University of Skövde, Sweden
Department of Computer Science
MSc Dissertation 1996

More than Meets the Eye:

THE ROLE OF HUMAN VISUAL SIGNALS IN COLLABORATION

PIERRE GANDER
(pierre@ida.his.se)

Supervisor: Agneta Gulz
Subsidiary supervisor: Mikael Johanneson

Submitted by Pierre Gander to the University of Skövde as a dissertation towards the degree of MSc by examination and dissertation in the Department of Computer Science, October 1996.

I certify that all material in this dissertation which is not my own work has been identified and that no material is included for which has already been conferred upon me.

Signed...

Tell me and I'll forget;
show me and I may remember;
involve me and I'll understand.

Chinese proverb

CONTENTS

LIST OF FIGURES

Abstract

The aim of this dissertation is to study the importance of human visual signals in human collaboration. Visual signals include, among other things, *gesture*, *facial expression*, and *direction of gaze*. The motive is to see if visual signals are important to include in systems for computer-mediated communication (CMC). Two approaches are used: a literature study and an empirical study. The hypothesis in this study is that visual signals improve communication in task collaboration, increasing task performance. The literature studied reveals two classes of tasks: first, where people cooperate towards a shared goal, and second, where people negotiate and work towards different goals. Only in the second class does the literature show any significant differences in task performance. The empirical study conducted contained 80 subjects who solved two tasks in pairs divided into two groups: one with visual signals (20 pairs), and one without (20 pairs). The tasks were: 1) making a drawing from another person's description of a photograph, and 2) answering questions on the contents of a narrative read and re-told by another person. Availability of visual signals was the independent variable, and the dependent variable was task performance in terms of judged similarity between drawing and photograph, and score on questions on the narrative. No significant differences in task performance were found between the two experiment situations. The conclusion is that visual signals would not be essential to include in a CMC system with cooperative tasks where people work towards the same goal. However, visual signals would be important when expression of emotion and values is needed, as in negotiations in distributed meetings.

Preface

One winter evening, me and a friend of mine were sitting at his place talking. Something in our conversation triggered my friend to get up and say: ”Let me show you a videotape of a virtual reality project I came into contact with.” The project was called DIVE and soon we were watching different demonstration clips on his VCR showing various three-dimensional landscapes. ”This one shows a distributed meeting”, my friend said. I could see several block-shaped objects centered around a table, and some sort of a whiteboard used for drawing on in the background. ”The blocks are the other meeting participants”, I was informed. When I looked closer I could see that each of the blocks had a simple, unchanging face, consisting of some dots and lines for eyes, nose, and mouth. For participants of a meeting, I thought they looked rather inexpressive. ”This can't work for a real meeting”, I said. ”You can't see who is talking, if the other participants agree, are angry, or bored. Maybe someone is trying to ask for permission to speak?”. Here, lots of people had spent lots of time to create this system to have distributed meetings in, and probably had the belief that it would work. My intuition told me otherwise. It was a question of how human visual signals function in communication: is visual information beyond that of the block-shapes shown in this system necessary for communication to function properly? This was the start for this investigation of the role of visual signals in communication. I wanted to study the literature on the sub-

ject, but also perform a large scale empirical study of visual signals in collaboration. This project is what is reported on in this paper.

This work has mainly been a single person's work, but on some occasions, I had the pleasure to collaborate (mostly visually) with a number of persons whom I wish to thank. Without Lee Sandberg, I wouldn't have got the idea in the first place. Pernilla Jensen, Johan Språng, Zoltán Biró, and Henry Rindebäck were helpful in the pilot study. The latter three are also remembered for interesting discussions. The latter two, along with Zayera Khan, Magnus Haake, Hanna Risku, and Ajit Narayanan, are thanked for commenting on drafts of this dissertation. Staff and MSc students at the department of computer science of University of Skövde gave useful criticism at my oral presentations. Ingvar Lind provided useful picture expertise. Finally, my two supervisors (Agneta Gulz and Mikael Johannesson) were very helpful, in spite of summer holidays, care-taking of infants, and events of burning cars.

1 INTRODUCTION

The White Rabbit put on his spectacles. "Where shall I begin, please your Majesty?" he asked. "Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

Lewis Carroll, *Alice's Adventures in Wonderland*

1.1 BACKGROUND

The situation where people who need to communicate are geographically spread is becoming more and more frequent. Being too far away from each other to communicate face-to-face, they need a way to mediate their communication. With the exception of personal messengers and mail, mediation of communication relies on technology. Earlier, this has often been achieved by telephone, or sometimes by using a TV camera in teleconferencing. Now, using computers to communicate at distances is becoming more widespread. This mediation of communication through computers is referred to as *computer-mediated communication* (CMC). CMC is becoming increasingly used, mostly because of the fast growth and spread of the Internet among different groups in society.

Also, the work place is becoming distributed. People work together separated by geographical distances, mediating their communication through technology. Using the computer as tool in cooperative work has resulted in the multi-disciplinary research field called *computer supported cooperative work* (CSCW). Mediating cooperation through computers leads to a number of questions: How should software be written to best facilitate cooperative work? In pursuing this, a more fundamental question arises: What basic mechanisms are at work in human cooperative work in the first place?

In this dissertation, one such issue is investigated: Is visual contact between participants essential in cooperative work? ”Visual signals” include such elements of communication as gestures, facial expressions, direction of gaze, and body posture (see section 2.1.4 for a proper definition).

When people communicate, they spontaneously use gestures, facial expressions, and other visual signals. Among other things, our bodies provide information about our intentions, our attitudes, reference to external objects, and who is to talk when in a conversation. A strong position that visual signals are a necessary part of communication, in real life as well as in computer interfaces, is expressed by Norman (1992). He claims that in order to have graceful communication, visual signals which control the interaction need to be present. Without them, communication breaks down. We would never know if people like what we say to them, if they have understood, if they are sad, happy, or bored. Maybe they want to interrupt and speak?

Computerized systems for human collaboration does not presently take these facts into account. Systems for distributed meetings, such as DIVE (Ståhl & Andersson, 1995) and MASSIVE (Greenhalgh, 1996), only supply a minimum of visual signals, and as hypothesized in this dissertation, too few to work properly.

Why are not visual signals included in systems for collaboration? And are they really important to include? These are the questions to be answered in this thesis. It is important to have an answer to whether visual signals are an essential part of successful collaboration. If they are, they should be included in a computer-mediated communication system.

A possible explanation of why visual signals have been neglected in computerized communication systems is the following. Because we are unaware of using gestures when we communicate, we tend to see them as unimportant when we sit down to think about what needs to be included when building a system for communication. The most salient aspect appears to be the transfer of ideas by the usage of *words*, and sometimes this is only what ends up being built into the system. Also, visual signals could have been neglected because they are technically difficult to implement. However, with the increased computational and graphical power of today's technology, this should no longer pose an insuperable problem.

1.2 AIM

The aim of the dissertation is to investigate the importance of human visual signals in computer-mediated communication (CMC). If visual signals are found to be important, they should be included in a computerized system for human communication.

1.2.1 Hypothesis

The hypothesis is that visual signals *are* an important aspect to include in CMC; that tasks are solved better when visual signals are present.

1.2.2 Delimitation

The focus of the study is on *the collaboration of people*. Thus, the technical issues of particular computer systems for collaboration will not be described in detail. This means that the study really is about human collaboration in general, not tied to any specific technology.

Details of *visual signal recognition* and *generation* will not be investigated. These are two large areas where much research is going on. The study performed does not make any assumptions on any features or limitations of visual recognition or generation. Rather, it is assumed that the technical problems in these areas can be solved, making them function in the same way natural recognition and generation does in humans (this assumption is also made by Chapanis, 1975).

A delimitation to study a special subset of visual signals could have been made, but as no reason for such a delimitation was seen, instead *all* visual signals were treated as a single group. There are some methodological problems involved in controlling only a subset of visual signals. If subjects are to act natural in a collaborative situation, they can not be instructed to use only certain visual signals and not other (because of the unconscious nature of visual signals, further described in the background in Chapter 2). Prevention of certain visual signals could be made technically, such as cutting holes at various places in a screen standing between subjects collaborating, making them see only selected parts of the other person's body. However, communication cannot be too heavily manipulated without making it unnatural. Short *et. al.* (1976) say: ”We must recognize that the communication stream is a total stream and it is dangerous to isolate items of information in individual channels from the total message” (p. 161).

The study was delimited to collaboration between pairs of persons. Studies of collaboration between two persons has been conducted by Chapanis and colleagues (Chapanis *et. al.*, 1972; Ochsman & Chapanis, 1974, Chapanis, 1975). Otherwise, mostly studies of groups larger than two persons have been made (e.g. Short *et. al.* 1976; Williams, 1977; Olson & Bly, 1991; Tang, 1991; Tatar *et. al.*, 1991; Valacich *et. al.*, 1991; Isaacs *et. al.*, 1995). Findings from studies of groups consisting of two persons can probably not be generalized to groups with more participants. This issue will be discussed further in Section 6.3.1.

The collaboration of interest in this study is assumed to take place in a work place environment. This assumption links the study to the field of CSCW. Other situations where distributed collaboration could be used are, for example, leisure situations in the home, where CMC could be used for social or entertainment purposes. These cases are however not covered in the study.

1.3 OBJECTIVES

The dissertation was carried out in two main parts. First, literature was studied to answer the question: Are visual signals important to include in CMC? Literature on *non-verbal communication*, *computer-mediated communication* (CMC), *computer supported cooperative work* (CSCW), *computer supported cooperative learning*, and related literature was studied to see if support to include visual signals was found there.

Secondly, a large scale study was carried out to empirically study the importance of visual signals. The focus was on ”hard” areas such as success in problem solving and reasoning, as opposed to more ”softer”, subjective ones like attitudes, feelings, and quality of life. In the first case, objective measures of the success of task performance can be obtained. The focus in the study was on how well persons *actually perform* on tasks. The second case only provides indirect measures of task performance, and was not chosen for this reason.

In order to investigate the importance of visual signals in collaboration, the general idea was to compare task performance in two situations: one with visual signals, and one without visual signals. Subjects were asked to carry out tasks under these two situations, and the results were compared. An existing computerized system was not used in the study. Instead, subjects were tested under ”technology-neutral” conditions, which has a number of advantages (further discussed in Section 4.5).

First, a pilot study in small scale was carried out with the purpose of evaluating the test (design, material, procedure, choice of subjects, etc.). Using experiences from the pilot study, the real study was then carried out on a larger scale.

1.4 COGNITIVE SCIENCE PERSPECTIVE

This paper studies collaboration and communication from a viewpoint of cognitive science. It draws on evidence from different areas, mostly *psychology* and *linguistics*. The study of *language use*, the focus of areas such as *psycholinguistics* and *sociolinguistics*, connects the areas of psychology and linguistics. Parts of the evidence discussed in this paper comes from *neuroscience*, especially McNeill's (1993) argument that speech and gesture are one and the same underlying system.

The results have practical implications for *computer science*, providing guidelines for how systems for human communication should de designed.

Other perspectives *not* applied in this paper, but occurring in the CSCW literature, are *management* and *organizational theory* points of view.

1.5 TARGET GROUP

This report is written with a target group in mind consisting of persons with knowledge corresponding to at least MSc level in one of the cognitive sciences, and with rudimentary knowledge of the others.

1.6 PREVIEW OF FORTHCOMING CHAPTERS

The remaining part of the dissertation is organized as follows:

In chapter 2, a background to human communication and collaboration is given in order to specify the conceptual framework for the dissertation. Both theoretical aspects as well as aspects of computerized systems are described.

Chapter 3 presents the results from the literature study in trying to answer the question of the role of visual signals in collaboration.

The empirical study is described in detail in chapter 4.

The results of the empirical study are presented in chapter 5, and are discussed in chapter 6.

Chapter 7 contains some final conclusions, while chapter 8 points to some future research.

2 BACKGROUND

There is more to speaking than just using words.

Ronald Wardhaugh (1993, p. 171)

This chapter provides the background to the dissertation. It sets the conceptual context and defines important concepts that will be discussed later in the paper. The *communication process* is described, with emphasis put on *non-verbal communication*. An early model of communication, that of Shannon and Weaver, is presented, followed by Jakobson's more recent one which takes culture and context more into account. Further, the fields of *computer-mediated communication* (CMC) and *computer-supported cooperative work* (CSCW) are introduced in order to give some background to distributed collaboration. Finally, as an application of CSCW, examples of collaborative systems using *virtual reality* are described.

2.1 WHAT IS COMMUNICATION?

Communication is a human activity of which everyone knows, but of which few can give a good definition. Fiske (1982) makes five general assumptions about communication (p. 11):

- It is possible to scientifically study communication.
- All communication includes signs and codes. Signs are things or acts which refer to something other than themselves. Codes are the systems in which signs are organized, and which give rules for how signs relate to each other.
- Signs and codes are transmitted or made available to others, and transfer or receiving of signs/codes/communication is equal to practice of social relations.
- Communication is of major importance for life in our culture: without it, no culture can survive. Consequently, the study of communication includes study of the culture in which it is integrated.
- Behind these assumptions is a general definition of communication as ”social interaction by use of messages”.

A number of theories of communication, created for different purposes, have been developed.

The study of communication can be divided into two main schools (Fiske, 1982). The first one is the *process school*. Here, communication is seen as a transfer of messages, and weight is put on how sender and receiver codes and decodes, and how they use channels and media in communication. The second school sees communication as *creation and exchange of meaning*. Emphasis is put on the *function* of communication and on the cultural context in which communication exists. It is grounded in the area of *semiotics*, the discipline concerned with signs and their relation to reality.

Short *et. al.* defines communication from a social psychology point of view as ”the physical signals whereby one individual can influence the behaviour of another” (Short *et. al.*, 1976, p. 153).

This present study agrees with the process school in the study of communication. Communication is considered an active two way process in which information is transferred between persons. However, the importance of culture and context is also recognized.

2.1.1 Fundamental concepts in communication

There are some concepts commonly used when discussing communication (from Fiske, 1982):

- *Channel*. The channel is defined as the physical substrate in which the signal is transferred.
- *Media*. The media is defined as the technical means for transforming the message to a signal that can be transferred via the channel. Media can be divided into three main categories:
  1. Presentational media: voice, face, body. These media use the expression of the spoken word, facial expression, gesture, etc. The communicator must be present, since he or she is the media. They are limited to the here and now. They produce *communication acts*.
  2. Representational media: books, paintings, photographs, etc. They can record the presentational media (category 1) and exist independently of the communicator. They produce *communication products*.
  3. Mechanical media: telephone, radio, television. They function as carriers of presentational and representational media. The major differences, as described by Fiske, between representational and mechanical media is that mechanical media use channels created by technology which suffer from technical limitations and are more exposed to noise than representational media.
- *Sign*. A sign is defined as a physical signal which stand for something other than itself.

- *Code*. Code is defined as a system of meaning common to members of a culture or subculture. It consists of signs together with rules or conventions for how and in which situations the signs can be combined to form more complex messages.
- *Redundancy*. Redundancy is defined as the result of high predictability; what is predictable or conventional in a message. That a message is redundant does not mean it is superfluous. Redundancy is not only useful in communication, it is absolutely necessary in practice (Fiske, 1982). There are two major functions of redundancy in communication. Firstly, a technical function defined by Shannon and Weaver (1949), and secondly, an expansion of their concept to the social dimension. The first way redundancy function is to compensate for errors in transmission. For instance, in a noisy communication situation, such as over a poor telephone line, we can repeat what we just said. By having redundancy in the message also means that receivers with different background knowledge are able to decode and understand the message. The other way redundancy functions is in social context, where it is used in *phatic communication* to keep already open communication channels open and ready. For instance, saying ”hello” to a friend in the street transfers a message of high redundancy. The message does not transfer any new information, it only serves to keep and strengthen an existing social relation.
- *Feedback*. The concept of feedback originates from the area of cybernetics, where it is defined as a flow of information back to its origin; a circular causal process in which a system's output is returned to its input (Miller, 1995). In communication, feedback means transfer of the receiver's reaction back to the sender. Feedback is essential in communication and helps the communicator to adapt the message to the receiver's needs. Channels allow varying degrees of feedback. For instance, telephones only allow use of voice for feedback, while communicating face to face allow people to send with their voice and receive with their eyes.

2.1.2 Models of communication

Shannon and Weaver's (1949) ”Mathematical Theory of Communication” is considered one of the most important seeds from which communication theory has grown (Fiske, 1982), see Figure 2.1. They described communication as a simple, linear process of sending a signal from sender to receiver. The focus is on the *precision* and *efficiency* of communication. In their model, Shannon and Weaver defined *information* technically as a measure of the predictability of the signal; the number of possibilities open to the sender. Nothing is said about the contents or the meaning of the message. Further, they defined the concepts *redundancy* and *entropy* (a measure of the disorder of a message) mathematically as inversely proportional to each other. A message with high redundancy, has low entropy, and has low information content. Conversely, a message with low redundancy has high entropy and is information-rich.

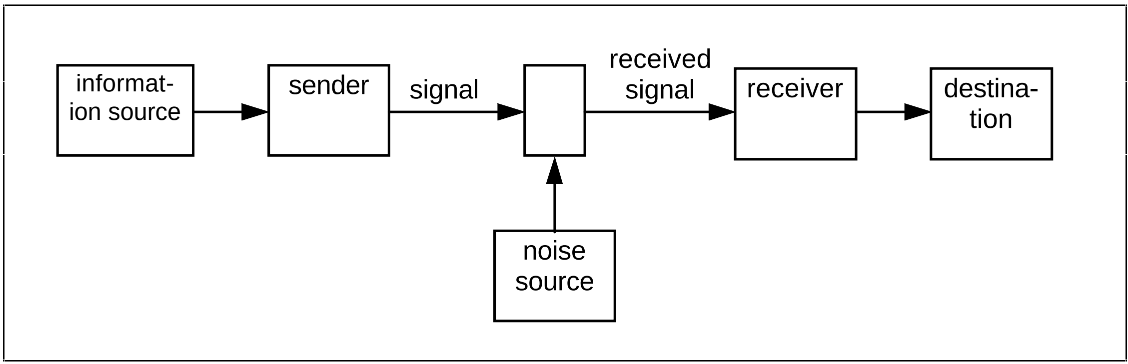

Figure 2.1. Shannon and Weaver's model of communication.

Shannon and Weaver's model of communication has a number of shortcomings. The meaning is assumed to be captured by the message alone, no reference to context or culture is given, although the meaning can be considered to reside as much in the culture as in the message (Fiske, 1982). Shannon and Weaver's concept of *noise* includes everything added to the signal which is not intended by the sender, comes from another source, or makes the message harder to decode correctly. This way of classifying noise would lead to the undesirable case where parts of communication, which are considered important, must be categorized as noise (Fiske, 1982).

A model that goes beyond the pure process-oriented view of communication, into the semiotic school, is that of Jakobson's. It will be described here because it builds on Shannon and Weaver's model, while at the same time emphasizing culture and context as important parts of communication. Also, in Jakobson's model, unintentional visual codes need not be classified as noise, as would be the case in Shannon and Weaver's model. These features made Jakobson's communication model suitable to the present study.

First, Jakobson describes the *fundamental factors* in a communication act, illustrated in Figure 2.2. These six factors must be present for communication to be possible. Next, he describes which functions the communication act performs for each of these factors. In his model, a *sender sends* a *message* to a *receiver*. The message refers to the *context*. Two other fundamental factors are *contact* (the physical channels and the psychological connections between sender and receiver) and *code* (a common system of meaning).

Sender	Context Message Contact Code	Receiver

Figure 2.2. Jakobson's fundamental communication factors (adapted from Fiske, 1982).

Emotive	Referential Poetic Phatic Metalinguistic	Conative

Figure 2.3. Communication functions in Jakobson's model (adapted from Fiske, 1982).

Each of these factors decide a particular function of language. These six functions are: emotive, conative, referential, poetic, phatic, and metalinguistic, as shown in Figure 2.3. The *emotive* function describes the relation of the message to the sender. The *conative* function is the effect of the message on the receiver. The *referential* function is how the message relates to reality. The *phatic* function serves to keep already open channels for communication open. The *metalinguistic* function identifies the type of the message. Finally, the *poetic* function is the aesthetic relation of the message to itself.

2.1.3 Non-verbal communication

The term non-verbal communication refers to everything in communication which is not the words themselves (Fiske, 1982). This includes gesture, facial expression, and tone of voice, among other things. A popular term often used is ”body language”, which is the part of non-verbal communication which can be expressed using the body. Wertheim (1996) emphasizes the importance of the non-verbal codes by noting that the verbal part of the message actually means less than the non-verbal part.

Language can only be understood if its relationship and interaction with its non-verbal part is recognized (Linell, 1978). Verbal codes differ from non-verbal codes in that the latter can only be used to refer to the present situation; they can only give information about the ”here and now” (Fiske, 1982).

Non-verbal communication has two main functions (Fiske, 1982). Firstly, it gives information about the speaker and his or her situation: identity, emotions, attitudes, social status, etc. These are socio-emotional signals. Secondly, non-verbal codes control the interaction. By using certain gestures, body posture, and tone of voice, the speaker can dominate others, bond with them, or reject them. Subtle non-verbal signals indicate things in a conversation such as turn-taking and that one part wants to end the talk. McNeill (1992) also mentions that gestures can be used by the communicators to refer to objects, both concrete and abstract.

The human body is the primary way of transferring non-verbal codes. Fiske (1982) lists ten codes and show what meaning they can carry:

1. *Orientation*. How a person is oriented towards another sends messages of the relation between them: there is a difference between facing someone and turning one's back to them.

2. *Appearance*. This includes aspects under one's control, like hair, clothes, make up, and jewelry, and such not so easy to control like height and weight.

3. *Head nods*. These are mainly used to control turn-taking in interaction.

4. *Facial expression*. This code can be subdivided into: eyebrow position, eye shape, mouth shape, and size of the nostrils. These combine to give facial expressions, following a well-defined set of rules. Facial expressions have many similarities across cultures.

5. *Gesture*. The primary sources of gestures are the hands and arms, but gestures with the feet and head are also important. Gestures are closely linked to speech, and are mostly not under conscious control. The specially developed sign languages of the deaf are not considered to be gestures in this meaning. They have full linguistic expressive power, while the gestures naturally occurring when people speak do not contain the properties of language. McNeill (1992) describes five types of gestures:

    - *Iconics*. These gestures bear a close formal relationship to the semantic content of speech, exhibiting the same act as referred to in speech, giving a further illustration of it.

    - *Metaphorics*. These gestures present a pictorial description of an abstract idea or concept. For instance, a cartoon could be shown as if it was a physical object in the space between a person's hands.

    - *Beats*. The hand moves along with the rhythm of speech. Beats provide crucial information about the meta-level of what is referred to in speech.

- *Cohesives.* These are repetitive gestures that signal continuity when the contents of speech needs to be interrupted and explained.
- *Deictics*. These gestures point to objects and events, both concrete and abstract.

6. *Body posture*. The way people sit, stand, or lie communicate attitude (friendliness, hostility, etc.) and mood (degree of tension or relaxation, etc.).
7. *Direction of gaze and eye-contact*. When, how often, and how long we meet other people's gaze is a way of sending important messages about relations. Eye-contact also controls start, turn-taking, and endings in interaction.
8. *Proximity*. How close we stand to a person we talk to leaves clues to which relation we have to the person. The distance around individuals can be divided into zones, marking intimate, personal, and official distances (varying much across cultures).
9. *Body contact*. Who we touch and where and when we touch someone convey important signals about relations. This code is the one with most variation across cultures.
10. *Non-verbal aspects of speech*. These is further divided into two categories:
    - Changing the meaning of an utterance by use of prosodic codes such as pitch and stress.
    - Conveying such information about the speaker as mood, personality, class, social position, and attitude towards the listener, by use of intonation, loudness, accent, speech defects, and speed.

The first eight of these codes rely solely on the eyes as media. *Body contact* is received with a combination of the senses of vision and touch. The *non-verbal aspects of speech* is exclusively auditory.

Non-verbal signals are sometimes called *extra-linguistic* or *non-linguistic* signals. In this paper, these terms are used to refer to the same set of signals.

2.1.4 Visual signals

Visual signals can now be defined, for purposes of this dissertation, as a subset of the non-verbal codes. From the list presented above, they include *orientation*, *appearance*, *head nods*, *facial expression*, *gesture*, *body posture*, and *direction of gaze and eye-contact*. The codes *proximity* and *body contact*, even though they involve vision, are not included in the concept *visual signals* as used in this dissertation, since they do not apply to the experiment situation used in the study.

2.2 COMPUTER-MEDIATED COMMUNICATION

As the word suggests, computer-mediated communication (CMC) is about communication mediated through computers. There is no agreement on a single definition of CMC, but December (1996) proposes the following as a definition of the field:

> ”Computer-Mediated Communication (CMC) is the process by which people create, exchange, and perceive information using networked telecommunications systems that facilitate encoding, transmitting, and decoding messages.”

He continues to characterize CMC as a research field:

> ”Studies of CMC can view this process from a variety of interdisciplinary theoretical perspectives by focusing on some combination of people, technology, processes, or effects. Some of these perspectives include the social, cognitive/psychological, linguistic, cultural, technical, or political aspects; and/or draw on fields such as human communication, rhetoric and composition, media studies, human-computer interaction, journalism, telecommunications, computer science, technical communication, or information studies.”

CMC can be either *synchronous* or *asynchronous*. Synchronous CMC is when individuals ”*at the same time* can carry out a task such as communicating, making a decision, planning a new initiative, structuring a proposal, writing a paper, or sketching a design” (Baecker, 1993, p. 3). Asynchronous systems ”*transcend the limitations of time* and *space*, allowing [...] communications and problem solving among groups of physically dispersed individuals” (Baecker, 1993, p. 3) An example of asynchronous communication is *electronic mail*.

In this paper, the terms ”networked telecommunications system”, ”computer system”, ”computerized system”, and sometimes in context, only ”system”, is used to mean the same thing, that is, technology through which communication between people is mediated.

2.3 COMPUTER SUPPORTED COOPERATIVE WORK

2.3.1 Definition of CSCW

Computer supported cooperative work (CSCW) is defined by Greenberg (1991) as:

> ”The study and theory of how people work together, and how the computer and related technologies affect group behavior.” (p. 133)

There is no agreed upon definition of CSCW, although Greenberg's definition presented here is representative. CSCW is a multidisciplinary field involving *computer science*, *psychology*, *sociology*, *anthropology*, *ethnography*, *management*, and *management information systems* (Greenberg, 1991). Although CSCW first appeared with its current name in 1984, its history goes back over forty years. Some related areas of CSCW in the field of education are *computer supported cooperative learning* and *distance education*.

2.3.2 Groupware

Groupware is a term denoting software developed to support the working of groups. Among the many kinds of groupware available, the most successful one has to date been *electronic mail* (Baecker, 1993). Other technologies are *computer conferencing*, *desktop videoconferencing*, *collaborative writing*, and *group decision support systems*.

Desktop videoconferencing, supplying both audio and image of the other conference participants, might at first glance appear to solve the problems concerning mediated communication. However, several problems with videoconferencing have been pointed out. Benford and Fahlén (1994) mention two problems: 1) participants cannot easily establish common spatial terms of reference, and 2) lack of embodiment: users have no representation within the common work area, which may, among other things, cause social problems. Isaacs and Tang (1993) found that in video communication, it can be difficult to notice peripheral cues, control the floor, have side conversations, point to things or manipulate real-world objects. Short *et. al.* (1976) describe how eye-gaze, which controls interaction, breaks down in video communication.

Software building on the principle called *What You See Is What I See* (WYSIWIS) also has severe drawbacks. In WYSIWIS, users see the same objects, in the same ways, in the same places, and users often have their own graphical pointer in the workspace. Benford & Fahlén (1993) argue that WYSIWIS inhibit collaboration by placing too many constraints on people's freedom to move and act as individuals.

To solve the problems associated with both video communication and WYSIWIS, researchers have proposed the idea of a common spatial frame within which users are *embodied* (Benford & Fahlén, 1994; Isaacs & Tang, 1993). That a user is *embodied* means that he or she is given a graphical representation in the common virtual space users share. This enables other users to see, for example, the identity, location, orientation and activity of the user.

A wisdom from the field of CSCW is that the technology itself should not contain any control mechanisms for interaction between humans. Instead, the technology should be flexible enough to support the natural emerging interaction behaviors of people collaborating (Greenberg, 1991; Benford & Fahlén, 1994).

2.3.3 Tasks in CSCW

The tasks involved in CSCW are of many different kinds. McGrath (1984) proposes a typology of tasks, based on the work of a number of researchers, which he calls the *task circumplex*. The model (illustrated in Figure 2.4) includes four general processes: to Generate (alternatives), to Choose (alternatives), to Negotiate, and to Execute. These are represented in the four quadrants of the model and indicate what the group is to do. Each quadrant can further be subdivided into processes as shown in Table 2.1. The sectors of each task type is positioned in the two dimensions of the model. Along the vertical axis, the element of cooperation in tasks increases upwards, and the element of conflict increases downwards. On the horizontal axis, the conceptual character of tasks increases to the left, while the behavioral character of tasks increase to the right.

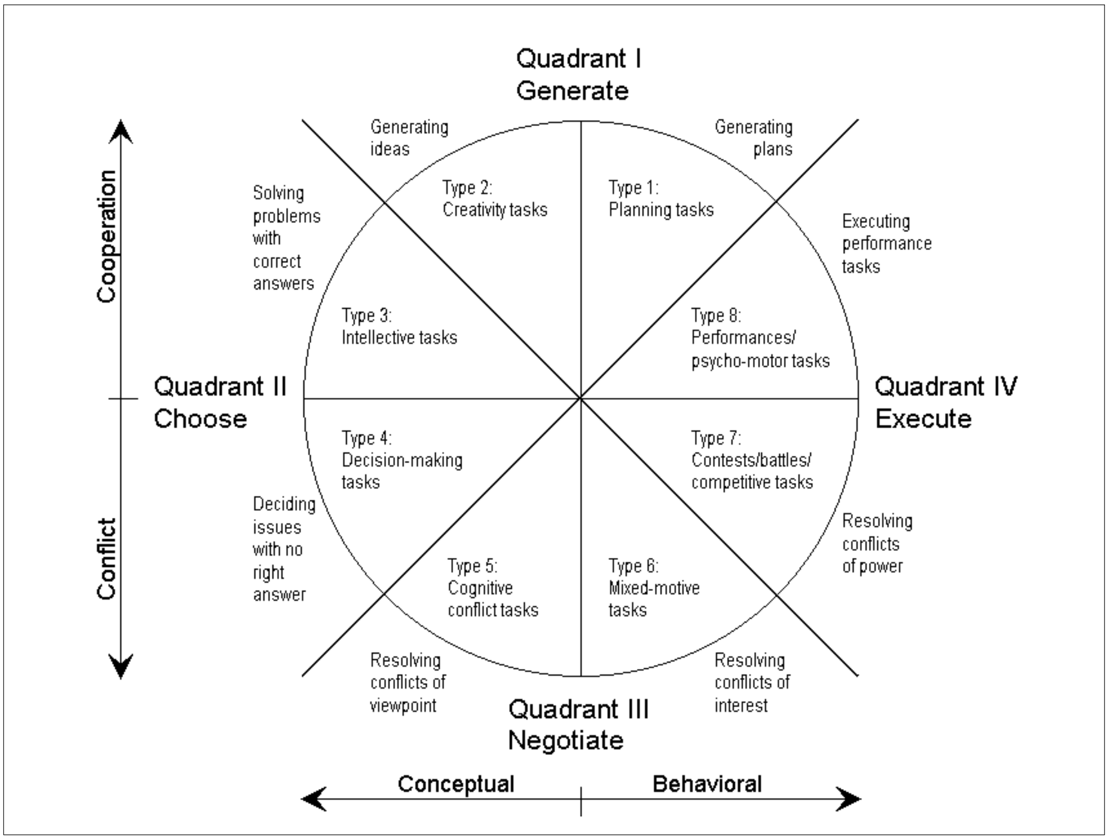

Figure 2.4. McGrath's circumplex model of group task types.

Quadrant I: GENERATE

Type 1.	Planning Tasks: Generating plans. Key notion: Action-oriented Plan.
Type 2.	Creativity Tasks: Generating ideas. Key notion: Creativity.

Quadrant II: CHOOSE

Type 3:	Intellective Tasks: Solving problems with a correct answer. Key notion: Correct Answer.
Type 4.	Decision-Making Tasks: Dealing with tasks for which the preferred or agreed upon answer is the correct one. Key notion: Preferred answer.

Quadrant III: NEGOTIATE

Type 5.	Cognitive Conflict Tasks. Resolving conflicts of viewpoint (not of interests). Key notion: Resolving policy conflicts.
Type 6.	Mixed-Motive Tasks: Resolving conflicts of motive-interest. Key notion: Resolving pay-off conflicts.

Quadrant IV: EXECUTE

Type 7.	Contests/Battles: Resolving conflicts of power; competing for victory. Key notion: Winning.
Type 8.	Performances: Psychomotor tasks performed against objective or absolute standards of excellence, e.g., many physical tasks; some sports events. Key notion: Excelling.

Table 2.1. Quadrants, task types, and key concepts for the McGrath Task Circumplex (adapted from McGrath, 1984).

The tasks used in the empirical study are described in Section 4.6 and related to McGrath's model in Section 4.6.8.

2.3.4 CSCW research

It is difficult to scientifically study the impact of technology on working groups. Greif (1988) points to three reasons. Firstly, users have different experiences and react to the user interface in different ways. Secondly, there is the difficulty of staging a realistic group-work setting in a laboratory and have volunteers to use the system to produce meaningful data. Thirdly, methods used to study interaction in single-user systems do not apply well to group support systems.

2.4 VIRTUAL REALITY

Aukstakalnis and Blatner (1992) defines *virtual reality*:

> ”Virtual Reality is a way for humans to visualize, manipulate and interact with computers and extremely complex data”

Virtual reality can be used for many purposes, for example, visualization and entertainment, but it does also supply a common spatial frame which makes it suitable for human collaboration (Benford & Fahlén, 1993).

DIVE (Distributed Interactive Virtual Environment) (Ståhl & Andersson, 1995) is a virtual reality system, developed at the Swedish Institute of Computer Science (SICS).

In DIVE, people who are geographically separate can cooperate, for example, in conferences and meetings. The participants meet in a simulated visual environment that in certain ways resemble reality. Objects in the environment can be manipulated, ”whiteboards” can be used by participants to give and share information by ”drawing” on them.

Users are represented by graphical 3D-bodies or icons. The icons resemble rough blocks, sometimes with a simple, fixed face. Figure 2.5 depicts a meeting in DIVE showing some of the most advanced user embodiments. Recently, there has been discussions to add to the DIVE environment a whole range of features of user embodiment, including facial expressions and gestures (Benford *et. al.*, 1995). This has not yet been implemented.

A system similar to DIVE is MASSIVE (Greenhalgh, 1996), also used for distributed meetings. How users are represented in MASSIVE can be seen in Figure 2.6.

No other system which incorporate a wider range of user embodiment than that provided by DIVE was found in the literature.

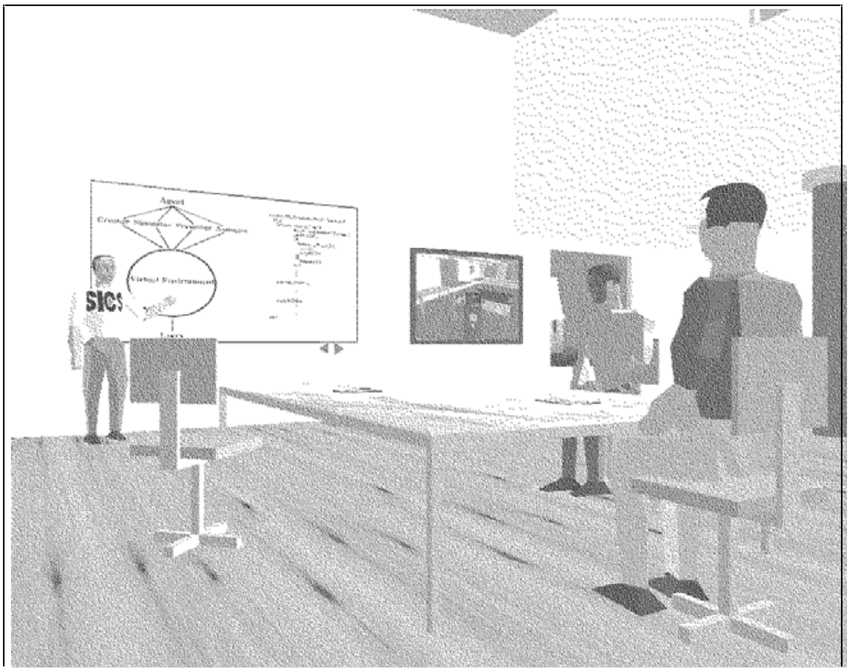

Figure 2.5. A distributed meeting in DIVE.

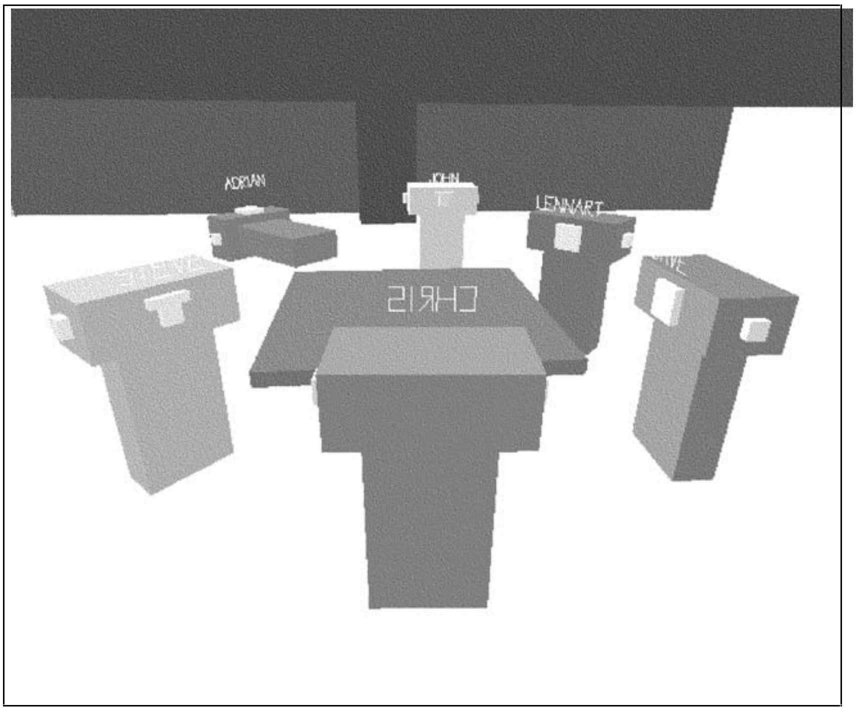

Figure 2.6. A distributed meeting in MASSIVE.

3 LITERATURE STUDY

Ju längre tillbaka du kan se
desto längre framåt ser du troligen.

Winston Churchill

This chapter provides a look at the literature of earlier studies of the role of visual signals in communication. In doing so, the question of whether the richer media always is the better is raised. Further, some anticipations of how lack of visual signals could affect task performance is presented. Related literature, on collaborative systems and human-computer interaction, is also discussed with focus on the issue of visual signals.

3.1 EARLIER STUDIES OF THE ROLE OF VISUAL SIGNALS IN COMMUNICATION

In a series of pioneering studies of human communication and collaboration, Chapanis and colleagues compared, among other media, face-to-face interaction and voice-only interaction (Chapanis *et. al.*, 1972; Ochsman & Chapanis, 1974, Chapanis, 1975). They found only slight differences in task performance between face-to-face and voice-only interaction.

Williams (1977) gives a review of research comparing face-to-face and mediated communication. The amount of impact of the visual channel was found to depend on the type of task. In cooperative tasks, no difference was found. However, conflictful tasks show differences between face-to-face and audio-only interaction. People cooperate more, and easier comes to an agreement in face-to-face interaction, compared with voice-only interaction.

Mostly focusing on the impact of visual signals on social interaction, Short *et. al.* (1976) discuss the effects of lack of visual signals. Mutual attention and responsiveness of participants in a conversation is not likely to be affected by lack of visual signals, since these functions, usually handled by eye-gaze, head nods, and gestures, can also be handled by the audio channel with ”yes”, ”ah ha”, ”ummm”, etc. Short *et. al.* consider this particularly likely since people in general are practiced in telephone conversation. Channel control—who is to speak and for how long—regulated with head nods and eye movements, is more likely to be affected by lack of visual signals. Short *et. al.* give some evidence of sex-related differences in perception of unintentional cues: ”In mixed-sex interactions, sex-related differences were found in the relative reliance on the visual and auditory channels for such signals, suggesting the possibility of sex differences in reaction to the removal of the visual channel” (Short *et. al.*, 1976, p. 160) Short *et. al.* concludes:

> ”In most cases, the functions of the non-verbal cues have been in some way related to forming, building or maintaining the relationship between the interactants. The absence of the visual channel reduces the possibilities for expression of socio-emotional material and decreases the information available about the other's self-image, attitudes, moods and reactions. So, regarding the medium as an information transmission system, the removal of the visual channel is likely to produce a serious disturbance of the affective interaction; one would expect the transmission of cognitive information to be relatively unaffected. Thus, the tasks which would be expected to be most sensitive to variation in the medium of communication are tasks in which the expression of emotion (and perception of this emotion) is an important part of the interaction, tasks with a great need for timing and coordination of the speaker's activity with the responses of the other, and finally, tasks on which there is the greatest need to manipulate others.” (Short *et. al.*, 1976, p. 161)

3.2 IS ALWAYS THE RICHER MEDIA THE BETTER?

Is always face-to-face interaction superior to other kinds of media? On the one hand, this seems reasonable, since a richer media provide more information. This conclusion is supported in the literature. On the other hand, one could imagine, that people somehow could be distracted by visual signals during interaction, but this conclusion is not supported in the literature.

However, there are some arguments against the ”as rich media as possible” hypothesis. Discussing the area of computer-supported design, Kvan (1994) notes:

> ”While it is true that communication in some cultures relies heavily upon gesticulation, effective communication can be achieved without these ancillary references. [...] This hypothesis appears to be supported by the telecommunications market - for example, AT&T has recently supplemented their videophone offerings with the 'Picasso' still video system, reflecting a market void not filled by synchronous video communications systems.” (Kvan, 1994, section *Communication for Collaborative Design*, paragraph 3)

People react differently to face-to-face compared with more technology-based media. Isaacs *et. al.* (1995) compared users' perceptions of talks in face-to-face versus distributed settings (live broadcast of video, audio, and slides). The study showed that audiences preferred to watch talks in the distributed setting (it should be noted however that speakers preferred to *give* talks in the local setting). The study does not give any measure of how, other than through user's perception, the distributed talks actually were better.

So, even though there are some objections, in summary, the experimental literature supports the conclusion that no other media is superior to face-to-face interaction in terms of measured performance.

3.3 HOW COULD VISUAL SIGNALS FACILITATE TASK PERFORMANCE?

Task performance could be facilitated by visual signals in at least three ways: control of interaction, use of illustrative gestures, and enhancing understanding.

First, visual signals help to control the interaction. Without efficient handling of turn-taking, conversation would take longer. Lack of visual signals would also mean a longer wait before feedback is given to the speaker, because feedback would have to be sent in the auditory channel, which is busy during speech. This means that tasks that have a time-limit could be affected so that performance is lowered when visual signals are absent. However, as described by Short *et. al.* (1976) these visual cues could be taken over by voice cues, such as ”yes” and ”ummm”.

Second, using visual signals, explicit illustrative gestures could be made illustrating certain aspects of the task. In a spatial task, shapes and spatial relations could be communicated using gestures. This would facilitate the spatial task since more exact information concerning the task could be conveyed.

Third, gestures could be used to enhance understanding. In tasks including elements of retelling a sequence of events, gestures have several functions of switching between discourse and meta-level of discourse, for example, *beats* and *cohesives* (see p. 14) (McNeill, 1992).

3.4 RELATED LITERATURE

3.4.1 Systems for collaboration

Towell & Towell (1995) reports from a study of conferencing on the Internet implemented in a Multi User Dungeon (MUD) system as synchronous, text-only interaction.

They are enthusiastic about MUDs as a medium for conferences, and do *not* consider effects of lack of visual signals.

Tang (1991) reports on findings from observational studies of collaborative work. He mainly considers the deictic function of gestures, that is, the pointing at other persons or places in the workspace. Tang notes that participants use gesture even when all visual contact is absent and the other persons can't see them.

Tatar, Foster and Bobrow (1991) report from experiences of some time's use of a real system for collaboration, Cognoter. In the beginning, an inappropriate model of conversation which they call ”parcel-post” model was used. This model is a unidirectional model of communication which resulted in frequent breakdowns of communication. When the ”parcel-post” model was replaced by an interactive model, communication was improved.

In summarizing the experiences from the study of a distributed research group, Olson and Bly (1991) argue that CMC may not necessarily make social interaction ”less” but perhaps only ”new”.

3.4.2 HCI, interfaces, and intelligent agents

When examining recent literature on human-computer interaction (HCI), one gets the impression that a minor revolution is about to take place. It is the revolution of graphical, conversational, human-like interfaces. This is the next step in the evolution of interfaces after command language and direct manipulation interfaces: conversational interfaces.

One key notion here is that of *intelligent agent*. An intelligent agent helps the user interact with the system. Research on intelligent agents are carried out at Xerox PARC, MIT's Media Lab, Microsoft, and by researchers at other companies and universities (see Ibister & Layton, 1995).

Recent research at Microsoft is concerned with implementing a conversational interface to a computer application which plays compact discs (Ball *et. al.*, 1996). The system takes user speech as input and an animated speaking parrot carries a dialog with the user. Microsoft also sells a home management system called Microsoft Bob (Microsoft, 1996) which includes the intelligent agent metaphor.

Quek (1994; 1995) reports on advances in non-verbal vision-based interfaces including human face recognition, facial expression interpretation (also described in detail by Ekman *et. al.* (1992)), lip reading, head orientation detection, eye gaze tracking, three-dimensional finger pointing, hand tracking, hand gesture interpretation, and body pose tracking.

Huang & Orchard (1992) point out the importance of visual input when they describe man-machine interaction in the 21st century:

> ”We would like these interactions to mimic the ways we interact with objects and people in the physical world, and, to achieve this goal, we believe that it is essential to consider the exchange of video data into and out of the computer” (p. 428)

Bates (1994) argues that emotion is crucial to include in conversational interfaces. Since expression of emotion is made non-verbally, including by visual means, Bates can be said to argue for visual signals in the interface.

Graphical simulation of human conversation, including verbal and non-verbal aspects of interaction have been implemented by Cassell and colleagues (Cassell *et. al.*, 1994a; Cassell *et. al.*, 1994b).

There are several sources for these developments within the area of HCI. Technology is furnished with more and more computational and graphical power. There have also been advances in computer vision. Another reason, which connects the fields of HCI and CSCW, is that more computer systems are becoming multi user systems.

A computerized system for human collaboration can accomplish nothing if its interface is not powerful enough to handle the input and output of signals needed for human interaction. Therefore the development of CMC systems depends much on the advances in interface design.

4 METHOD

Though this be madness, yet there is method in 't.

William Shakespeare, *Hamlet*

The method of the experiment is the focus of this chapter. First, an overview of the experiment is given. The hypothesis and the independent and dependent variables are described. Further, the method, tasks and measurements are described and motivated. The experimental procedure is described in detail, followed by how the data were analyzed. Last in the chapter, two pilot studies that were carried out is described.

4.1 OVERVIEW OF THE EXPERIMENT

The basic chain of logic behind the experiment (details omitted but given later in this chapter) is as follows:

The general question is: Does the use of visual signals improve collaboration in computer-mediated communication (CMC)?

The general question is tested by having two persons communicate under conditions of CMC in two different situations: one *with* visual signals, and one *without*.

More specifically, this is done by having two subjects at a time cooperate on two selected tasks (*describing and drawing a photograph*, and *re-telling a narrative*) in the two situations. By then comparing the results one can see whether the tasks are more successfully solved in the situation with visual signals or not. The tasks were selected as representative of tasks occurring in CMC: *the drawing* is spatial and *the narrative* is symbolic.

At a lower level of description, this translates the following: Do subjects draw pictures more similar to the photograph (spatial task), and do subjects answer the questions about the narrative better (symbolic task), in the situation *with* visual signals, than in the situation *without* visual signals?

In terms of statistical computation, this is equivalent to: does the two groups' means (on both tasks) differ significantly, as calculated by a t-test?

If so, then the chain can be followed upward again, finally giving an affirmative answer to the general question.

4.2 HYPOTHESIS

The hypothesis is that visual signals improve collaborative task performance.

In more detailed terms, this equals that the means for the performance of the two tasks should be significantly higher in the experiment situation with visual signals present.

4.3 VARIABLES

4.3.1 Independent and dependent variables

The independent variable that is manipulated in this study is the presence or absence of visual signals. To be able to claim that the effects observed really are caused by the independent variable, one has to create an experiment such that this is *the only thing* that varies. In this experiment, this means that everything has to be held constant except that in one situation, visual signals are permitted, and in the other, visual signals are prevented. What this means is that the communication in the second experiment situation—without visual signals—has to be a *proper subset* of the first experiment situation—with visual signals. This relationship is shown in Figure 4.1.

If one were to have two situations where one was *not* a proper subset of the other (in terms of communication signals present), causal reasoning would be made more difficult. For instance, if one experiment situation was face-to-face interaction with visual signals, and the other one was communication via electronic mail, one could not say what caused the effects obtained in such a study. Was it the lack of visual signals in electronic mail? Or perhaps it was the fact that electronic mail is not interactive in real time (asynchronous), or that voice is unavailable in electronic mail? The proper subset relation does not hold between these two situations, and therefore no causal connection can be established between availability of visual signals and the observed phenomena.

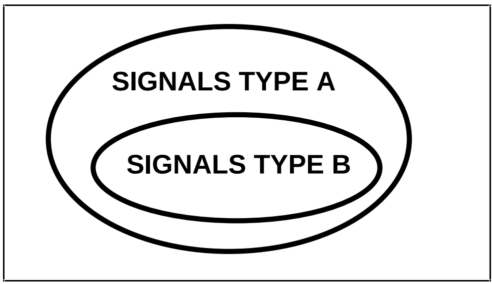

Figure 4.1. Signals of type B is encompassed by signals of type A.

When excluding visual signals in the second experiment situation, several different kinds of visual signals are excluded. As noted, these will not be treated separately in this study, but will rather be handled as a single group labeled *visual signals*. An analysis of the various types of signals in the experiment situation is made in Section 6.1.1.

4.4 CHOICE OF METHOD

The study situation chosen is a laboratory study. Studying people in a realistic task situation, and being able to control the variables of interest, was considered a good choice. In order to obtain measures of task performance, a quantitative analysis was chosen.

To test the hypothesis, a number of alternative methods could have been used.

People could have been observed in natural situations where they use CMC. A field study of this kind has the advantages that the conditions are realistic and the results obtained would be of high ecological validity. However, a field study has several drawbacks. It is difficult to carry out. Where should people collaborating using CMC be found, and how should they be observed? Even if it would be possible, it would take more time than was available for this study to get the right experiment conditions to show up.

Also, interviews could be made with people using CMC. They could be asked whether they find that visual signals improve collaboration. However interesting answers could be obtained from such an approach, the results are likely to be unreliable. People might *think* things are in a way other than they really are. This would give results of low validity.

4.5 WHY NOT USE A COMPUTER IN A CMC STUDY?

One possibility would be to implement facial expressions and gestures in a system like DIVE and carry out the test in that environment. This was not done for the following four reasons:

First of all, an investigation of whether it would be technically possible to do such an implementation would have to be conducted. Then, even if it turned out to be technically feasible, an implementation would require more resources than were available for this study.

Secondly, a more important drawback would be that if the hypothesis was tested in a certain system, one will not know whether the results are due to some quirks in the implementation or to the user interface of that system. It will be uncertain whether the results could be applied to a different system. By abstracting the actual system and performing a test in a technically simpler way, one would be able to get results with higher generality. A ”real” situation is more general, and from any results obtained here one can draw the conclusion that they will hold in any computerized system for collaboration, fulfilling some minimal set of criteria (see Chapter 7, *Conclusions*).

Thirdly, using a technology-neutral approach adheres to the idea within CSCW that technology of group interaction should not contain any explicit mechanisms for control of interaction (Greenberg, 1991; Benford & Fahlén, 1994).

Finally, there is the advantage of not having to consider individuals' differing reactions towards the interface, as pointed out by Grief (1988).

There appears to be no evidence in the literature that group interaction is different when computerized. Valacich, Dennis, and Nunamaker (1991) argue that it *might* be different:

> ”Much of the theory for small group interaction [...] [should] be re-examined in a computer-supported context, as the use of computer technology may profoundly affect group and organizational process, structure and outcomes” (p. 262)

However, there seems to be no real support for this claim. Generalization from collaboration in a technology-neutral to a computerized setting thus seems reasonable.

4.6 CHOICE OF TASKS

4.6.1 Tasks used in earlier studies

In a review of research on comparisons between face-to-face and mediated group communication, Williams (1977) gives some examples of tasks used in earlier studies:

- Finding on a map the nearest physician to a given address.
- Assembling a trash-can carrier from instructions.
- Identifying newspaper articles referring to a given topic.
- Identifying a light socket that matches another one.
- Discovering the factors that distinguish profitable and nonprofitable hypothetical factories.
- Deciding how to deal with an older worker who is holding up a production line.
- Scheduling on a hypothetical college timetable.

Williams differentiates between *cooperative tasks* and *conflictful tasks*. In cooperative tasks, participants work towards a shared goal, while in conflictful tasks, participants have different goals, interests, or viewpoints, and must negotiate to reach a consensus.

Another study conducted by Valacich, Dennis, and Nunamaker (1991) used *idea generation* and *decision making* as tasks.

4.6.2 General task criteria

Possible tasks to include in the study should fulfill the following criteria:

- Tasks should be representative of tasks normally performed in a real collaborative situation. Having realistic tasks gives high ecological validity.
- Tasks should be collaborative, so that they are only solvable if subjects work together. If a subject could carry out the task himself/herself, effects of collaboration—which is the topic of interest in the study—would be lost. This means that there must be some material which collaborating subjects are not allowed to show each other.
- Tasks should not be too easy or too difficult. All or near all subjects should be able to complete tasks, but not with top performance. This is to ensure a spread of values on the variables measured.

- Task should not require any special skill or specialized knowledge. This ensures that subjects with special skills are not facilitated.
- Tasks should be directed toward neither men nor women, so that neither sex is facilitated.
- Subjects should not have any pre-knowledge of the tasks. Tasks should not have appeared anywhere else before. Pre-knowledge of the tasks would facilitate subjects who had it.
- Tasks should be straightforward to evaluate, and measurements of task performance should be accurate and reliable.
- The tasks must be interactive. Both speaker and listener must be able to adapt to the situation (see for example Nilsson, 1993). For instance, having a pre-recorded video of a person speaking is not desirable, since it makes communication artificially non-interactive. Another non-desirable situation would be the combination of the two experiment situations (visual signals versus no visual signals), for example, having one speaker talk with one subject behind a screen (visual signals absent) and to another subject with no screen (visual signals present) at the same time. The problem with this setup is the question of who the speaker ”really” is communicating with: the hidden person or the visible person? The behavior of the speaker depends on whether he or she has visual contact or not.

4.6.3 Tasks and cognitive processing

The selection of tasks should represent the types of cognitive processing humans carry out. Gärdenfors (1992) claims that the two main models of knowledge representation are the *spatial* model and the *symbolic* model. With this in mind, one *spatial* task and one *symbolic* task were developed. If the results would hold over both tasks, this would give firmer evidence than if only one kind of cognitive task had been used.

Having two tasks makes it possible to see whether exclusion of visual signals has different effects on different types of tasks (spatial versus symbolic). If results would show that tasks are solved unevenly, it would also be possible to infer what kind of information the visual signals convey. One guess is that visual signals ought to facilitate the spatial task in particular, so that task performance is higher for that task compared with the symbolic task. If this is true, it would indicate that visual signals convey spatial information.

4.6.4 The spatial task

In choosing the spatial task, several alternatives were examined. Before describing the spatial task used in the study, three alternative tasks that were considered will be described.

One possible spatial task would be having one subject instruct another subject to fold a piece of paper in a special way. Advantages are that it requires spatial thinking, is quite easy to handle practically, and that the subjects' solutions could be saved for later evaluation. But the disadvantages outweigh. The task of folding paper could be said to be a quite unrealistic task, being too artificial. It is also hard to find any folding procedure people have not been exposed to before that is not too easy or too difficult. Finally, the task is difficult to evaluate; it is unsure how to check when a subject has completed the folding, and how should partial solutions be judged?

A second alternative considered is an assembly task. This was used by Chapanis (1975). Here, one subject could instruct another to build a specified model using, for instance, LEGO. LEGO is suitable for building since it can be handled by nearly everyone and is easy to assemble and disassemble. The instructing subject could have as his or her source a ready-built model, a step-by-step instruction sheet, or a picture of a completed model. This task is realistic and represents such real-world tasks as managing repairs in hazardous environments (e.g. in space or under water) involving geographically separate persons using remote communication. However, this task was not selected for the study because of a number of disadvantages. If subjects' models are saved and evaluated later, this requires a very large supply of material (LEGO blocks). To avoid this, models could be photographed, but then they need to be shot from three different angles. Restricting the model to two-dimensions makes photographing easier (only one photograph needed), but misses the idea of building a model; it should be three-dimensional. Finally, the task is also difficult to evaluate; when have subjects finished the task, and how are partial solutions evaluated?

A third possibility considered, but not used in the final study, is in some way an abstraction of the LEGO assembly task. One subject could guide another subject in drawing some simple geometrical shapes. The first subject could have a picture with geometrical shapes, and the second subject could have a blank paper and pencil, trying to draw the first subject's picture from his or her description. This design was abandoned when it was discovered in the pilot study that subjects, in the experiment situation where visual signals are allowed, used their fingers to in a detailed way draw the geometrical shapes in the air in front of the other subject. Explicitly instructing subjects not to do this was judged to impose unnatural restrictions on the task. Using this task—without the restrictions mentioned—would mean that the study would confirm the hypothesis in a trivial manner. The outcome would be known beforehand, and then there would be no need to carry out the study.

The task used in the study was a drawing task, like the one described above, but with the difference that the motive was a photograph (depicted in Appendix A). One subject has access to the photograph which he or she, in any way preferred, describes to the other subject. The second subject makes a drawing of the photograph using a blank paper, pencil, and eraser as aids. The disadvantages with the previous drawing task mentioned is overcome because the photograph is sufficiently complex not to make subjects gain anything by ”drawing in the air”. The drawing is saved and later judged by separate judges on how similar it is to the photograph. Five minutes was chosen as a time limit on this task since pilot studies showed that sufficiently rich drawings were produced in that time. The task fulfills the criteria listed above, with the possible exception that drawing skill can be considered a special skill possessed by certain subjects. However, this ”injustice” is present in subjects tested in both experiment situations and should therefore have no impact on the results. The results need give only the relative difference between the two experiment situations, not in an absolute way show how good subjects are at drawing.

4.6.5 The symbolic task

Language is the kind of symbolic representation people use most (Gärdenfors, 1992), so consequently a symbolic task should include language. People often tell stories; things that happened to them or to someone else. Stories even play a role in corporate decision making (Norman, 1993). As the symbolic task, the task of retelling a story was chosen. One subject reads a story, followed by a period of verbally sharing this story with the other subject. The second subject then answers some questions on the contents of the story. By using multiple-choice questions, a measure of the performance of this task could easily be obtained. The task fulfills the task criteria listed above.

This task is complex and involves many cognitive abilities. One could ask what the performance on this task shows. Is it a measure of how well subjects *remember?* Or is it dependent on how well the story is *communicated* by the first subject? These questions need not be answered here, since what is tested is the whole task, the whole situation. The various cognitive abilities used will be involved in a real situation where this task is performed as well. Effects that occur in this task during the study (whatever they can be attributed to) will also be present in a real collaborative situation.

When answering the multiple-choice questions, subjects were instructed to guess if they did not know the answer to a question. The reason for this is twofold. First, some subjects are likely to make guesses anyway, without explicit instructions. Second, in a subject's guess there is likely to be some degree of knowledge, even if they are unaware of it.

4.6.6 Restrictions in tasks

Tasks which needed as few restrictions as possible with respect to experimenter control were selected, so that natural mechanisms of collaboration and communication could emerge. This gives a higher ecological validity, compared with if subjects were to be more restricted to perform unnaturally restrained tasks.

4.6.7 Subjects' motivation to solve tasks

Why would subjects want to perform the tasks? What makes them continue and not give up after a short amount of time?

The main motivation lies in the social setting. They have agreed to be part of an experiment, so they are expected to complete their part of the responsibility towards to experimenter.

Another reason of social nature is that subjects may also be concerned with providing a good image of themselves, both to the other subject in the pair, and to the experimenter.

There might also be intrinsic motivation in the tasks. Subject might want to do as well as possible because it is rewarding in itself.

Finally, one could argue that motivation is not often even found in real situations concerning collaborative tasks, so why should it be considered in the laboratory? If subjects were heavily motivated to perform well (by use of money or other rewards), the situation in the laboratory might turn into a unnatural one, compared with a realistic situation.

4.6.8 McGrath's circumplex model of tasks

Turning to McGrath's circumplex model of tasks (described in Chapter 2 on page 20) it can be seen that neither of the two tasks used in the experiment can be classified easily. The spatial task can be categorized as belonging in Quadrant IV: ”Execute: Type 8.

Performances” because the goal is to produce as good a drawing as possible. But the spatial task also has features of Quadrant I: ”Generate: Type 2. Creativity Tasks” since the subject who does the drawing has to be creative. The symbolic task would belong in Quadrant II: ”Choose: Type 3. Intellective Tasks” because it has well-defined answers. However, the task as a whole—reading, retelling, and remembering a story—is not a pure intellectual task as given by the task circumplex model.

4.7 CHOICE OF MEASUREMENTS

To decide if task collaboration is better in one situation than in another, one needs a way of measuring *how well* a task is solved.

The dependent variables in this study are *picture similarity*, judged on a scale from 1 to 7 (for the spatial task) and *narrative recapitulation*, from 0 to 8 (for the symbolic task).

It would have been possible to study the number of errors subjects made while working on a task. However, this is difficult to measure, both technically—being able to spot actions performed at high speed—and conceptually: what constitutes an error? These shortcomings make a study of errors unreliable. Adding video recording would solve the technical difficulties (but not the conceptual ones) and would enable a detailed study of errors. However, it was judged that the large amount of time needed to analyze video records did not match up with what could be obtained from such an approach.

Another possible way to measure quality of collaboration is to record task completion times. This measure is both technically and conceptually simple. Yet, task completion times was not chosen as a measure since it was found difficult to combine with a suitable task where it would really measure task performance. However, task completion times were recorded for control purposes. For example, if a pair would have highly unusual values on the dependent variables (e.g. a score of zero on the narrative recapitulation), the task completion time could be checked to see if it was extremely short, thereby providing an explanation to the unusual value.

One further possibility not utilized is to let subjects make their own records as they perform a task. This could be realized by letting subjects put indications of how well they think they are doing on a scale on different time intervals. This was not used for two reasons. First, it was judged to disturb the subjects too much. Second, it would only provide the *perceived* task performance, which could be totally different from the *actual* task performance. It would thus not be a valid measure.

4.8 CHOICE OF MEASUREMENT TECHNIQUES

Having settled on measuring *picture similarity* (for the spatial task) and *narrative recapitulation* (for the symbolic task), these measurements were realized in the following way.

Picture similarity was measured by letting judges rate on a 7-point scale how similar they thought the pictures were to the photograph.

Narrative recapitulation was measured by having subjects answer a questionnaire with knowledge control questions on the narrative. A score was then calculated by counting the number of correctly answered questions on the questionnaire (see Appendix D for the questionnaire and the correct answers).

Task completion times were measured in minutes and seconds using a standard digital wristwatch. The accuracy of the watch was ensured by checking it against another time source. The precision on the watch was higher than the precision needed for the measurement.

4.9 SUBJECTS

Eighty subjects (40 pairs) participated in the study for pay[1]. Subjects' ages varied from 19 to 64 years, with a median of 25 years. Of the subjects, 50 were male (62.5%), and 30 were female (37.5%). Subjects were recruited at the University of Skövde or at their home (in student apartment areas), with the result that most subjects were either students or staff at the University. Subjects were randomly assigned to the two experiment situations, as well as to the roles in the collaboration pair.

No subject was included in more than one pair. All subjects understood, spoke, and read Swedish fluently (this also gives some guarantee that subjects within a pair have roughly the same social conversational behavior, through their similar cultural background). Subjects had reasonable good vision (visual aids were used if needed). None of the subjects had any problems with moving their face and upper body. Subjects did not have any pre-knowledge on the task they would solve (see discussion in section 4.6, *Choice of tasks*). These conditions of language and physical factors of subjects were checked at recruitment time.

Subjects had roughly the same social relationship to each other in all pairs; all pairs consisted of subjects who did not know each other (this was ensured at recruitment time). If some subjects would be close friends, this might affect the study in the way that these pairs would more easily communicate and be able to interpret the other part's visual signals.

Further subject variables are sex and age. Communication between two people of different sex or of a high difference in age is likely to be different from when these conditions do not hold. These subject variables were not controlled in the study, since it was considered to be something occurring in natural situations as well. By not limiting composition of groups, a result more valid to generalize should be obtained. Age and sex

[1] Subjects who were students or staff at the University were not paid. Only subjects recruited outside of the University were paid, due to department policy.

were still recorded for each subject, to ensure that the values of the variables of the subjects tested under the two situations were balanced.

Other subject variables not controlled, but balanced between the two groups, are cultural and individual differences in communication, especially use of non-verbal codes. If the two subjects speak the same language, but come from cultures with differing use of non-verbal codes, misunderstanding is likely to occur, making communication more difficult. The same disruption of communication could happen if subjects had large individual differences in the use of non-verbal codes.

For *each* subject, the following subject variables were recorded (using a questionnaire which subjects were asked to fill in):

- Age (year of birth)
- Sex (male or female)
- Occupation (job title or ”student”)
- How often subject attends conferences (multiple choice: 1...3)
- How often subject attends meetings (multiple choice: 1...4)
- Subject's preference of expression: *writing, speech* or *no preference*

Recorded for each *pair* of subjects:

- Subjects acquaintance (*acquainted* or *not acquainted*)
- Tasks completion times (one for the spatial task and three for the symbolic task, in minutes and seconds). These were not recorded as a measure, but in order to be able to spot possible ”abnormal cases” during analysis later, that is, extremely short or too long completion times (longer than the time limit).

The two groups of subjects tested under the two experiment situations had balanced distributions of values of all recorded subject variables compared with each other.

The subject profile obtained makes generalization from the subjects used in the study to the average Internet user appear reasonable. A typical user of the Internet was judged to be a reasonably good representative of a person who would use CMC in collaboration. Statistics from The Fourth WWW User Survey in October 1995 by the GVU Center (Pitkow & Kehoe, 1995) reveal that the average World Wide Web user is 32.7 years old, and the distribution of female and male users are 29.3% and 70.7% respectively.

The purpose of having *many* subjects perform *few* tasks is because generalization over subjects were considered more important than over tasks. If few subjects would be used, who perform a wide range of tasks, extreme values on one or more influencing subject variables would affect *all* tasks carried out by that subject. Since subject variables such as experience and individual differences in communication has an impact on task performance, it is preferred to have a large sample of subjects.

4.10 EXPERIMENT MATERIAL

Effort was made to ensure that subjects' material should be easily understood: it should be visibly and conceptually clear. All papers used were of A4 size and white. A standard pencil and eraser was used.

The screens covering each subject's material were made of cardboard and sufficiently high to hide the material from the other subject and low enough to allow visual signals such as facial expressions, gestures, and body posture to be easily observed by the other subject.

The screen preventing visual signals in one of the experiment situations was made of cloth. No check was made on the acoustic transparency of the cloth since sound could still pass on the sides as well as under it.

The photograph used for the spatial task (available in Appendix A) was taken from Lindström (1994). It was photocopied into black and white, because colors were judged to unnecessarily complicate the task. The criteria for the photograph was that it may not be of a widely known motive, and should contain easily recognizable objects. The photograph was fastened to a screen.

The narrative (available in Appendix C) was the opening chapter from Werner (1972). The criteria for the narrative was that it should be easy to read, and not widely known (stories by famous authors would not do). The narrative was retyped on two white A4 stapled papers.

4.11 THE LABORATORY

The experiment was carried out in an ordinary room at the University of Skövde. The laboratory excluded noises (being closed) and was free from disturbances (people were instructed not to disturb by a sign on the door outside. The room's glass wall was covered to prevent spectators from looking in and subjects from looking out). Further, the laboratory had normal room conditions, in terms of lightning (window curtains were drawn and ceiling lights were used) and temperature. Following recommendations of Valacich *et. al.* (1991), conditions of air conditioning and ergonomics were also ensured to be fulfilled.

4.12 DESIGN

The laboratory setting is showed in Figure 4.2.

In the experiment situation *with* visual signals, subject 1 and subject 2 sat on opposing sides of a table, facing each other. Visual signals such as gestures and facial expressions were visible to both subjects. Subject 2 had the task of replicating a picture which was visible to subject 1 only (spatial task), and re-telling a narrative (symbolic task). Subjects could communicate freely. Each subject could only see his or her own material.

The experiment situation *without* visual signals was the same as the situation described, except that a screen was located between the subjects, shielding any visual information.

The experimenter sat beside the two subjects, seeing both subjects from the side.

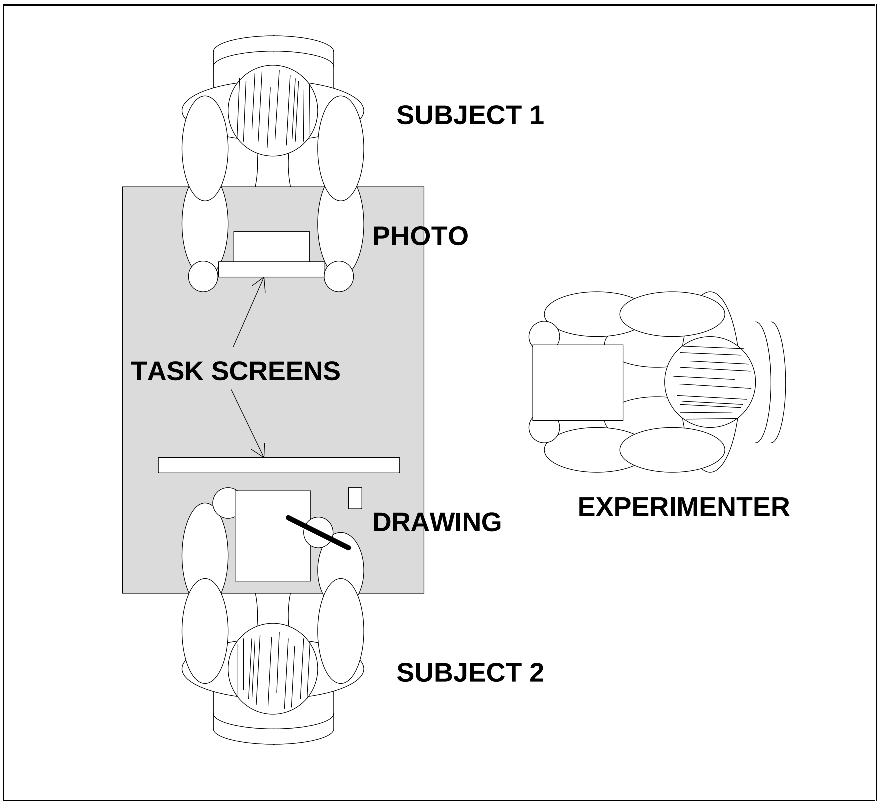

Figure 4.2. Basic laboratory setting. Subjects sit facing each other, cooperating on a task. In the situation with visual signals (depticted), two low task screens ensures that subjects see only their own material. In the situation without visual signals, a screen prevents all visual contact between the subjects.

4.13 PROCEDURE

The procedure for the two experiment situations was exactly the same.

For each pair of subjects, the spatial task was performed first, directly followed by the symbolic task. The order of tasks was arbitrarily chosen, and was kept the same throughout the study in order to keep things constant between the two experiment situations. No reasons for varying the order of tasks were seen.

Note that subjects were not at any time instructed to use non-verbal communication.

The experimenter remained neutral with respect to gestures, facial expressions, and body posture. Direction of gaze, that is, looking at subjects when addressing them, was used by the experimenter in order to be able to give appropriate instructions to subjects.

Notes were taken by the experimenter during the experiment in case of any unusual events.

Now follows a chronological account of the experimental procedure.

4.13.1 Before the tasks

Two subjects were shown into the laboratory. Subjects were asked if they were acquainted with each other:[2]

”KÄNNER NI VARANDRA?” / *”DO YOU KNOW EACH OTHER?”*

The answer was noted. Subjects were shown to their seats. Subject 1 (who shall see the photograph) was seated at one end of the table, and subject 2 (who will make the drawing) was seated at the other. The experimenter was seated beside the table, facing the subjects. Subjects were instructed:

”NI SKA NU UTFÖRA TVÅ SMÅ UPPGIFTER TILLSAMMANS.” / *”YOU WILL NOW CARRY OUT TWO TASKS TOGETHER.”*

---

[2] As the experiment was conducted in Swedish, the experimenter's instructions will here be written in Swedish, directly followed by an English translation in italics.

4.13.2 Spatial task

1. A blank paper, pencil, and eraser was placed on the table in front of subject 1.

2. Subjects were informed verbally that they will solve a task together:

   ”DEN FÖRSTA UPPGIFTEN GÅR UT PÅ ATT DU [TURNING TO SUBJECT 1] SKA BESKRIVA EN BILD PÅ ETT FOTOGRAFI DU HAR DÄR BAKOM SKÄRMEN, OCH DU [TURNING TO SUBJECT 2] SKA FÖRSÖKA ÅTERSKAPA DETTA FOTO SÅ BRA DU KAN. DU [TURNING TO SUBJECT 1] FÅR BESKRIVA DET HUR DU VILL, OCH DU [TURNING TO SUBJECT 2] FÅR FRÅGA VAD DU VILL MEDAN NI HÅLLER PÅ. DEN ENDA RESTRIKTIONEN ÄR ATT NI INTE FÅR SE VARANDRAS MATERIAL. NI HAR FEM MINUTER PÅ ER ATT GÖRA DETTA. JAG SÄGER TILL NÄR DET ÄR EN MINUT KVAR. HAR NI NÅGRA FRÅGOR?” / *”THE FIRST TASK CONSISTS OF YOU [TURNING TO SUBJECT 1] DESCRIBING A PICTURE OF A PHOTOGRAPH YOU HAVE BEHIND THAT SCREEN. YOU [TURNING TO SUBJECT 2] SHOULD TRY TO RECREATE THIS PHOTO AS WELL AS YOU CAN. YOU [TURNING TO SUBJECT 1] CAN DESCRIBE IT IN ANY WAY YOU WANT, AND YOU [TURNING TO SUBJECT 2] CAN ASK ANYTHING WHILE YOU ARE WORKING. THE ONLY RESTRICTION IS THAT YOU CAN'T SEE EACH OTHER'S MATERIAL. YOU HAVE FIVE MINUTES TO DO THIS. I WILL NOTIFY YOU WHEN THERE IS ONE MINUTE LEFT. DO YOU HAVE ANY QUESTIONS?”*

3. Questions were answered by the experimenter, if asked.

4. The screen with the photograph was turned up so that it could be seen by subject 1.

5. The timer was started. Starting instructions were given to the subjects:

   ”NU KAN NI BÖRJA.” / *”YOU CAN START NOW.”*

   Subjects were monitored during the experiment so that they did not see each other's material.

6. Subjects were informed when one minute of the time was left:

   ”NU ÄR DET EN MINUT KVAR.” / *”NOW IT'S ONE MINUTE LEFT.”*

7. The task ended and the timer was stopped. Tasks could end in one of two ways:

   (i) The time limit of 5 minutes was reached

   (ii) Subjects said they were finished ahead of time

   In the first case, the subjects were told that the time was up:

   ”NU ÄR TIDEN UTE.” / *”TIME IS UP.”*

8. The drawing was removed and the photograph was put face down on the table.

9. Subjects were informed:

   ”DET VAR DEN FÖRSTA UPPGIFTEN.” / *”THAT COMPLETES THE FIRST TASK.”*

10. On the back of the drawing, some details were noted in order to be able to identify the drawing later: date, time, the experiment situation, and the task completion time were noted. The experiment part of the spatial task was now finished.

4.13.3 Symbolic task

1. Subjects were seated in the same way as before the spatial task.

2. Subject 1 was informed verbally:

   ”DIN UPPGIFT ÄR NU ATT TYST LÄSA EN KORT BERÄTTELSE, OCH FÖRSÖKA KOMMA IHÅG DEN. DU SKA SEDAN FÖRSÖKA ÅTERBERÄTTA HANDLINGEN TILL [HONOM/HENNE]. DU FÅR INTE HA NÅGRA HJÄLPMEDEL. DU HAR FEM MINUTER PÅ DIG. JAG SÄGER TILL NÄR DET ÄR EN MINUT KVAR. DU KAN SÄGA TILL OM DU ÄR KLAR TIDIGARE. DU FÅR ALLTSÅ EN SEPARAT TID SEDAN FÖR ATT ÅTERBERÄTTA HANDLINGEN — JAG SÄGER TILL. HAR DU NÅGRA FRÅGOR?” / *”YOUR TASK IS NOW TO SILENTLY READ A SHORT STORY AND TRY TO REMEMBER IT. YOU SHOULD THEN TRY TO RETELL THE CONTENTS OF THE STORY TO [HIM/HER]. NO AIDS ARE ALLOWED. YOU HAVE FIVE MINUTES TO DO THIS. I WILL ANNOUNCE WHEN ONE MINUTE IS LEFT. YOU CAN LET ME KNOW IF YOU ARE FINISHED EARLIER. YOU WILL GET ANOTHER SEPARATE TIME PERIOD LATER TO RETELL THE STORY — I WILL TELL YOU WHEN. DO YOU HAVE ANY QUESTIONS?”*

3. Questions were answered by the experimenter, if asked.

4. The timer was started. Subject 1 was told to begin:

   ”NU KAN DU BÖRJA.” / *”YOU CAN START NOW.”*

   Subject 1 was monitored so that he or she did not take written notes.

5. Meanwhile subject 1 read the text, subject 2 was shown the photograph, with the purpose of letting subject 2 have something to do while he or she waited.

6. Subject 1 was informed when one minute of the time was left:

   ”NU ÄR DET EN MINUT KVAR.” / *”NOW IT'S ONE MINUTE LEFT.”*

7. The task ended and the timer was stopped. Tasks could end in one of two ways:

   (i) The time limit of 5 minutes was reached

   (ii) Subject 1 said he or she was finished ahead of time

In the first case, subject 1 was told that the time was up:

”NU ÄR TIDEN UTE.” / *”TIME IS UP.”*

8. The paper with the story is removed by the experimenter. The reading time was noted.

9. The subjects were instructed:

”DIN [TURNING TO SUBJECT 1] UPPGIFT ÄR NU ATT ÅTERBERÄTTA HANDLINGEN I DEN BERÄTTELSE DU NYSS LÄST, SÅ VÄL DU KAN. DIN [TURNING TO SUBJECT 2] UPPGIFT ÄR ATT SEDAN ENSAM SVARA PÅ NÅGRA SKRIFTLIGA KRYSSFRÅGOR OM BERÄTTELSEN. NI FÅR PRATA OCH FRÅGA VARANDRA FRITT UNDER TIDEN. NI FÅR INTE HA NÅGRA HJÄLPMEDEL. NI FÅR 3 MINUTER PÅ ER. JAG SÄGER TILL NÄR DET ÄR 30 SEKUNDER KVAR. HAR NI NÅGRA FRÅGOR?” / *”YOUR [TURNING TO SUBJECT 1] TASK IS NOW TO RETELL THE CONTENTS OF THE STORY YOU HAVE READ, AS WELL AS YOU CAN. YOUR [TURNING TO SUBJECT 2] TASK IS LATER TO ON YOUR OWN ANSWER SOME WRITTEN MULTIPLE CHOICE QUESTIONS ABOUT THE STORY. YOU MAY SPEAK FREELY AND ASK EACH OTHER WHAT YOU WANT WHILE YOU RETELL THE STORY. NO AIDS ARE ALLOWED. YOU HAVE THREE MINUTES TO DO THIS. I WILL LET YOU KNOW WHEN THERE IS 30 SECONDS LEFT. DO YOU HAVE ANY QUESTIONS?”*

10. Questions were answered by the experimenter, if asked.

11. The timer was started. Starting instructions were given to the subjects:

”NU KAN NI BÖRJA.” / *”YOU CAN START NOW.”*

Subject 1 retold the story to subject 2, during a phase of unrestricted communication, with the exception that subject 2 was monitored so that he or she did not take written notes. The subjects were informed when 30 seconds of the time was left:

”NU ÄR DET TRETTIO SEKUNDER KVAR.” / *”NOW IT'S THIRTY SECONDS LEFT.”*

12. The timer was stopped. Tasks could end in one of two ways:

(i) The time limit of 3 minutes was reached

(ii) Subject 2 said he or she was finished ahead of time

In the first case, the subjects were told that the time is up:

”NU ÄR TIDEN UTE.” / *”TIME IS UP.”*

The time taken to retell the story was noted.

13. Subject 2 was given the questionnaire, a pencil and an eraser. Subject 2 was instructed:

    ”DIN UPPGIFT ÄR NU ATT SVARA PÅ NÅGRA SKRIFTLIGA KRYSSFRÅGOR OM DEN BERÄTTELSE DU NYSS HÖRT. HJÄLPMEDEL ÄR PENNA OCH SUDDGUMMI. LÄMNA INGET BLANKT, UTAN CHANSA HELLRE ÄN ATT INTE BESVARA NÅGON FRÅGA. HAR DU NÅGRA FRÅGOR?” / *”YOUR TASK IS NOW TO ANSWER SOME WRITTEN MULTIPLE CHOICE QUESTIONS ABOUT THE STORY YOU HAVE HEARD. AIDS ARE PENCIL AND ERASER. DON'T LEAVE ANY QUESTIONS WITHOUT ANSWERING THEM, TRY TO MAKE A GUESS IF YOU DON'T KNOW THE ANSWER TO A QUESTION. DO YOU HAVE ANY QUESTIONS?”*

14. Subject 2 answered the questionnaire. This usually took up to two minutes.

15. The experiment situation was noted, along with answer time, on the questionnaire.

16. The experiment part of the symbolic task was finished.

When both tasks were finished, subjects were told to fill out a questionnaire which had the purpose of obtaining a subject profile (for a list of subject variables recorded, see section 4.9). After this, subjects were informed that the experiment was over, and were offered some sweets.

Subjects were told that the purpose of the study could not yet be revealed, since it might affect the study. The subjects were however informed that a debriefing notice would be put up on the department notice board after the study had been completed, giving details about the study: the purpose, the correct answers, and references to the material used in the study.

It was announced to subjects that the results of the study would be revealed at the usual presentations of the master degree dissertations at the department (details were to be announced later by the department).

Lastly, subjects were thanked for their participation and were shown out of the laboratory.

4.14 EXTRAORDINARY EVENTS DURING THE EXPERIMENT

The experiment ran smoothly with the following exceptions:

About halfway through the series of experiment sessions, a subject found a spelling error in the story text. He corrected it with a pencil, and the correction remained there for the rest of the experiment. No reason was seen to remove it, as it might be spotted by another subject again later. This change of the experiment material was judged not to have any impact on the experiment conditions.

On one occasion, subject 2 might have been able to read parts of the story from a distance, as subject 1 unintentionally held the papers in a position to allow the possibility of reading. Even if reading occurred, this should not significantly have affected the study.

By mistake, subjects were not notified of the remaining time on two occasions. This was not considered to be a significant error.

All subjects completed all tasks and answered all questions except two subjects who did not answer question number 6 on the questionnaire. These were simply handled as if the incorrect answer had been given.

4.15 DATA ANALYSIS

After the experiment was completed, all 40 groups (pairs of subjects) were randomly assigned a number from 1 to 40 in order to be able to reference them.

4.15.1 Spatial task

The 40 drawings were marked with their corresponding group number previously assigned.

In order to see how similar the drawings were relative to the photograph, the drawings were given absolute judgements using category scaling. The judgements were made by five separate judges, unaware of in which experiment situation the drawings were created. Judges consisted of five students from the University of Skövde (two female and three male, studying different courses) selected at random, who performed the judgement for pay. The judges decided how similar they thought the drawings were to the photograph, on a 7-point scale, where 1 is least and 7 is most similar. A 7-point scale was chosen because it was considered to give judges a sufficient degree of discrimination without at the same time be too fine-grained to pose problems of which grade should be chosen (the scale 1 to 100 would make if difficult for judges to discriminate between, say, a rating of 67 and 68). The judges were instructed to first pick out the least and most similar drawings and give them a grade of 1 and 7, respectively. This was made to anchor the judgements, that is, to ensure that the entire range of grading was used.

The reason to use five judges instead of only one, is that a judge might have an odd way of grading the drawings. Depending on one judge only would mean that the entire result on the spatial task would be affected by possible unusual characteristics of that judge. For instance, a judge who grades the drawings in a random-like manner is not desirable. There was no particular reason behind choosing exactly five judges. The number five was chosen as a number large enough to balance out any odd judgements.

The grades were written on a separate grading form. The mean value of all five judges was then calculated for each drawing, followed by total means for each of the two groups. When all judgements had been collected, they were analyzed statistically with a t-test to tell if the two groups' means differed significantly.

4.15.2 Symbolic task

The questionnaires were corrected and given a score between 0 and 8 points, depending on how many questions were correctly answered.

The total means for the two groups were calculated.

When all experimental data had been collected, they were analyzed statistically with a t-test to tell if the two groups' means differed significantly.

4.16 PILOT STUDIES

Before the experiment was carried out on a large scale, two pilot studies were conducted in order to test the appropriateness of the material and procedure.

4.16.1 First pilot study

A first pilot study was performed in laboratory setting, with purpose of testing the material, experimental settings, and procedure. The pilot study included one pair of subjects, the experimenter, and in addition, an external observer (a student trained in empirical studies). The external observer's role was to observe the entire study, including the experimenter's actions in order to discover any mistakes in experimental procedure.

This first pilot study resulted in adjustments of tasks and material used. Initially, a simple picture consisting of a few geometrical shapes were used in the study. It was realized that the picture had to have enough detail to stop subjects from—in the experiment situation with visual signals—draw the picture in the air with their finger in front of the other subject. The external observer was able to give some general feedback concerning the experimental procedure to the experimenter.

4.16.2 Second pilot study

A second pilot study involving two pairs of subjects was run in a non-laboratory setting (in ordinary rooms) to test the changes made after the first pilot study. Primarily, the material and procedure were tested. The spatial task was tested with one pair of subjects and the symbolic task was tested with two pairs of subjects. The chosen photograph worked well as material in the spatial task: the subject had enough time to be able to make a rough sketch of the photograph described to him. When testing the symbolic task, the questions on the narrative was found to be slightly too easy. Two additional questions—which were judged to have a high degree of difficulty—was added, making it eight questions in total.

5 RESULTS

Att förlora en illusion innebär
alltid att bli en sanning rikare.

Arthur Schnitlzer

The mean task performance in the two experiment situations can be seen in Figure 5.1. As this figure shows, there is little difference in performance between the two experiment situations. Figures 5.2 and 5.3 show the frequency distribution of the two tasks and give a more detailed picture of the results. No significant differences were found between the performance of the groups under the two experiment situations, as shown in Table 5.1 (the probabilities returned by the t-tests should be .05 or lower if the difference can be said to be significant). For some examples of drawings produced by subjects, the reader is referred to Appendix B.

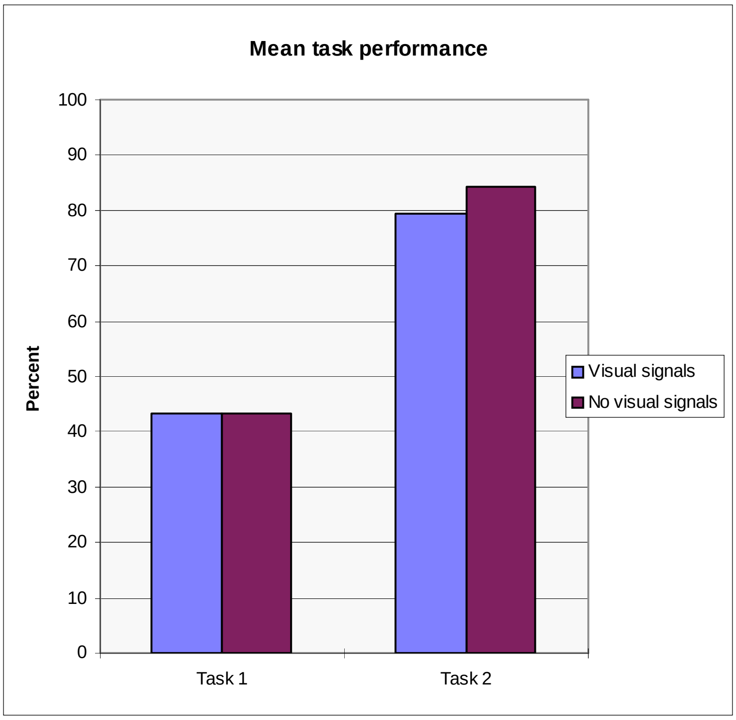

Figure 5.1. Mean task performance of the two tasks.

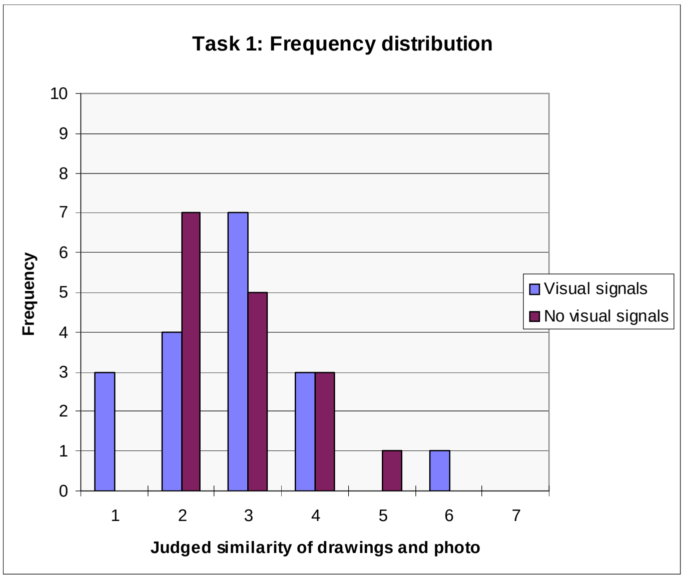

Figure 5.2. Frequency distribution of the spatial task. (The mean ratings of five judges.)

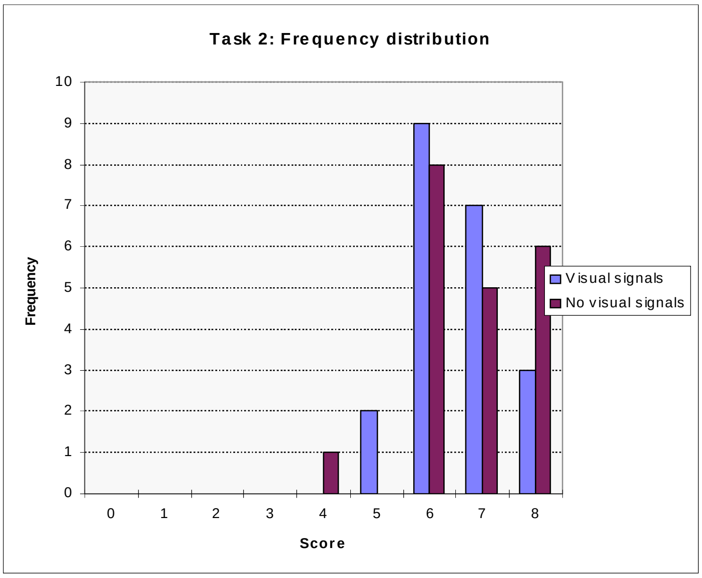

Figure 5.3. Frequency distribution of the symbolic task.

SPATIAL TASK		
	Mean	Standard deviation
Visual signals	**3.04 (43.4%)**	**1.351**
No visual signals	**3.03 (43.3%)**	**.856**
t-test	**.490**	

SYMBOLIC TASK		
	Mean	Standard deviation
Visual signals	**6.35 (79.4%)**	**1.040**
No visual signals	**6.75 (84.4%)**	**.875**
t-test	**.119**	

Table 5.1. Means, standard deviations, and t-test returned probabilities for the two tasks.

6 DISCUSSION

There is reason in the roasting of eggs.

English proverb

6.1 RELATION TO EARLIER STUDIES

As the results show, no differences in task performance between voice-only and face-to-face interaction were found. This result is consistent with the studies of Chapanis and colleagues (Chapanis *et. al.*, 1972; Ochsman & Chapanis, 1974, Chapanis, 1975), and also the various work described by Williams (1977).

Contrary to what was expected, the lack of visual cues for turn-taking and feedback did not affect task performance. This result agrees with Short *et. al.* (1976) who argue that these visual cues are taken over by the audio channel.

The possibility to make explicit, illustrative gestures on the spatial task, when visual signals were present, did not improve task performance, compared with when visual signals were absent. During the experiment, only about three cases of illustrative gestures were noted.

Gestures for retelling a narrative, as described by McNeill (1992) did not improve performance on the story task. Gestures for switching between meta-level and back in the narrative do not appear to have improved understanding.

As noted by Tang (1991) subjects used gestures in the situation where visual signals were prevented, even though they knew that the other person could not see them.

6.1.1 Visual signals analyzed

Benford *et. al*. (1995) describe 19 concepts that can be used to analyze systems for collaboration. In their article they show how these concepts apply to collaborative systems by analyzing two virtual reality systems, a computer game, and ordinary video communication.

When examining these concepts in relation to this experiment, ten of them were seen to be present, namely: *presence*, *location*, *identity*, *activity*, *viewpoints*, *actionpoints*, *availability*, *gesture*, *facial expression*, and *voluntary versus involuntary expression*. Some of them are present in both experimental situations, while some are available only when visual signals are permitted. The ten concepts will now be briefly described and shown how they relate to the experiment, with the purpose of giving a more detailed picture of the experiment situation.

Concepts present in both experiment situations:

- *Presence*. (If and how other persons are aware of the someone's presence in the environment.) This factor is available to subjects in the experiment. In both experiment situations, both subjects know that the other person is sitting on the other side of the table, ready to carry out the task.

- *Identity*. (To be able to recognize who a person is.) This is not necessary in the experiment since it only involves two persons at a time. There is no need to be able to differentiate persons when collaborating.
- *Activity*. (Convey a sense of on-going activity.) In a broad sense, subjects in the experiment know that the other is involved in some activity—that is the task they are currently working on (but see *Availability* below).
- *Voluntary versus involuntary expression*. (Some expressions are consciously controlled by persons, such as stretching out their tongue. Others are not consciously controlled, such as looks of shock, anger, and fear.) Since no mediating technology is used in the experiment, all expression naturally available to people will be conveyed. The exception is involuntary visual expressions, which will of course not be available in the experiment situation without visual contact.

Concepts available only when visual signals are permitted:

- *Location*. (Conveying position and orientation of a participant. This can be used for leaning towards or back from someone, or turning one's back on someone, etc.) This is only available in the experiment situation where visual signals are permitted.
- *Viewpoints*. (Where a person is attending; closely related to gaze direction). This can only be seen by the other person when visual signals are permitted. It is totally absent in the situation without visual contact.
- *Actionpoints*. (Where in space a participant is manipulating.) When no visual contact is available, it is not possible to see where a person is manipulating something. Since small screens were used to prevent subjects from seeing each other's material, the *exact* actionpoint of one person could not be seen by the other person, even when visual signals were permitted.

- *Availability*. (To see how busy and/or how interruptable a participant is.) In the experiment when visual signals are prevented, there is no way of seeing when the other person behind the screen is finished and is waiting, or is busy doing something (this is mainly relevant in the spatial drawing task). However, this could, and was in the experiment, conveyed by subjects using explicit questions like ”Are you finished?”.
- *Gesture*. (Use of limbs to convey meaning.) Fully supported in the experiment situation with visual signals, and totally absent from the situation without visual contact.
- *Facial expression*. (Use of the face to convey meaning.) Fully supported in the experiment situation with visual signals, and totally absent from the situation without visual contact.

6.2 NATURE OF TASKS

6.2.1 Cooperative and conflictful tasks

The tasks were cooperate in nature. That no differences were found is consistent with earlier studies involving cooperative tasks. The story used in the study had two questions on emotion of the characters, and the performance of these could be expected to be lower when visual signals are absent. It is however unsure whether the subject retelling the story really expressed these emotions, as he or she would have done if experiencing the emotions himself or herself. Analysis of these two questions shows that there are no significant differences between the experiment situations. However, when analyzing question 8 in isolation, an almost significant reduction in performance was found, from 95% to 80% ($p = .08$).

Although the tasks had as few restrictions as possible in order to make them as natural as possible, this could have been carried even further. Tasks could have been more collaborative in character, making more of the natural effects of collaboration emerge. Giving people a chance to act naturally gives the study a higher ecological validity. For example, the spatial drawing task could have been designed so that both subjects could collaborate on the drawing. The task of retelling a narrative could have been modified so that both subjects read and discussed the narrative, and answered the questions together.

6.2.2 Other factors affecting task performance

The time available for subjects on tasks was limited for practical reasons: if each pair would have had unlimited time, the total time taken would be too long. What effect did time limits have on the solving of the tasks? On the symbolic task, almost all subjects were finished ahead of time, and the limit should therefore not affect the results. On the spatial task, almost every subject was interrupted in drawing when the time limit of five minutes was reached. This means subjects probably did worse than if they would have had more time to draw. However, this does not affect the study, since only the relative values are of interest in comparing the two experiment situations. That drawings are made less similar to the photograph because of lack of time has the same effect in both situations. More time on tasks (especially on the spatial task) could have affected the results. Perhaps if subjects had, say, ten minutes instead of five, subjects with visual contact might have produced better drawings, and the differences would have been greater between the two experiment situations.

There is the question of learning in tasks. Task performance in the experiment situation without visual signals improve over time. It might be unusual for subjects at first not to have visual contact, but they might get accustomed to the situation, making task performance gradually better. However, as the results show no differences in task performance between absence or presence of visual signals, learning in this way does not seem to be significant in task performance.

A potential source of error was the time passing between the three phases of the symbolic task. Time passed, firstly, between the reading and the retelling, and secondly, between the retelling and the answering of questions on the narrative. In the first case, subject 1 had just read the narrative and had it all fresh in memory. In the second case, subject 2 had just been retold the narrative and had that in memory. The amount of time passing in both these cases until subjects could start on the next phase would be critical to how much is remembered. Items from short-term memory will quickly be lost, especially as the subjects are distracted by the verbal instructions given to them. Effort was spent to see to that roughly the same amount of time passed in these cases for all subjects.

Finally, an observation related to the tasks is that the spatial task was not made harder compared with the story task by the absence of visual signals, as was expected.

6.3 COMPOSITION OF COLLABORATION GROUPS

6.3.1 Group size

The collaboration groups included in the study were made up of two persons. This was a conscious delimitation made before the study. In a natural CSCW setting probably more than two persons would be involved in collaboration most of the time. Chapanis (1975) also used groups of two people and says on the question on generalization to larger groups that ”the full implications of this kind of communication are not known” (Chapanis, 1975, p. 42).

In video communication, there are dramatic differences between two or more people interacting, as noted by Benford and Fahlén (1994). If a person at a remote site looks into the video camera when there are only one person on the other side there is no problem. But who is the person looking at when there are more than one person viewing the screen at the other end? All persons at the viewing end get the impression that the person is looking at them in particular.

These considerations make it highly questionable if results obtained from groups of two persons can be generalized to groups with more than two persons.

6.3.2 Social and cultural aspects

The subject variable *acquaintance* of pairs in each group was selected in the experiment so that no pairs had subjects who were acquainted with each other. This was considered to provide the most equal conditions for the pairs. In hindsight, this is not representative of most groups in natural settings. Instead, people working together probably know each other. The effect of acquaintance could be that when people know each other, they are better at interpreting each other's visual signals. This means that even if there is a difference between absence and presence of visual signals, there might not be a difference when groups consist of persons who do not know each other. On the other hand, one could compose pairs where subjects knew each other, but that would introduce the difficulty of knowing *how well* they know each other. This measure appears difficult to obtain, and without it this kind of alternative group composition could not be made with maintained control of the subject variables.

Among the subjects some pairs were male only, some female only; the largest part were mixed. Are there any differences between groups with mixed versus single sex composition? One might suspect that when groups are mixed, subjects might spend more time looking at each other in the situation where visual contact was possible, possibly resulting in lower task performance. Analysis of task performance in mixed-sex versus single-sex groups showed however no significant differences between the two.

Earlier literature suggest that there might be sex differences in the interpretation of visual signals so that females exhibit a higher sensitivity to reduced visual signals than males. This hypothesis could not be examined since the data did not contain enough groups of females only.

The results obtained in the study can only be surely generalized to people of the same cultural background as the subjects. If the same study were carried out in a country other than Sweden, let's say Japan, the results could have been very different. This is because the use of non-verbal codes differ much between cultures (see e.g. Fiske, 1982).

For example, in Sweden, people rely relatively little on the use of gesture. It could be that the results look the way they do because Swedes use relatively few gestures. In a culture where gestures are used more, one might get significantly better task performance when visual signals are available (see Chapanis, 1975).

6.4 EXPERIMENT SETTINGS

Although effort was made to make the experiment settings as natural and restriction-free as possible, some subjects might have been under the impression that they were not allowed to use gestures. Subjects might have consciously stopped themselves from using visual codes. This would mean that impact of gestures were lost, which would explain the results. Even if there was a difference between absence and presence of visual signals, it may not have been seen because of artificial laboratory conditions.

The experimenter sat beside the subjects throughout the experiment. Experimenter presence might lower the performance of groups, because it made the subjects pressured or nervous. As this happened in both experiment situations, it should however have no effect on the relative performance. There is no reason to believe that experimenter presence affects subjects in the use of visual signals, so that one situation would be favored.

The lack of differences in task performance could be attributed to compensatory effects. Subjects might have compensated the lack of visual signals by extra use of voice to eliminate any differences in performance between the two situations. However, the observation made was that the photographs were not described in any particularly different way between the two experiment situations (as perceived by the experimenter—no detailed data of interaction were recorded).

6.5 GENERALIZATIONS

The focus of this study is to find general principles on human communication, that are not tied to any special technology. The results can be said to be general if they: 1) hold outside of this particular experiment setting, and 2) hold across tasks used in the experiment.

Although arrangements can be made to as much as possible make the experiment setting natural, it is not possible to verify that the first condition is fulfilled. To see if the results appear possible to generalize to a natural setting, one has to relate this study to earlier studies in the literature. Doing so reveals that the results are consistent with earlier studies, and judging on the basis of that, the results obtained in this study appear to be general.

The second condition—that results hold across tasks—can be checked within the study simply by comparing the different tasks used. As can be seen, the results hold across both tasks used in the study. Thus, the results are general across tasks, at least those used in this study.

6.6 SUMMARY OF DISCUSSION

Things that might have reduced the difference between the two experiment situations:

- The time limit on the spatial task, making drawings too similar
- Artificial conditions might have stopped subjects from using visual signals
- Subjects might have compensated for lack of visual signals using voice

Lack of visual cues for turn-taking and feedback, lack of possibility to make illustrative gestures, and gestures for meta-discourse information did not affect task performance. Tasks used were cooperative, but could in retrospect have been even more cooperative and free of restrictions. On the question of group composition, it was noted that the results can probably not be generalized to groups of more than two persons. Further, that group members were unacquainted was considered unrealistic. No support for sex differences of task performance in groups were found, neither between single-sex versus mixed-sex groups, nor between female-only groups versus the rest.

7 CONCLUSIONS

The most exciting phrase to hear in science, the one that heralds new discoveries, is not "Eureka!" (I found it!) but "That's funny ..."

Isaac Asimov

Technology has a great impact on society. We want technology that facilitates our daily lives and our work, technology that helps us carry out the things we want to do. The underlying reasons for our design, what we should put in our technology, thus becomes a matter of greatest concern. If we do not even know what we need to put into our technology, there is little help to know how to accomplish even the most technically advanced feat. We need to know the *what* before we should even start to think about the *how*.

The aim of this study is to investigate if visual signals are important to include in technology for human collaboration. Results from examining the literature show that visual signals *are* important in collaboration, at least in situations which require the expression of emotion, such as negotiation. The empirical study carried out shows that absence or presence of visual signals have no significant effects on task performance in cooperative tasks. Thus, this falsifies the initial hypothesis that visual signals would facilitate task performance. The results are however consistent with earlier studies in the literature. Some doubts about the validity, regarding the composition of the collaboration groups and the nature of tasks were raised. These considerations do not make the results less firm, but constrain the conditions under which they hold.

Now, the initial questions raised in this dissertation can be answered. One question asked was what basic mechanisms are at work in human collaboration. The study contributes to a part of this large whole—the part of human visual signals—and provides results showing the role of visual signals in collaboration. Another question asked—concerning the area of CSCW—was how software should be designed to best facilitate collaboration. The practical implications that follows from the results of this study are that if a particular computerized system is to be used for cooperative tasks only, visual signals would not be important to include. But if the task involves the expression of attitudes, values or emotions, visual signals would be an important feature. So, for instance, a distributed meeting in the DIVE system *should* include highly detailed visual codes since a meeting often includes negotiation which depends on the expression of individual's differing motives.

This study has also raised several new questions which will be treated next, in Chapter 8, *Future research.*

8 FUTURE RESEARCH

Don't let the verbal people get away with it!

Ajit Narayanan

The study has raised several questions which point to further research.

All visual signals were treated as a single group. The reasons for doing so has already been discussed. If a good way of studying single aspects of visual signals, for example, gestures only, is found such a study would provide more detailed data. This way, it could be seen if certain visual signals are more important than others. One should however be careful to make a top-ten-list of which visual code is the most important, the second most important, and so on. There are holistic interaction effects between codes which makes such a rating of importance highly questionable (Short *et. al.*, 1976).

It would be interesting to investigate if and how visual signals can be implemented in real computerized systems. How can the various types of visual signals be incorporated into a system? Do some types pose special problems, and are some easier than others to implement? (Some work of this kind is described in Badler, 1995 and Lu, Yoshizaka & Miyai, 1996.)

Can the results obtained in this study be applied to any technology, and if not, under which conditions can it be applied? It would be useful to know exactly which minimal set of criteria such technology need to fulfill in order for the results found in this study to apply to it.

To obtain results which say something about the importance of visual signals in groups larger than two persons, studies with groups of more than two persons have to be conducted. Also, perhaps further division of group size, above that of two persons, needs to be made.

This study only included cooperative tasks. It would be interesting to see what differences would be seen if conflictful tasks were used.

Perhaps people learn to handle the absence of visual signals, even in tasks of negotiation? Studies where repeated tasks are used could be made to see if task performance is increased on consecutive tasks of the same type.

REFERENCES

Aukstakalnis, S. & Blatner, D. (1992). *Silicon Mirage: The Art and Science of Virtual Reality*, Peach Pit Press, ISBN 0-938151-82-7.

Badler, N. I. (1995). *Synthesizing Conversation Between Human-Like Cooperative Agents*. [Online] Available http://www.cs.tufts.edu/~jacob/isgw/Badler.html, August 27, 1996.

Baecker, R.M. (1993). *Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration*. Morgan Kaufmann Publishers, San Mateo, CA.

Ball, G., Ling, D., Kurlander, D., Miller, J., Pugh, D., Skelly, T., Stankosky, A. Thiel, D. Van Dantzich, M., & Wax, T. (1996). *Lifelike Computer Characters: the Persona project at Microsoft Research.* To appear in *Software Agents,* Jeff Bradshaw Ed., MIT Press.

Bates, J. (1994). *The Role of Emotion in Believable Agents*. Communications of the ACM, July 1994/Vol. 37, No. 7, pp. 122-125.

Benford, S. & Fahlén, L. (1993). *A Spatial Model of Interaction in Large Virtual Environments*. In Proceedings of the Third European Conference on Computer Supported Cooperative Work (ECSCW'93), Milan, September, 1993.

Benford, S. & Fahlén, L. (1994). *Viewpoints, Actionpoints and Spatial Frames for Collobarative User Interfaces.* In HCI'94, 1994, Glasgow, UK.

Benford, S., Bowers, J., Fahlén, L., Greenhalgh, C., & Snowdon, D. (1995). *User Embodiment in Collaborative Virtual Environments*. In Proceedings from CHI'95, Denver, USA, May 1995, ACM Press.

Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., & Stone, M. (1994a). *Animated conversation: Rule-based* *generation of facial expression, gesture and spoken intonation for multiple conversational agents*. In Proceedings of SIGGRAPH '94.

Cassell, J., Stone, M., Douville, B., Prevost, S., Achorn, B., Steedman, M., Badler, N., & Pelachaud, C. (1994b). *Modeling the interaction between speech and action*. In Proceedings of Cognitive Science Society, Atlanta, Georgia.

Chapanis A., Ochsman R. B., Parrish, R. N., & Weeks, G. D. (1972). *Studies in interactive communication: The effects of four communication modes on the behaviour of teams during cooperative problem solving*. Human Factors 14: pp. 487-509.

Chapanis, A. (1975). *Interactive human communication*. Scientific American, 232, 1975, pp. 36-42.

December, J. (1996). Computer-Mediated Communication Magazine, ISSN 1076-027X, Vol. 3, No. 7, July 1, 1996.

Ekman, P., Huang, T., Sejnowski, T., & Hager, J. (Eds.) (1992). *Report To NSF of the Planning Workshop on Facial Expression Understanding*. [Online] Available http://mambo.ucsc.edu/psl/nsf.txt, August 27, 1996.

Greenberg, S. (1991). *Computer-supported cooperative work and groupware: an introduction to the special issues*. International Journal of Man-Machine Studies, Vol. 34, No. 2, pp. 133-142.

Greenhalgh, C. (1996). *MASSIVE*. [Online] Available http://www.crg.cs.nott.ac.uk/~cmg/massive.html, August 27, 1996.

Gärdenfors, P. (1992). *Blotta tanken*. Nya Doxa: Falun.

Huang, T. S. & Orchard, M. T. (1992). *Man-machine interaction in the 21st century: New paradigms through dynamic scene analysis and synthesis*. Proceedings of SPIE Conference on Visual Communications and Image Processing '92, Vol. 1818 (pp. 428-429). Nov. 18-20, Boston, MA.

Ibister, K. & Layton, T. (1995). *Intelligent Agents: a review of current literature*. [Online] Available http://www.research.microsoft.com/research/ui/persona/isbister.htm, August 27, 1996.

Isaacs, E.A., T. Morris, T.K. Rodriguez, & J.C. Tang. (1995). *A Comparison of Face-to-Face and Distributed Presentations*. Proceedings of the Conference on Computer-Human Interaction (CHI) '95, Denver, CO, in press.

Isaacs, E.A. & Tang, J.C. (1993). *What video can and can't do for collaboration: a case study*. In Proceedings of the ACM Multimedia '93 Conference. Anaheim, CA.

Kvan, T. (1994). *Reflections on computer-mediated architectural design*, IEEE, Transactions on Professional Communications, December 1994, available http://arch.hku.hk/people/tkvan/acm-94.html.

Lindström, P. (1994). *Att resa i Estland*. Norstedts, Stockholm.

Lu, S., Yoshizaka, S., & Miyai, H. (1996). *Creating a Human-like Computer Character with distinctive features of body motion and appearance.* [Online] Available http://www.research.microsoft.com/research/ui/lcc/abs/Lu.txt, August 27, 1996.

Lundh, L-G., Montgomery, H., Waern, Y. (1992). *Kognitiv psykologi*. Lund: Studentlitteratur.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.

Microsoft. (1996). *Microsoft Bob version 1.0 for Windows*. [Online] Available: http://portal.nienschanz.ru/catalog/products/Bob/default.htm, August 27, 1996.

Miller, J. G. (1995). *Living Systems*. Colorado: University Press.

Nilsson, B. (1993). *I ord och handling*. Studentlitteratur: Lund.

Norman, D. A. (1992). *Turn signals are the facial expressions of automobiles*. Addison Wesley.

Norman, D. A. (1993). *Things that make us smart*. Addison Wesley.

Ochsman, R. B. & Chapanis, A. (1974). *The Effects of 10 Communication Modes on the Behaviour of Teams during cooperative Problem Solving*. International Journal of Man-Machine Studies 6: pp. 579-619.

Olson, M. H. & Bly, S. A. (1991). *The Portland Experience: a report on a distributed research group*. International Journal of Man-Machine Studies, Vol. 34, No. 2, pp. 211-228.

Pitkow, J. E. & Kehoe, C. (1995). *GVU's 4th WWW Survey*. [Online] Available http://www.cc.gatech.edu/gvu/user_surveys/survey-10-1995/, August 27, 1996.

Quek, F. (1994). *Toward a Vision-Based Hand Gesture Interface*. In Proceedings of VRST '94 – Virtual Reality Software and Technology (pp. 17-32). River Edge, NJ, USA: World Scientific Publishing Company.

Quek, F. (1995). *Non-Verbal Vision-Based Interfaces*. Keynote Speech, International Workshop in Human Interface Technology '95, IWHIT'95, October 12-13, Aizuwakamatsu, Fukushima, Japan.

Shannon, C. & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Illinois.

Short, J., Williams, E., & Christie, B. (1976). *Visual Communication and Social Interaction*. Reprinted in (Baecker, 1993).

Ståhl, O., & Andersson, M. (1995). *DIVE - A Toolkit for Distributed VR Applications*. [Online] Available http://www.sics.se/dce/dive/online/ercim.html, August 27, 1996. (See also the DIVE Home Page at http://www.sics.se/dive/)

Tang, J. C. (1991). *Findings from observational studies of collaborative work*. International Journal of Man-Machine Studies, Vol. 34, No. 2, pp. 143-160.

Tatar, D.G., Foster, G. & Bobrow, D. G. (1991). *Design for conversation: lessons from Cognoter*. International Journal of Man-Machine Studies, Vol. 34, No. 2, pp. 185-210.

Towell, J., & Towell, E. (1995). *Internet Conferencing with Networked Virtual Environments*, Internet Research, Vol. 5, No. 3, pp. 15-20.

Valacich, J.S., Dennis, A. R. & Nunamaker, Jr, J. F. (1991). *Electronic meeting support: the GroupSystems concept*. International Journal of Man-Machine Studies, Vol. 34, No. 2, pp. 261-282.

Werner, L. (1972). *Puck på internatskola*. Almgrens, Örebro. ISBN 91-7216-000-4.

Wertheim, E. G. (1996). *The Importance of Effective Communication*. [Online] Available http://www.cba.neu.edu/~ewertheim/interper/commun.htm, August 27, 1996.

Williams, E. (1977). *Experimental comparisons of face-to-face and mediated communications: a review*. Psychological Bulletin, 84, pp. 963-976.

APPENDIX A:

THE PHOTOGRAPH USED IN THE SPATIAL TASK

APPENDIX B:

THE LEAST-LIKE, MIDDLE-LIKE, AND MOST-LIKE DRAWINGS

DRAWING GIVEN SCORE 1.0 OF 7:

DRAWING GIVEN SCORE 4.0 OF 7:

DRAWING GIVEN SCORE 6.8 OF 7:

APPENDIX C:

THE TEXT USED IN THE SYMBOLIC TASK

Alltsammans började en regnig dag i det stora ritkontoret på femte våningen i ingenjörsfirman "Danaplan". Regnet trummade mot de snedställa fönsterna. Himlen var mörkgrå och dyster. Men lysrören i ritkontorets tak sörjde för att ingenjörerna kunde arbeta ostörda.

Ingenjör Jörgen Winther hade bråttom med några invecklade beräkningar, när telefonen bredvid hans ritbord ringde. Han lade pennan ifrån sig, vände sig om i stolen och tog telefonen.

– Hallå!

– Winther? Det är Bang. Kom ner till mig ett tag.

Luren lades på. Winther reste sig och gick ut ur ritkontoret. Strax efteråt stod han i överingenjör Bangs stora arbetsrum. Bang själv satt bakom ett jättestort skrivbord, upptagen av att sätta sitt namn på en hög brev. Han såg hastigt upp och sa:

– Var så god och sitt . . . jag är färdig om en sekund.

Winther satte sig och började stoppa sin pipa. Han hade nätt och jämnt blivit klar med det när överingenjören sköt breven från sig och lutade sig tillbaka i sin bekväma vridstol.

– Jag vill tala om något viktigt med er, sa han. – Det är fråga om en befordran! I själva verket borde jag lyckönska er . . .

Winther svarade inte. Han förstod inte riktigt vad det rörde sig om. Bang skar av spetsen på sin cigarr och gav sig god tid att tända, innan han fortsatte:

– Ni känner till hamnprojektet i Valparaiso. Ni har själv varit själen i planläggningen. Ni vet hur mycket som står på spel för oss och ni vet att det är en stor och mycket invecklad uppgift, var lösning är beroende av vem som leder arbetet . . .

Winther nickade:

– Ja, sa han, – det vet jag, men vi har ingenjör Henriksen där, och i bättre händer än hans kan det väl inte vara?

Bang nickade:

– Alldeles riktigt, sa han. – Christian Henriksen är duktig. Men vi har i dag fått ett telegram från Valparaiso om att han har blivit allvarligt sjuk. Han har lagts in på sjukhus med en reumatisk sjukdom som kommer att binda honom vid sängen i flera månader, kanske längre . . .

Jörgen Winther såg hastigt upp på överingenjören:

– Ni har väl inte tänkte skicka mig till Chile i Henriksens ställe? frågade han.

– Jo, det är precis vad vi har tänkt, svarade Bang. – ni var på tal som ledare av arbetet när projektet antogs, men av olika skäl valde vi till sist Henriksen. Men nu, när Henriksen har blivit sjuk, så måste vi ha er där.

Winther reste sig. Han hade blivit mycket allvarsam.

– Ja, men det är omöjligt, sa han. – Jag kan inte resa bort . . .

Det blev en liten paus. Bang betraktade intresserat sin cigarr. Winther lät handen glida genom sitt bruna hår. Till sist sa överingenjören:

– Jag är rädd för att firman inte kan ta några personliga hänsyn, Winther. Ni blir tvungen att åta er uppgiften.

– Ja men, utbrast Winther, – kan man inte finna en annan man till det där jobbet. Ni vet att jag har varit änkling i tre år, sedan min hustru omkom vid en bilolycka, och ni vet att jag har en liten dotter . . . jag . . .

– Jag vet allt det, sa Bang allvarligt, – men direktionen har sammanträtt i dag och har beslutat att skicka er till Valparaiso, och ni känner firman tillräckligt väl för att veta att ett direktionsbeslut inte går att rucka på. Förresten, tillfogade han med ett litet leende, så är det här ju som jag sa förut, en befordran, som jag egentligen borde lyckönska er till. Men jag vet att det blir svårt för er att skiljas från er lilla flicka, även om det bara är för en begränsad tid.

– Jag har inga släktingar, som kan ta sig an henne, medan jag är borta, sa Jörgen Winther olyckligt.

– Hade er hustru inte heller någon familj?

– Bara en syster, som bor i Jylland. Hon är gift med en lantbrukare, och det finns inte riktiga skolor i närheten. Jag måste se till att Bente får en god utbildning. Det är ju snart de enda värdefulla man kan ge sina barn. Och om hon nu plötsligt skulle vara utan en bra skola, kunde det inverka på hela hennes framtid. Ja, kort sagt: jag är inte vidare entusiastisk för den här planen, även om den rent arbetsmässigt naturligtvis tilltalar mig oerhört mycket. Måste firman ha sitt svar omedelbart?

– Nej, sa Bang, – det behövs inte. Tänk över det ett par dar och låt oss talas vid sedan.

Jörgen Winther nickade till sin överordnade och gick ut ur det stora rummet. I djupa tankar vandrade han genom yttre kontoret och ut till hissen.

APPENDIX D:

MULTIPLE-CHOICE QUESTIONS USED IN THE SYMBOLIC TASK (WITH CORRECT ANSWERS MARKED)

Besvara följande åtta frågor genom att sätta kryss vid det alternativ du tycker passar bäst!

1. Vilket yrke har huvudpersonen i berättelsen?

- ☐ Arkitekt
- ☒ Ingenjör
- ☐ Konstnär
- ☐ Verkstadsarbetare
- ☐ Snickare
- ☐ Sjöman

2. Vilka familjeförhållanden har huvudpersonen?

- ☐ Ingen familj alls
- ☒ Ingen fru och ett barn
- ☐ Ingen fru och två barn
- ☐ Fru och inga barn
- ☐ Fru och ett barn
- ☐ Fru och två barn

3. Vad händer i början av berättelsen?

- ☐ Huvudpersonen får besök
- ☒ Huvudpersonen får ett telefonsamtal
- ☐ Huvudpersonen får ett brev
- ☐ Huvudpersonen reser iväg för att träffa någon
- ☐ Huvudpersonen blir utsatt för ett brott
- ☐ Huvudpersonen bevittnar ett brott

4. Hur utvecklar sig berättelsen?

- ☒ Han erbjuds en befodran
- ☐ Han får höra att han ska bli avskedad
- ☐ Han får ett stort arv
- ☐ Han får reda på att han vunnit ett pris
- ☐ Han blir kallad som vittne
- ☐ Han gör en resa

5. Hur slutar berättelsen?

- ☐ Huvudpersonen bestämmer sig för att acceptera förslaget
- ☐ Huvudpersonen bestämmer sig för att inte acceptera förslaget
- ☐ Huvudpersonen kommer med ett eget, nytt förslag
- ☒ Huvudpersonen får tid att tänka över beslutet
- ☐ Huvudpersonen vägrar lyssna på förslaget

6. Vilken var orsaken till Henriksens problem?

- ☐ Bilolycka
- ☐ Fängelsestraff
- ☐ Försvunnen
- ☐ Hjärtattack
- ☐ Cancer
- ☒ Reumatism
- ☐ Lunginflammation

7. Hur är huvudpersonens känsloläge när berättelsen slutar?

- ☐ Bestämd
- ☐ Arg
- ☒ Tveksam
- ☐ Irriterad
- ☐ Ledsen
- ☐ Glad

8. Hur är huvudpersonens samtalspartners känsloläge?

- ☒ Bestämd
- ☐ Arg
- ☐ Tveksam
- ☐ Irriterad
- ☐ Ledsen
- ☐ Glad