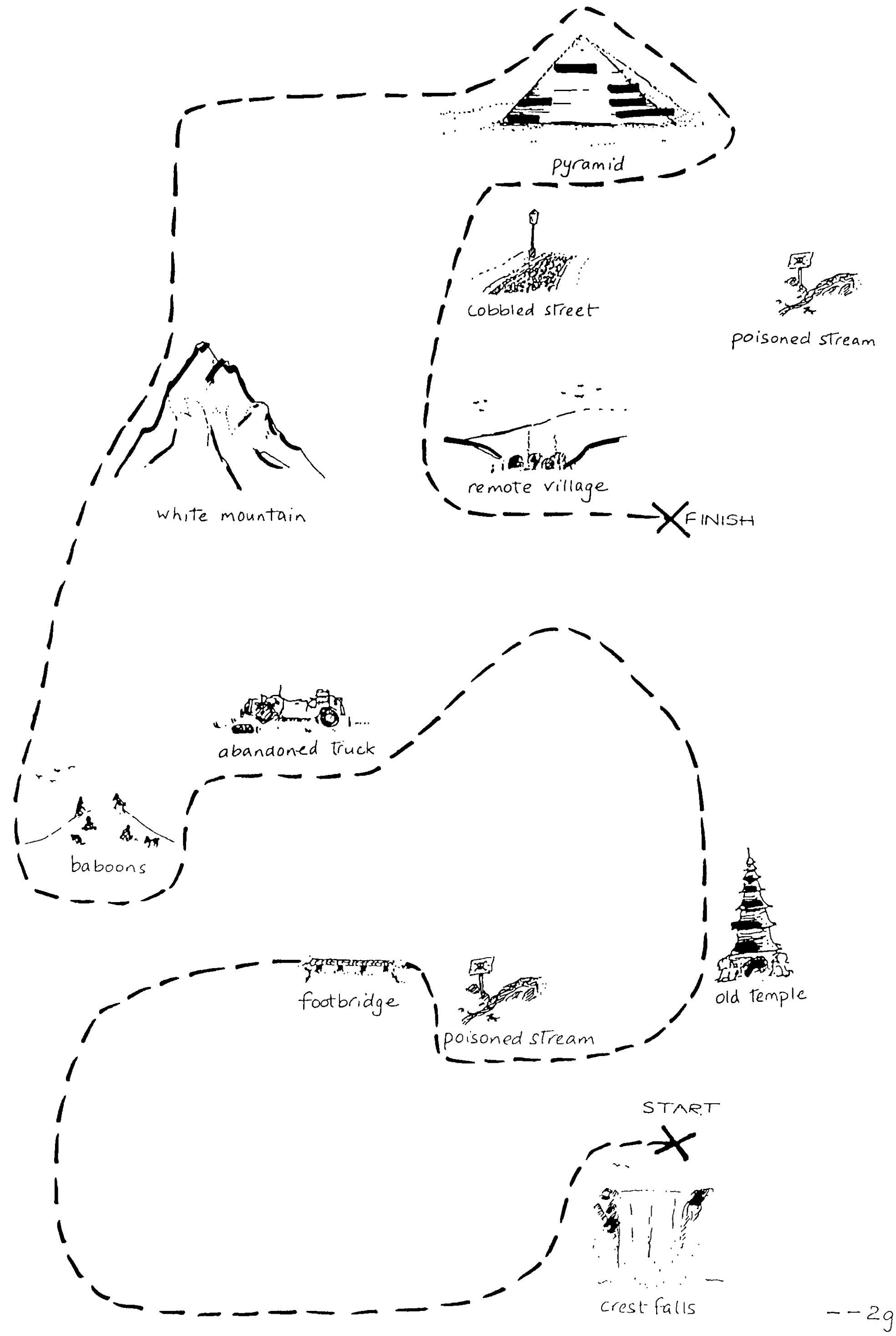


A.4 Map -+ Instruction Followers' Map



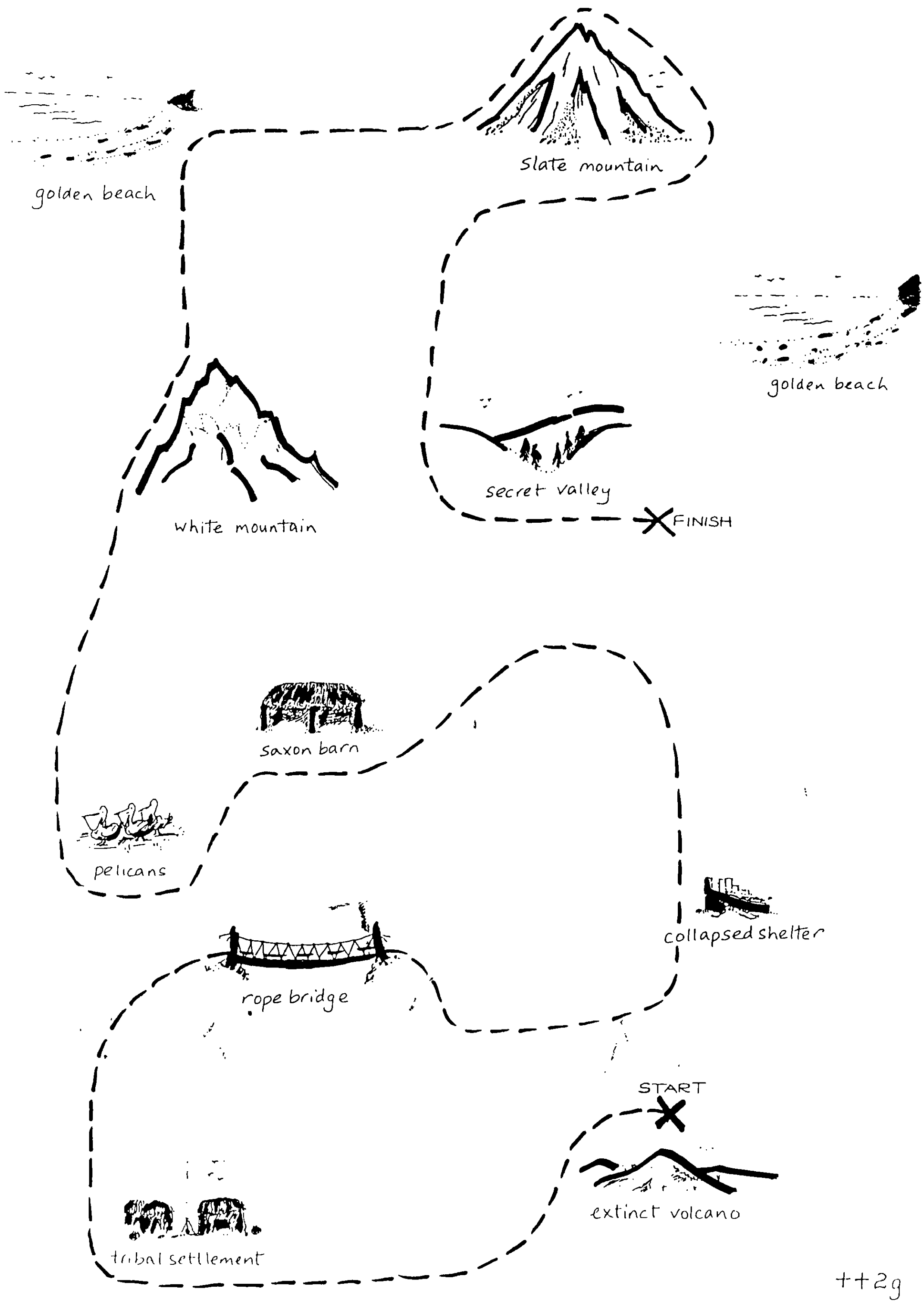
—+1f

A.5 Map --2 Instruction Givers' Map

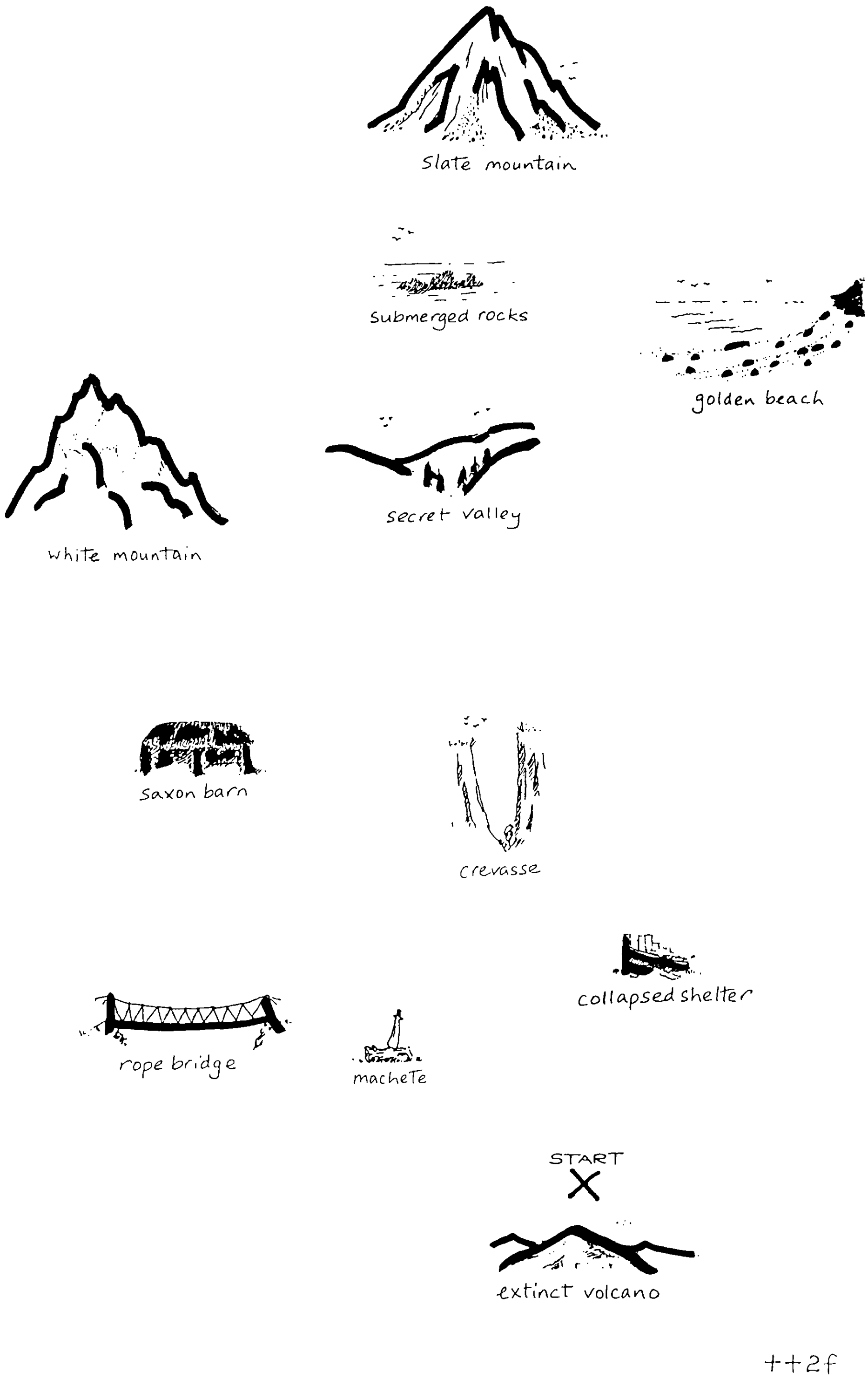


A.6 Map --2 Instruction Followers' Map

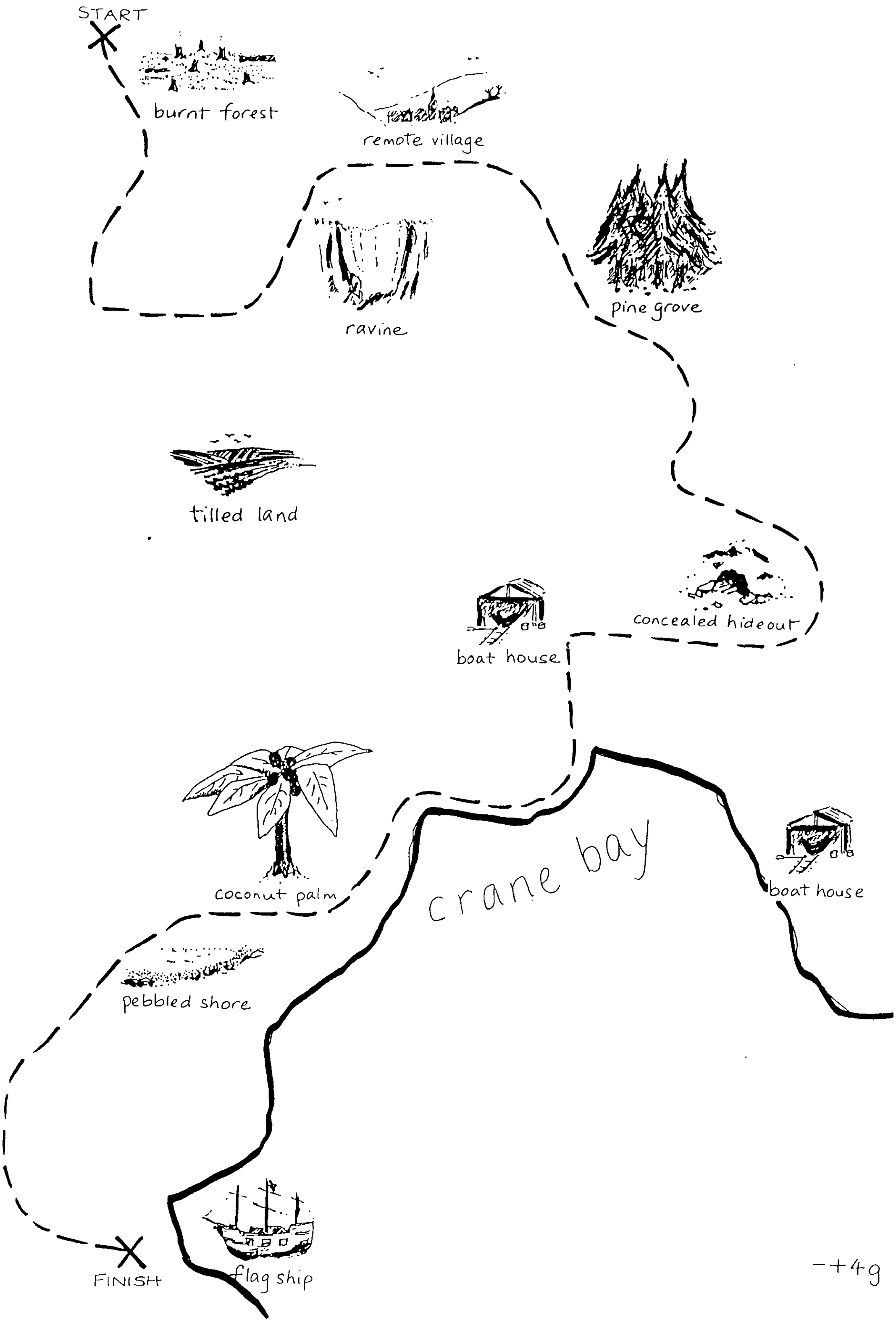
A.7 Map ++2 Instruction Givers' Map



A.8 Map ++2 Instruction Followers' Map

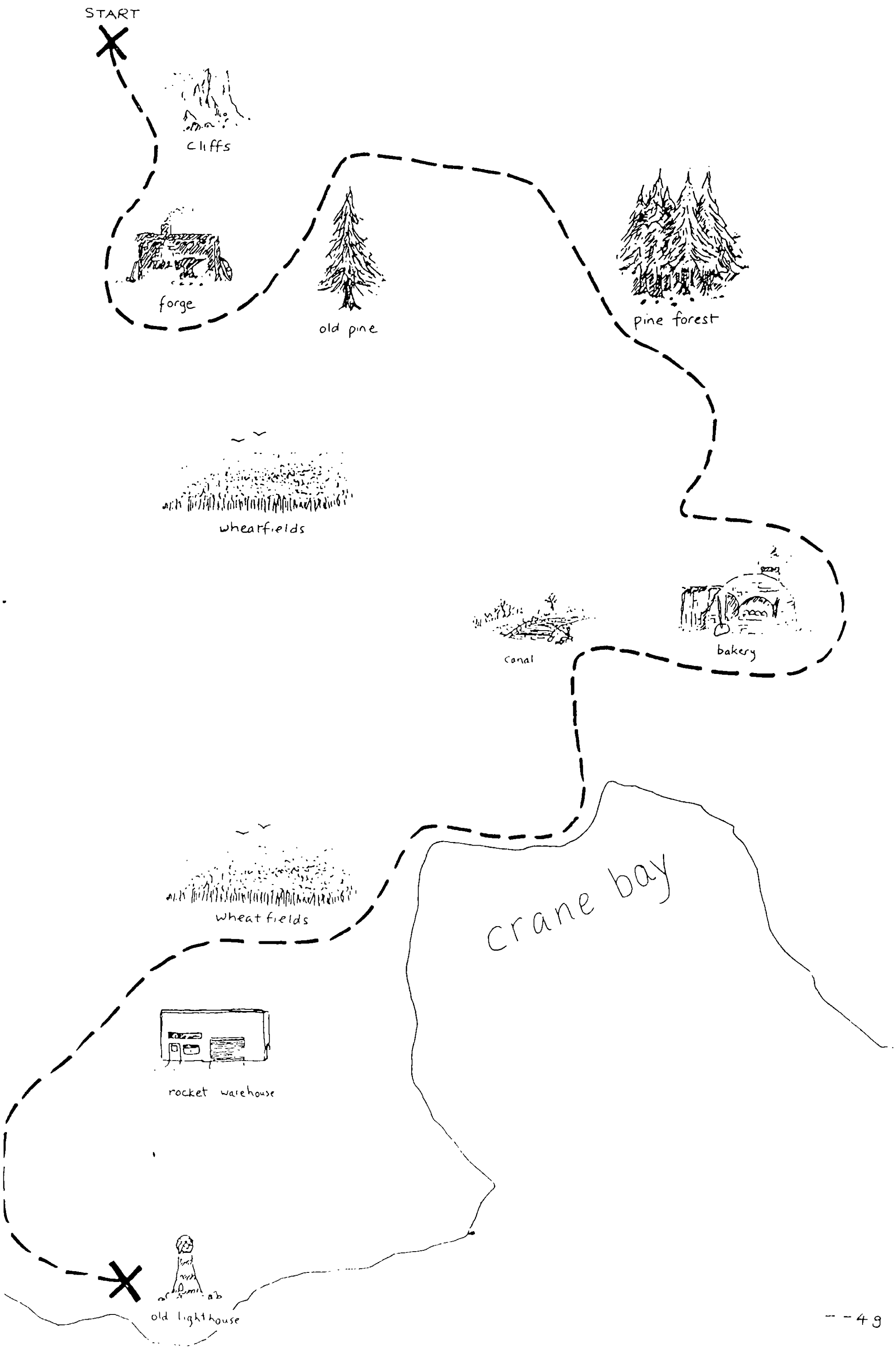


A.9 Map -+4 Instruction Givers' Map



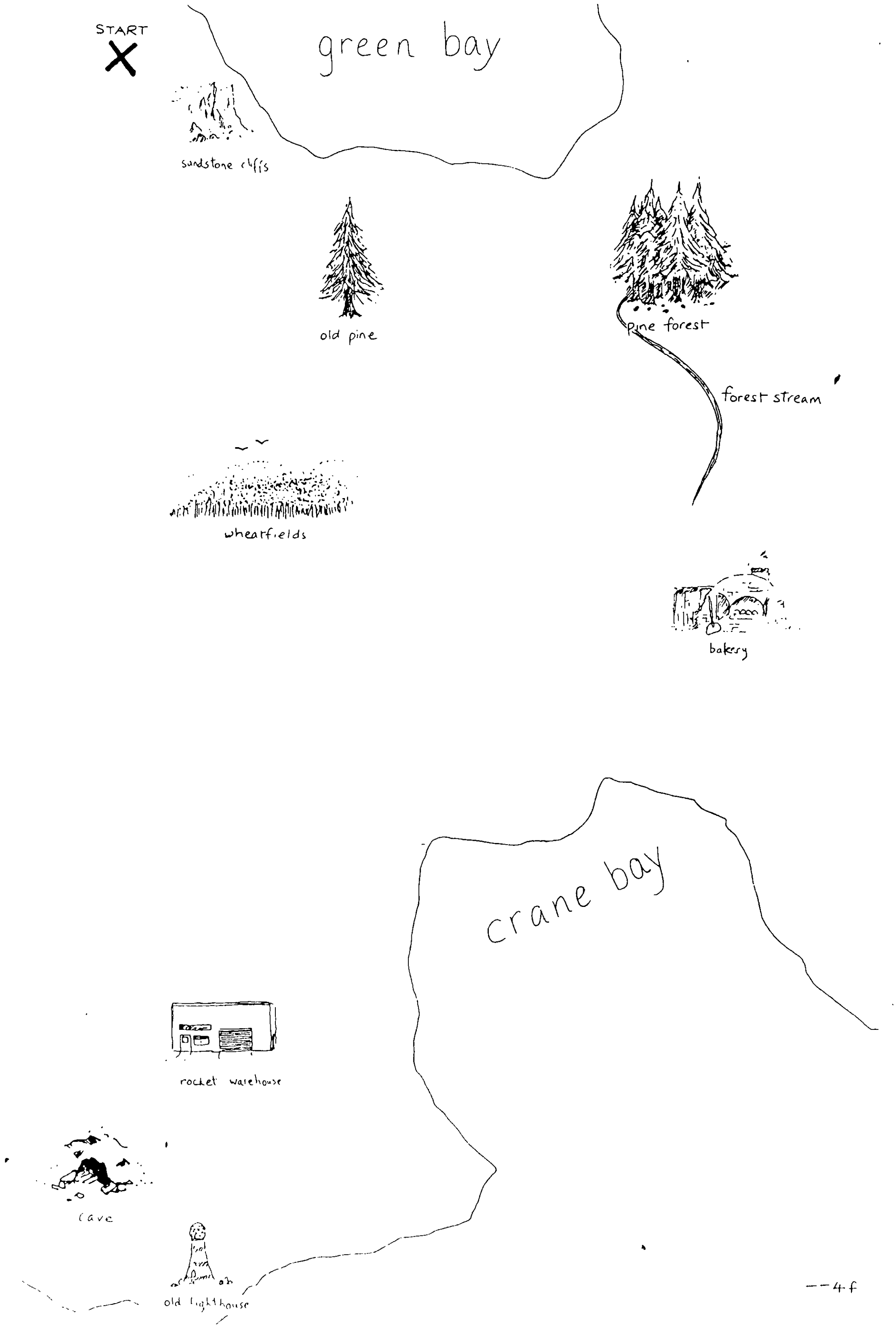
A.10 Map -+4 Instruction Followers' Map

A.11 Map --4 Instruction Givers' Map

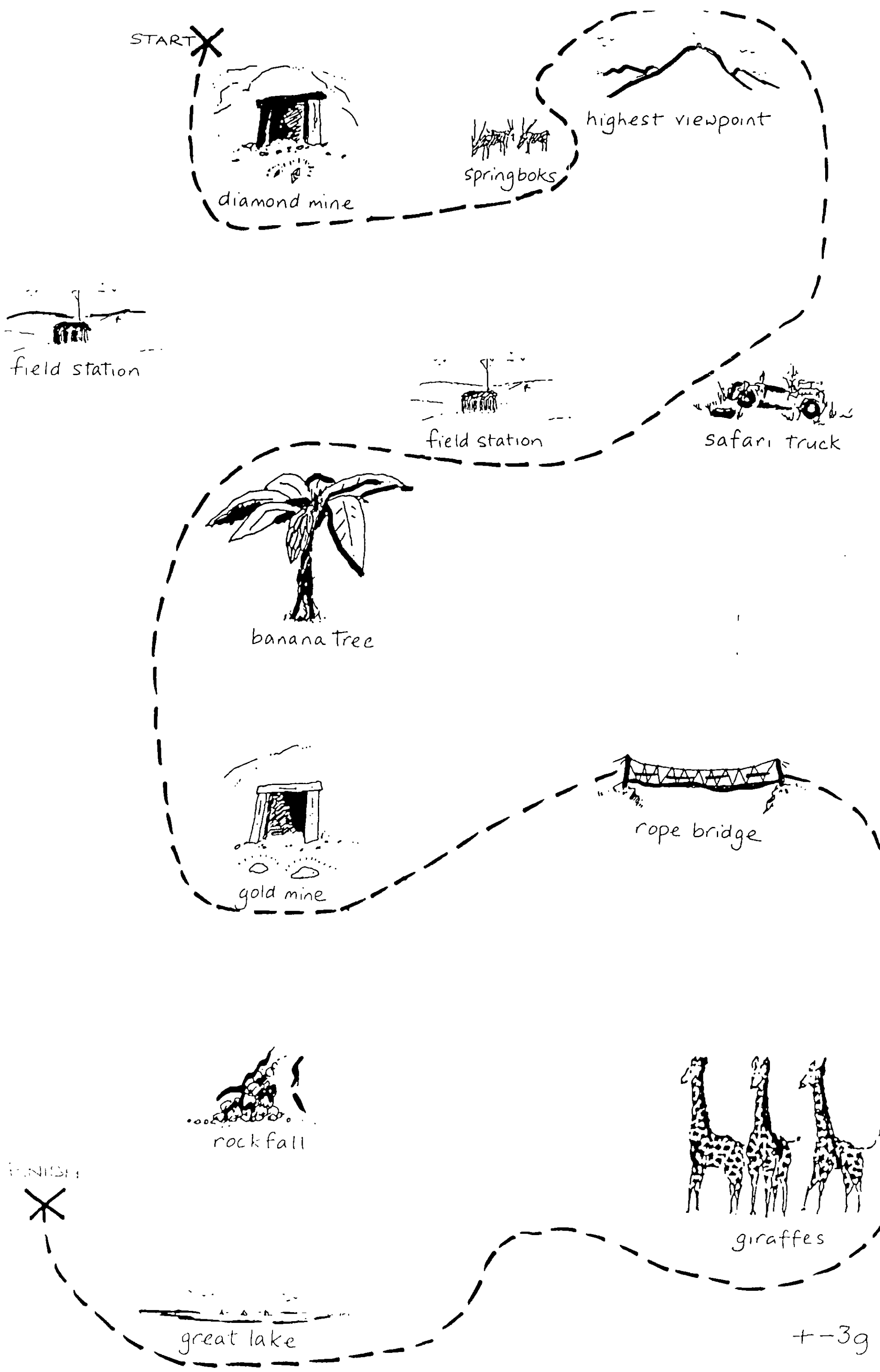


--4 g

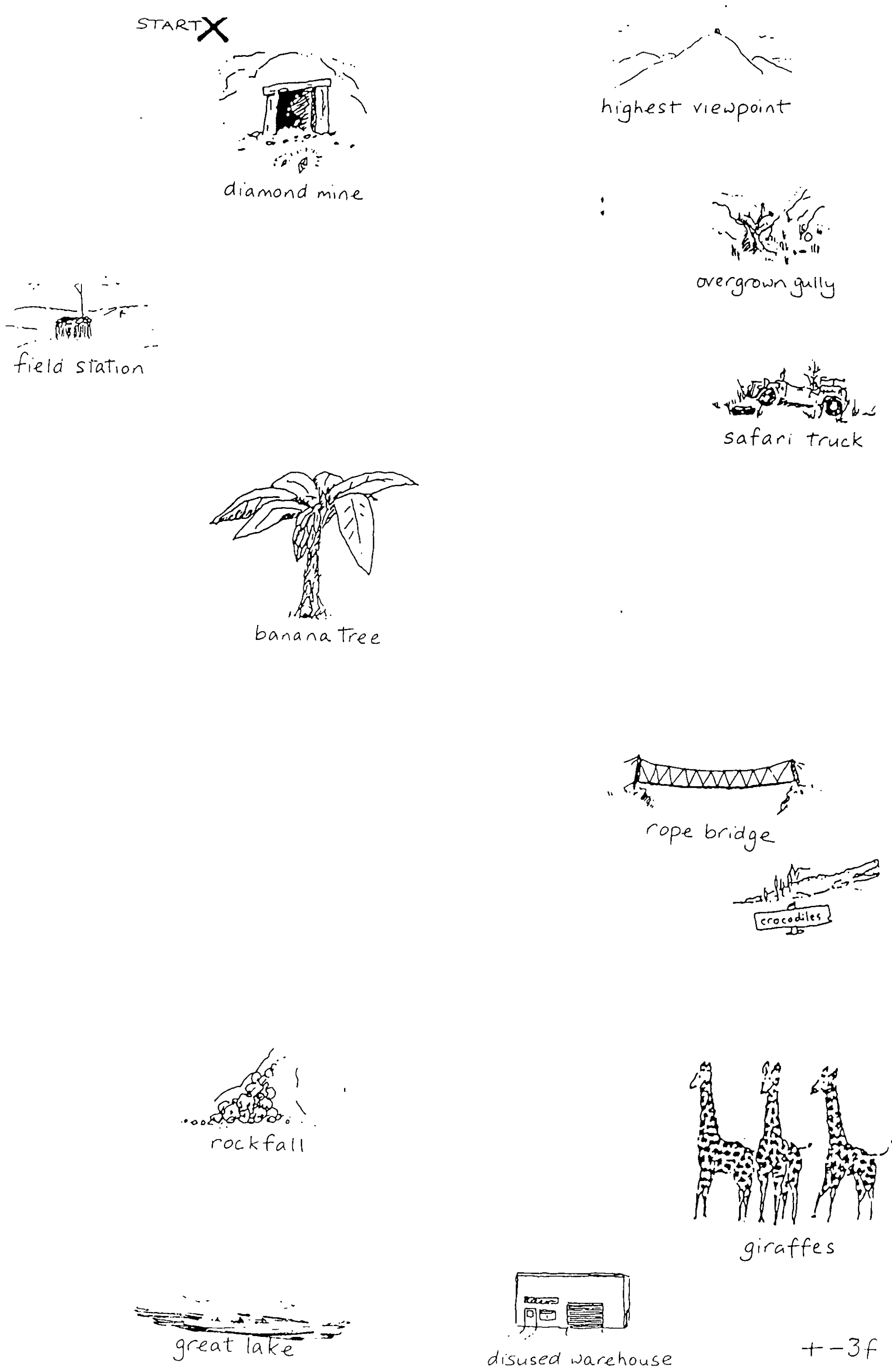
A.12 Map --4 Instruction Followers' Map



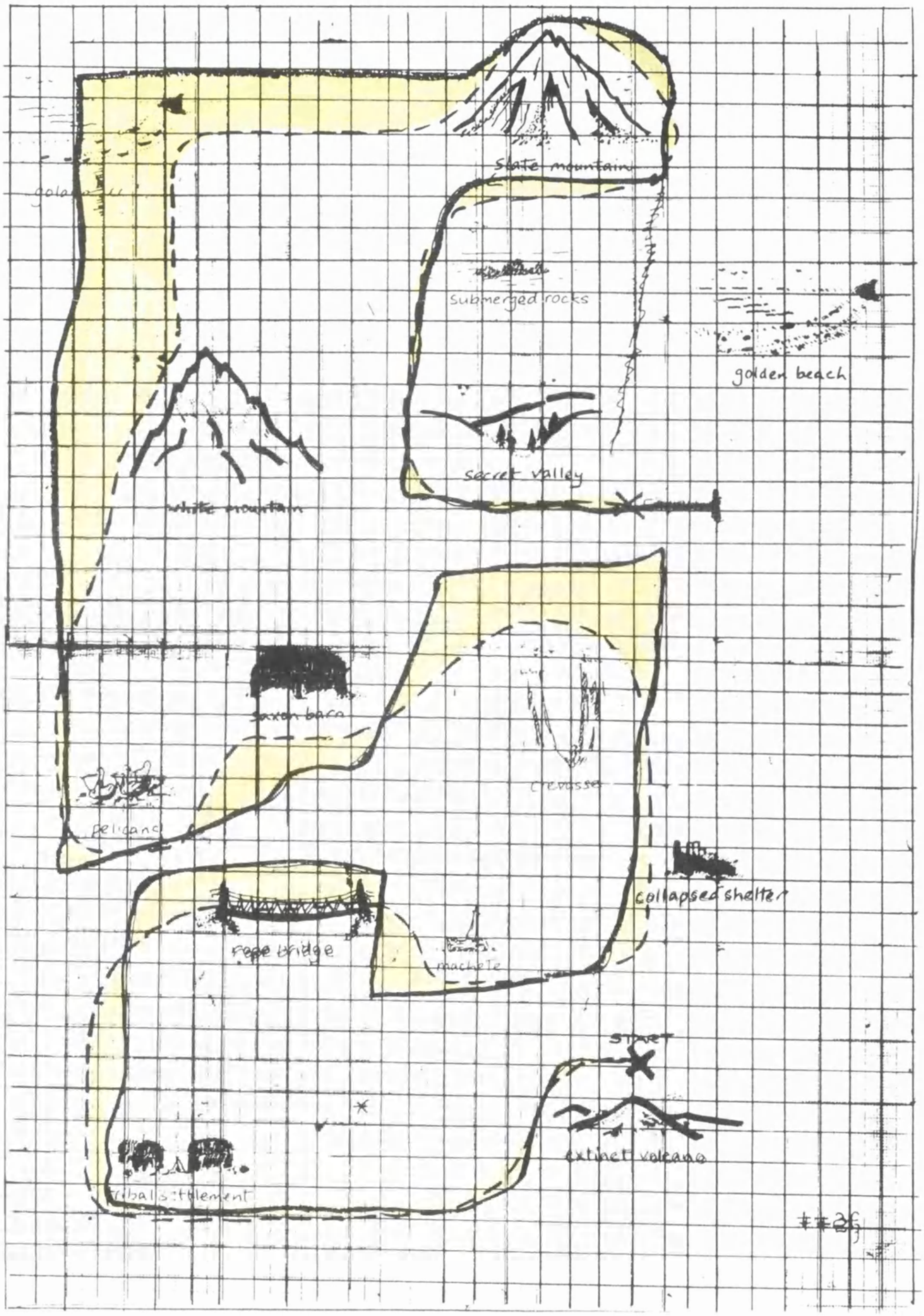
A.13 Reserve map +-3 Instruction Givers' Map



A.14 Reserve Map +- 3 Instruction Followers' Map

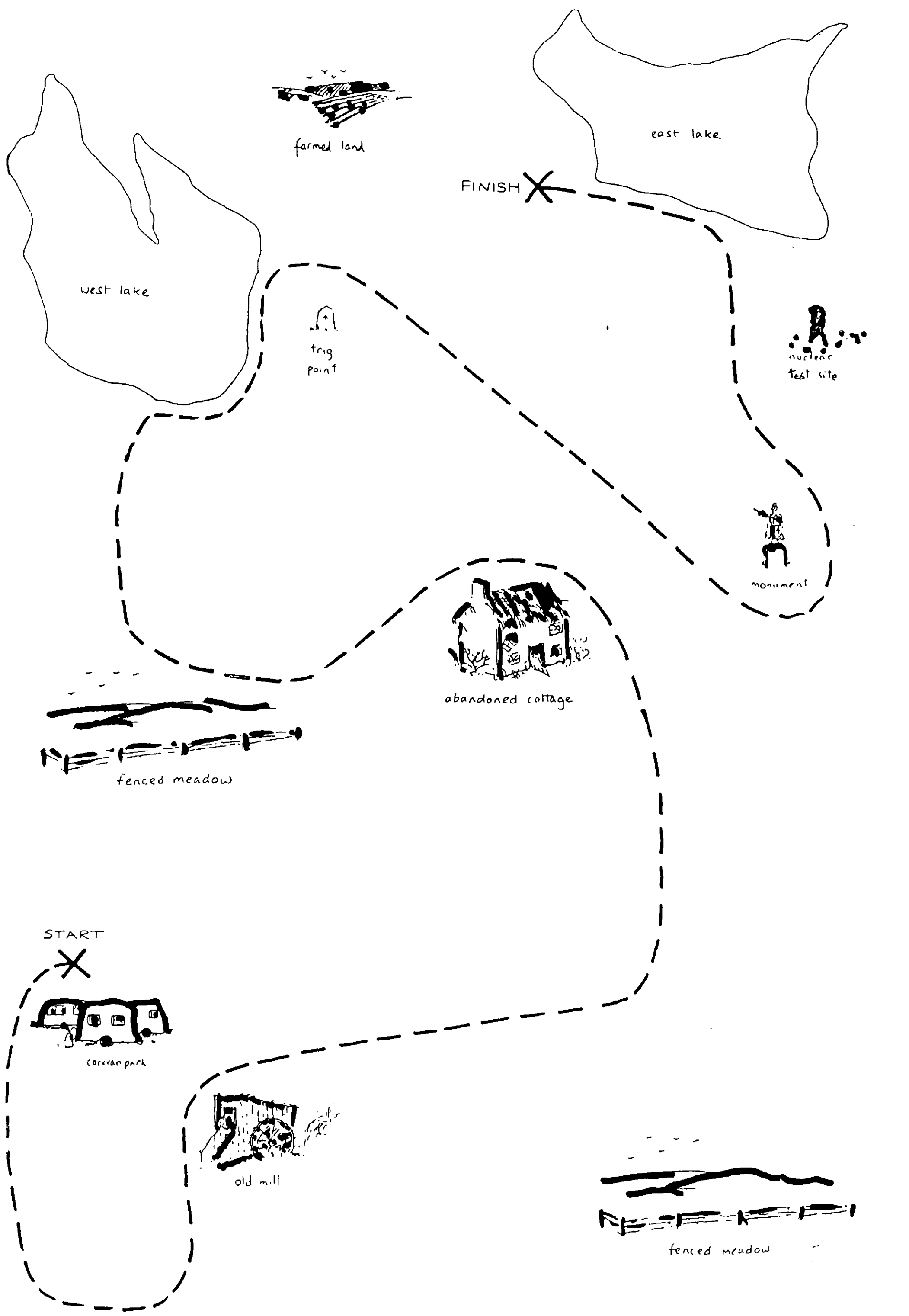


Appendix B: Study 1. Calculating Route Accuracy Scores for the Map Task. The continuous line is the Instruction Follower's route, the dashed line the original route, and the shaded areas show the deviance in square centimetres.



Appendix C: Study 2. Examples of HCRC Maps used in Study 2.

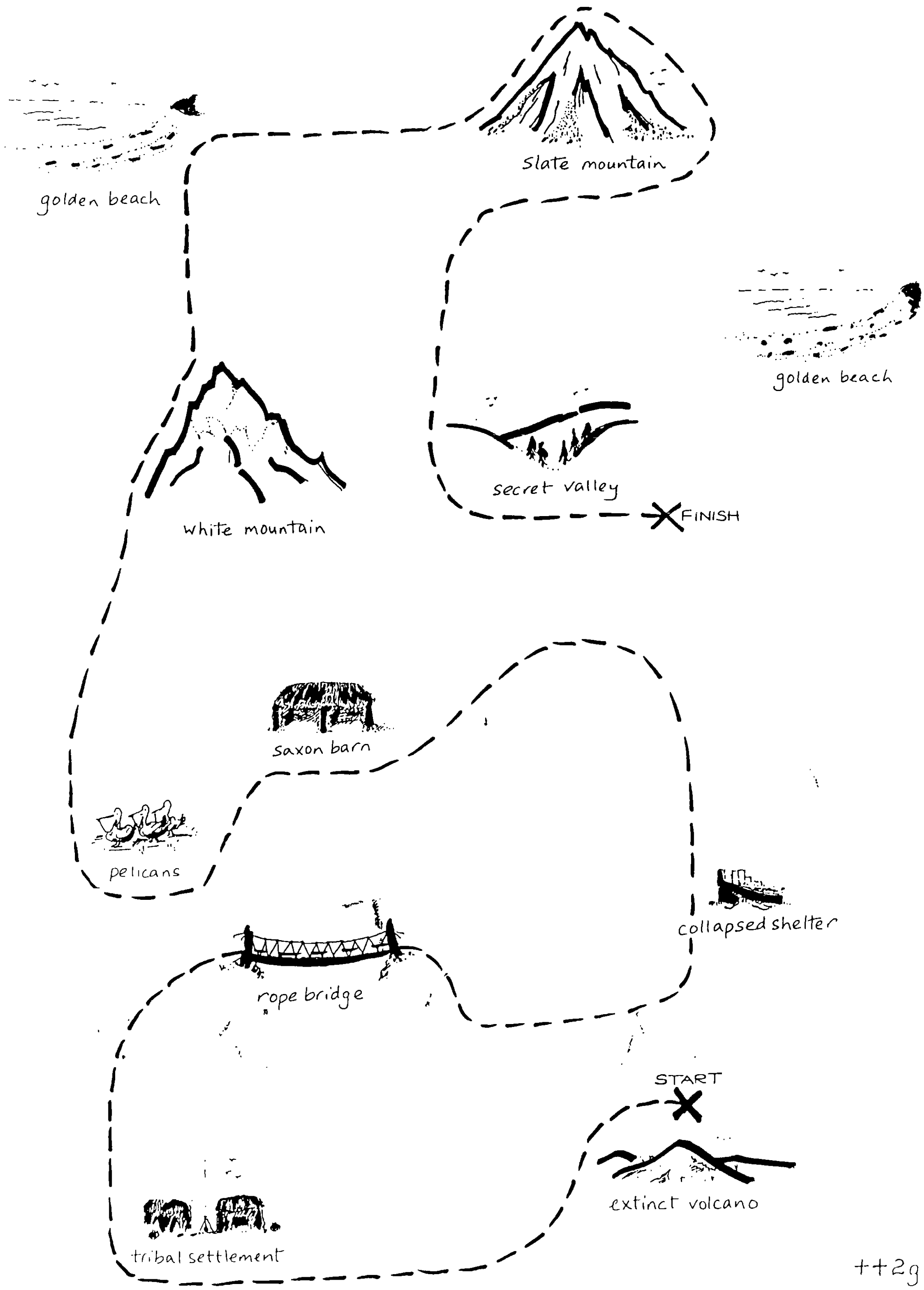
C.1 ++ 1 Instruction Givers' Map



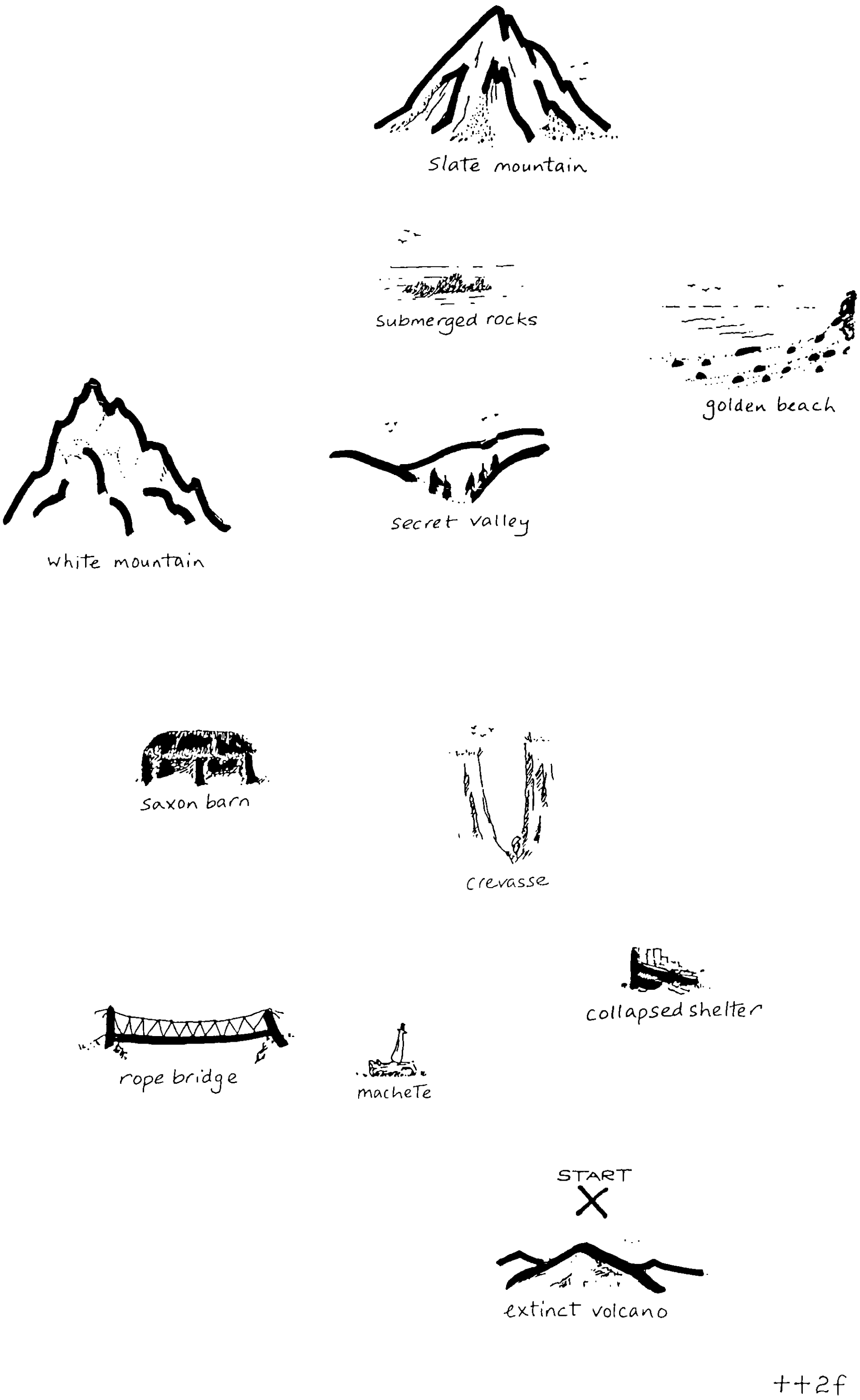
++13

C.2 ++1 Instruction Followers' Map

C.3 ++2 Instruction Givers' Map

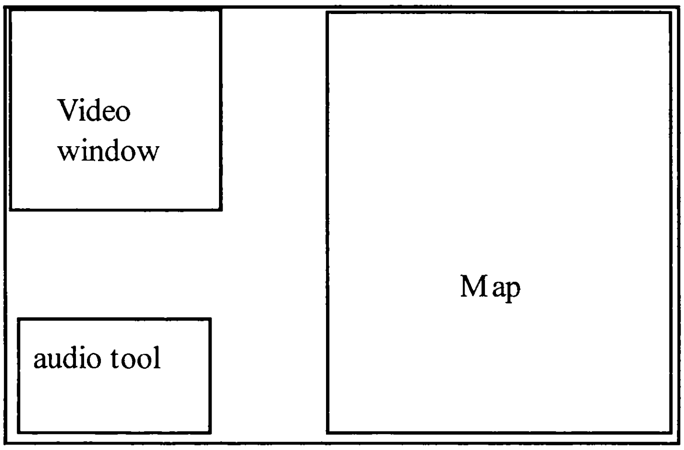


C.4 ++2 Instruction Followers' Map



Appendix D: Study 2. Monitor Configuration for VMC context

A line illustration of the monitor configuration for study 2.



Appendix E: Study 2. Questionnaire on User satisfaction with VMC context.

<u>Video Mediated Communication : Study 2</u>

Subject Pair .... date .... Role: IG IF

Familiarity of Partners. Approximately how long have you known your partner?

Years............. Months.........................

Ease of Communication: Please tick one answer for each of the questions below.

1. How easy did you find it to communicate with your partner?

a) very easy

b) fairly easy

c) neither easy nor difficult

d) fairly difficult

e) very difficult

2. How easy was it to take turns in the conversation?

a) very easily

b) fairly easily

c)average

d) with some difficulty

e) with much difficulty

Appendix E (cont)

3. Did you feel you were interrupting your partner?

a) very frequently	(i.e. every time you spoke)
b) moderately frequently	(i.e. every other time you spoke)
c) not very frequently	(i.e. occasionally)
d) rarely	(i.e. only once or twice)
e) not at all.	(i.e. none)

4. How often did you look at your partner?

a) very frequently	(i.e. every time you spoke)
b) moderately frequently	(i.e. every other time you spoke)
c) not very frequently	(i.e. occasionally)
d) rarely	(i.e.only once or twice)
e) not at all.	(i.e. none)

Appendix F: Study 2. Summary Tables of Responses to the Questionnaire on User satisfaction of VMC contexts.

Question 1. Ease of communication.

	Click To Speak	Open Channel
very easy	3	7
fairly easy	12	8
neither easy/diff	3	3
fairly difficult	2	2
very difficult	0	0

Question 2. Ease of turn-taking

	Click To Speak	Open Channel
very easy	6	13
fairly easy	8	7
neither easy/diff	4	0
fairly difficult	2	0
very difficult	0	0

Appendix F cont.

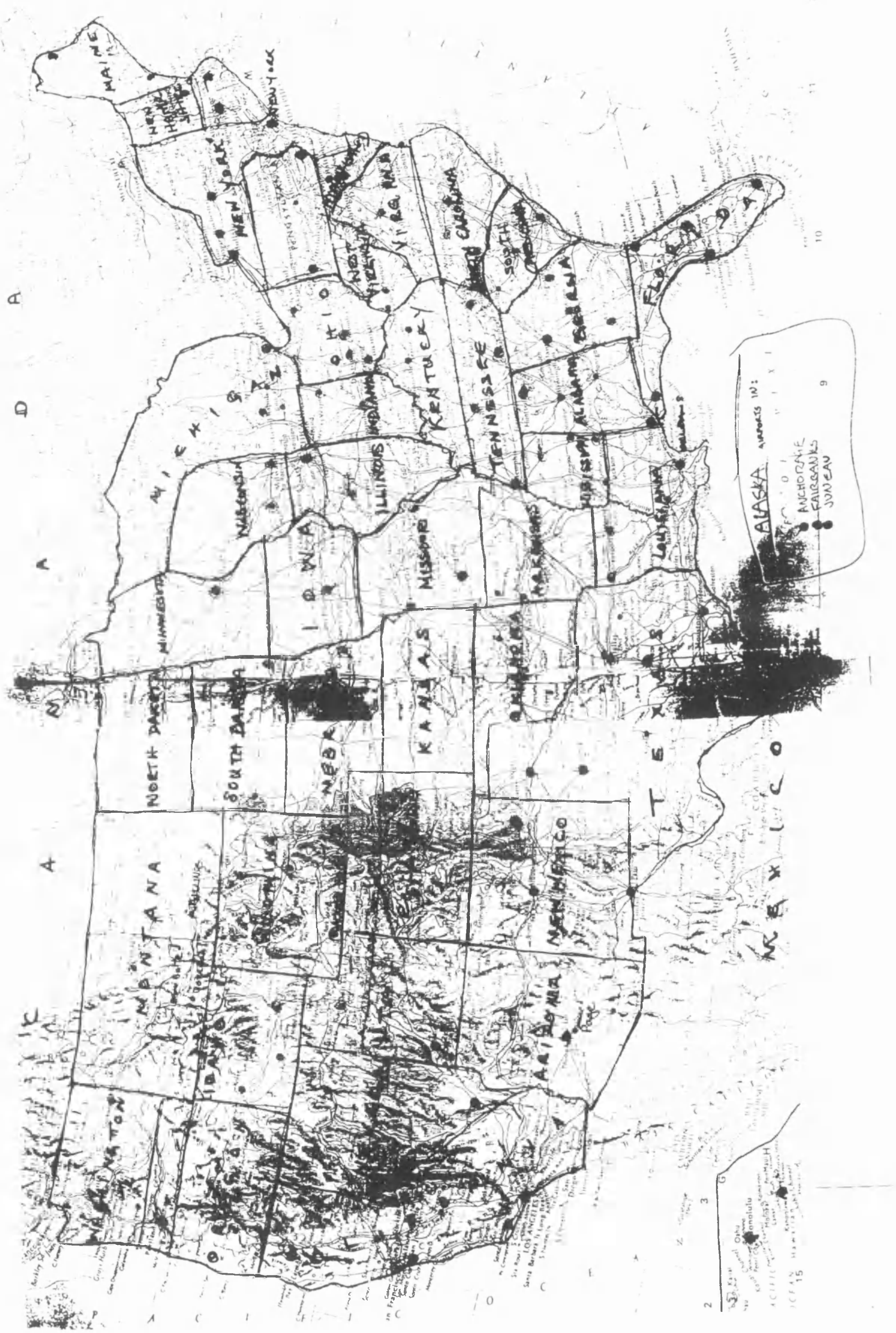
Q3. Subjective rating of frequency of interruptions

	Click To Speak	Open Channel
very often	0	0
moderately often	2	4
not very often	11	6
rarely	7	6
never	0	3

Q4. Frequency of looking at partner

	Click To Speak	Open Channel
very often	5	0
moderately often	3	5
not very often	8	8
rarely	2	7
never	2	0

Appendix G. Study 3. Map of USA for Face-to-face Spoken Travel Game.



Appendix H. Study 3. Questionnaire on Computer Experience of VMC participants.

1.How much experience have you had of using computers?

a) very experienced

b) experienced

c) moderately experienced

d) a little experience

e) no experience

Please give some indication of the type of experience you have had
e.g.. used word processor (Mac or PC), some programming skills etc.

....

....

2. Have you had experience of using a 'mouse' ?

a) very experienced

b) quite experienced

c) moderately experienced

d) a little experience

e) no experience

3. Have you had experience of text-based communication via computers ?

e.g. Have you used e-mail? **yes / no**

Appendix H. (cont)

4. Have you used any other types of computer-supported communication?

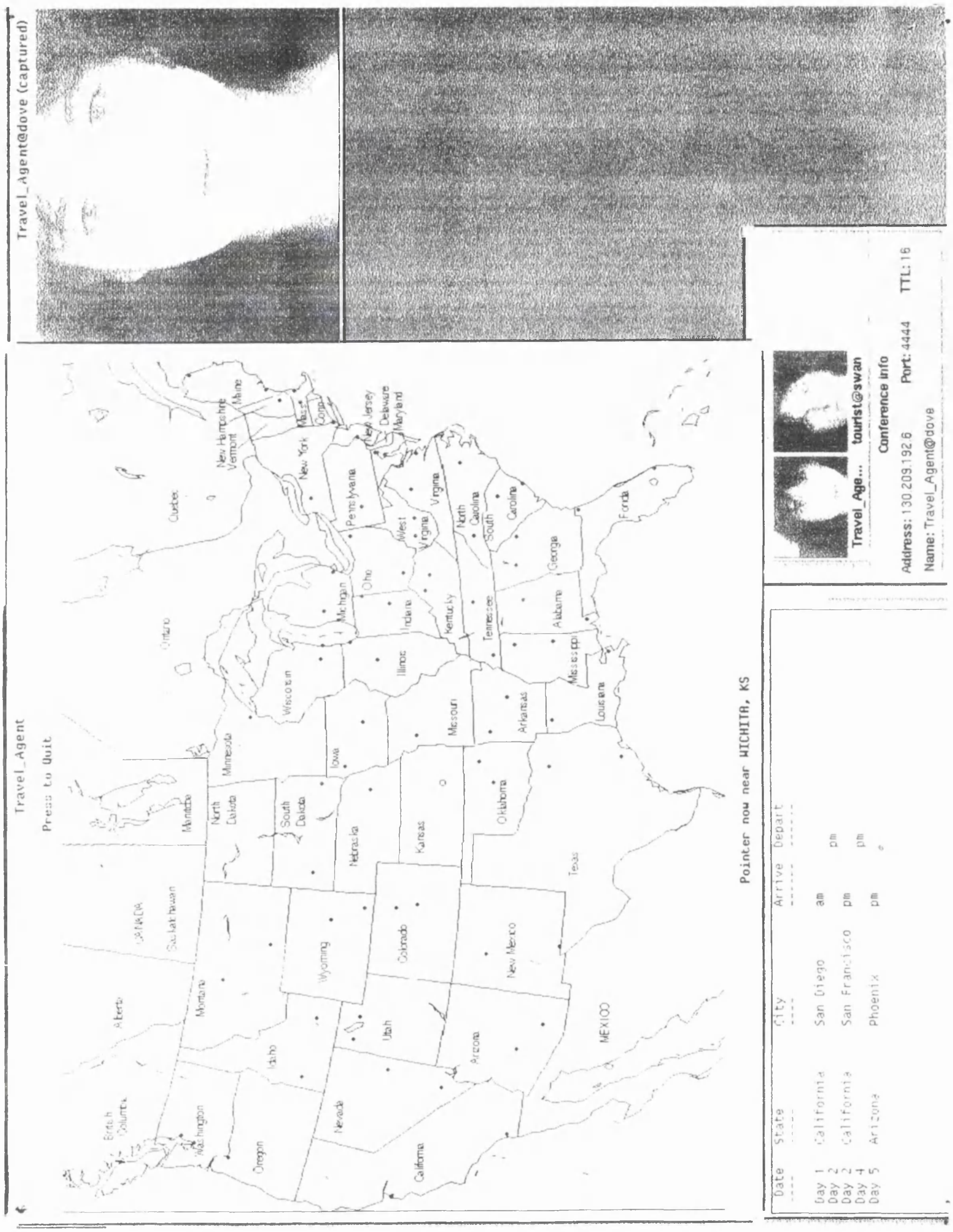
e.g. video-conferencing? **yes / no**

if **yes,** please give a few details of the type of communication and degree of experience:

....

....

Appendix J. Study 3. Monitor set-up for VMC version of Travel Game Task.

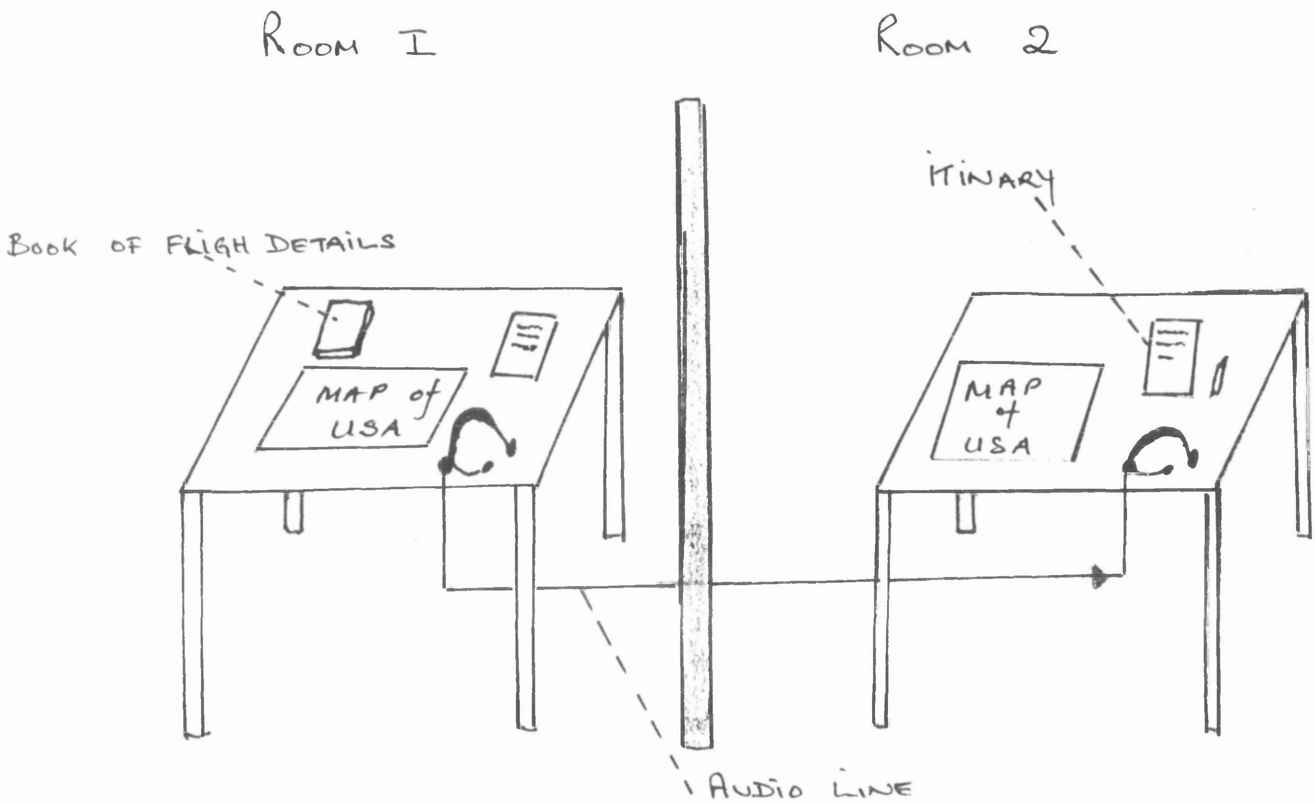


Appendix K. Study 3. Mean Freguency of Initiating Moves per dialogue. Raw scores

Group mean frequency of each Initiating Move per dialogue (raw scores)

Initiating Moves	**VMC**	**face-to-face**
Iinstruct	3.50	7.80
Directive	1.10	2.60
Explain	45.00	43.70
Query-yn	26.10	24.60
Query-w	17.30	35.5
Align	5.60	3.60
Check	9.00	19.10

Appendix L Layout for Travel Game 1. Spoken-only Context

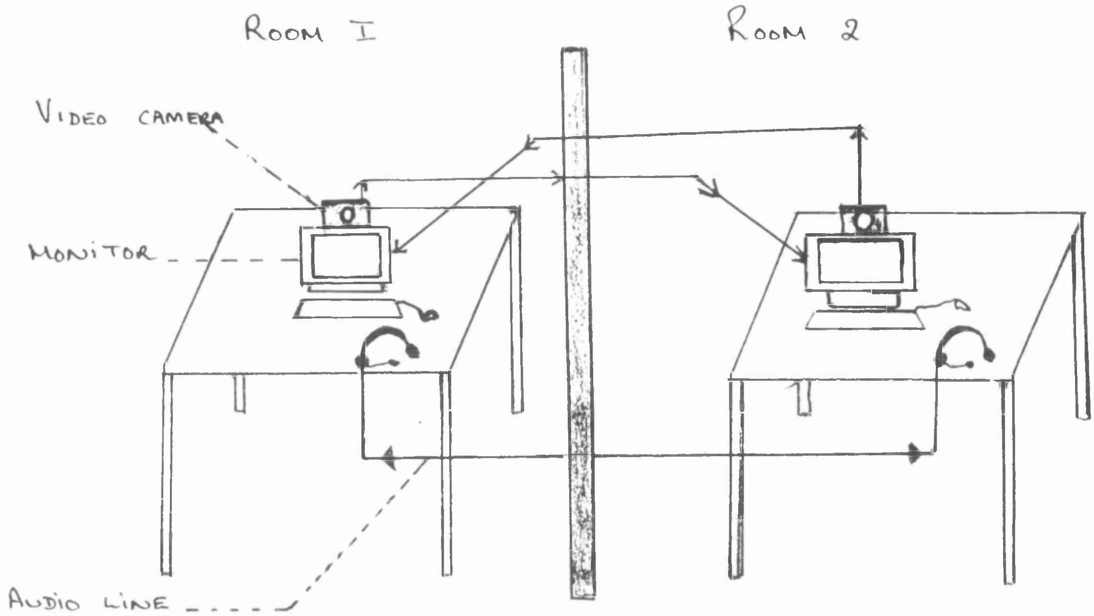


Appendix M. Study 3. Mean Number of Initiating Moves by Role of Participant. Raw Scores.

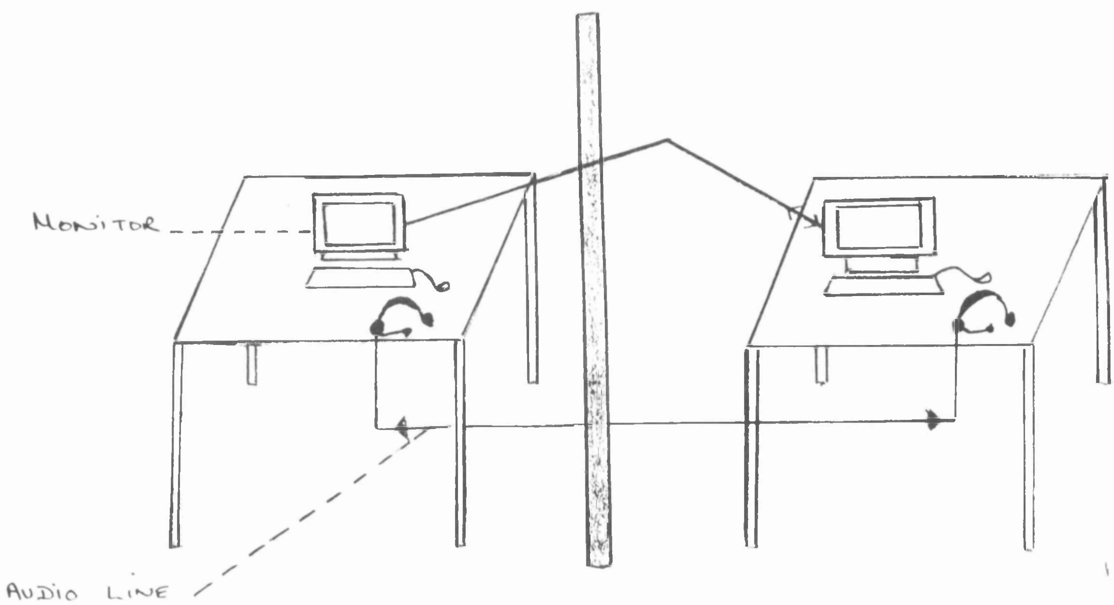
Mean number (raw scores) of Initiating Moves by Travel Agent and Client in the Face-to-face and VMC contexts.

Context	**Face-to-face**		**VMC**	
Role	**Travel Agent**	**Client**	**Travel Agent**	**Client**
Instruct	0.40 (0.69)	7.40 (6.04)	0.60 (0.84)	2.90 (4.58)
Directive	0.10 (0.32)	2.50 (4.06)	0.00 (0.00)	1.10 (0.99)
Explain	36.60 (11.09)	8.20 (11.31)	39.90 (12.32)	5.80 (6.36)
Query-yn	10.90 (3.14)	13.20 (10.98)	5.70 (2.36)	20.40 (12.18)
Query-w	26.90 (7.82)	8.90 (5.30)	11.50 (4.09)	5.80 (4.16)
Align	2.30 (1.71)	1.20 (1.38)	4.80 (3.99)	0.80 (1.23)
Check	4.90 (3.26)	14.80 (11.27)	4.60 (4.06)	4.40 (3.06)

Appendix N. Experimental Set-up for Travel Game: Video Mediated Contexts.



EXPERIMENTAL SET-UP: VIDEO MEDIATED CONTEXT



EXPERIMENTAL SET-UP: AUDIO-CONFERENCE CONTEXT