Instructional support for individual and collaborative demands on net-based problem-solving in dyads
Diese Arbeit wurde gefördert vom Virtuellen Graduiertenkolleg (VGK) der Deutschen Forschungsgemeinschaft „Wissenserwerb und Wissensaustausch mit neuen Medien“.

Dekan: Prof. Dr. Dr. Jürgen Bengel

Erstgutachter: Prof. Dr. Hans Spada

Zweitgutachter: Prof. Dr. Alexander Renkl

Datum der Disputation: 13.12.2006
I worked on this dissertation between April 2003 and July 2006. During this time, many people helped and supported my work in various ways.

First, I want to thank my advisor Prof. Hans Spada. His useful comments and ideas as well as fruitful discussions enriched my work in a sustainable way. I learned a lot from him over these three years and I am very thankful for his trust and support.

Furthermore, my gratitude goes to all the professors of the VGK, especially Prof. Rainer Bromme, Prof. Friedrich Hesse, Prof. Alexander Renkl, and to all the VGK doctoral students, in particular Anna Ertelt, Jessica Philipps and Mareike Florax. I thank all the colleagues in the research group “Cognition, Emotion, Communication”, especially Nikol Rummel, Sabine Hauser, Anne Meier, Hannah Swoboda and Julia Kern, for many interesting discussions, helpful comments on the manuscript and the nice coffee breaks and funny evenings we shared together. In terms of technical support, I’m very grateful to Jörg Spirik, Michael Stumpf, Paul Hüttner and Rolf Schwonke. For providing support during the empirical portion of the work, my thanks go to Carmen Heckmann.

Moreover, I thank the Deutsche Forschungsgemeinschaft DFG for the financing of the dissertation through the Virtual PhD program VGK. I further thank Ravensburger© for the permission to publish pictures from the Ravensburger© children’s game Differix.

Additionally, I want to thank my family – my parents, brother and grandparents – as well as my friends for their loving support; of course not only over the last three years. Thank-you for countless hours of babysitting! I especially thank my husband Wasa Hansen for supporting and encouraging me in everything I do. I thank him and our daughter Pia for making my life worthwhile and beautiful.
# Table of Contents

List of Figures and Tables .....................................................................................................................6

Abstract .................................................................................................................................................10
Zusammenfassung .................................................................................................................................13

Introduction and Overview ..................................................................................................................16

Chapter 1: Theoretical Background ......................................................................................................19
1.1 Individual Cognitive Demands in Collaboration ...........................................................................21
1.2 Demands due to Interaction: Coordination and Communication .....................................................22
1.2.1 Coordination ................................................................................................................................22
1.2.2 Communication ............................................................................................................................24
1.3 Additional Demands of Net-Based Settings .....................................................................................35
1.4 Supporting Collaboration ..................................................................................................................37
1.4.1 Improving the Technical Environment .......................................................................................37
1.4.2 Structuring the Interaction Process: Collaboration Scripts ..........................................................41
1.4.3 Enhancing the Skills: Model collaboration ..................................................................................49

Chapter 2: The Collaborative Task ........................................................................................................51
2.1 The Task Demands ..........................................................................................................................53
2.1.1 Individual Cognitive Demands: Visual Search and Object Recognition ........................................54
2.1.2 Demands due to Interaction: Installing Referential Identity .......................................................55
2.1.3 Additional Demands of Remote Collaboration via Audio-link .....................................................55
2.2 Support Measures Implemented ..................................................................................................56
2.2.1 Study 1: Shared Applications .......................................................................................................57
2.2.2 Study 2: Model for Observational Learning from Model Collaboration and Collaboration Script ......57

Chapter 3: Study 1: Test for the Influence of Individual and Collaborative Demands on Process and Outcome ..................................................................................................................61
3.1 Method of Study 1 ............................................................................................................................61
3.1.1 Design ........................................................................................................................................66
3.1.2 Participants ..................................................................................................................................66
3.1.3 Procedure ....................................................................................................................................67
3.1.4 Measures .....................................................................................................................................67
3.1.5 Hypotheses ..................................................................................................................................70
3.2 Results of Study 1 ............................................................................................................................71
3.2.1 Process Data ...............................................................................................................................71
# Table of Contents

3.2.2 Performance Measures ..................................................................................................................... 74  
3.2.3 Use of the Shared Application .......................................................................................................... 76  
3.2.4 Summary of the Results (Study 1) .................................................................................................... 79  
3.2.5 Relation between Process and Performance Data ............................................................................... 80  
3.3 Discussion of Study 1 .................................................................................................................... 83  

## Chapter 4:  
Study 2: Instructional Support for Individual and Collaborative Demands .......... 87  
4.1 Method of Study 2.......................................................................................................................... 88  
4.1.1 Design ........................................................................................................................................... 89  
4.1.2 Participants.................................................................................................................................... 96  
4.1.3 Procedure ......................................................................................................................................... 96  
4.1.4 Measures............................................  
4.1.5 Hypotheses ..................................................................................................................................... 101  
4.2 Results of Study 2......................................................................................................................... 102  
4.2.1 Process Data.................................................................................................................................. 102  
4.2.2 Performance Measures .................................................................................................................... 109  
4.2.3 Summary of the Results (Study 2) .................................................................................................. 111  
4.2.4 Further Analyses ........................................................................................................................... 113  
4.3 Discussion of Study 2 .................................................................................................................. 121  

## Chapter 5:  
Overall Discussion ....................................................................................................... 126  
5.1 Summary and Discussion of Results .........................................................................................126  
5.1.1 Types of Demands..........................................................................................................................126  
5.1.2 Dialog vs. Monolog ........................................................................................................................128  
5.1.3 Support of Collaboration ................................................................................................................131  
5.2 Perspectives for Further Research and Practical Implications ..............................................133  

References………………………………………………………………..135  

## Appendices:  
Appendix A: Study 1 .......................................................... 152  
A.1 Materials from Study 1................................................................................................................152  
A.2 Examples for the two Types of Use of the Shared Applications in Study 1 ..........155  
Appendix B: Study 2.................................................................................. 157  
B.1 Materials from Study 2................................................................................................................157  
B.2 Examples of Collaboration Script for Speaker and Addressee in Study 2 ..........161  
Appendix C: Hardware Configuration in Study 1 and 2 .............................................. 162  

5
List of Figures

Figure 1: Different understandings of communication.........................................................25
Figure 2: The speaker’s and the addressee’s display (here one task from study 2).................52
Figure 3: The speaker’s display in four different tasks..........................................................62
Figure 4: Process data measures with significant differences between concrete and abstract  
sets of pictures..................................................................................................................72
Figure 5: Process data measures with significant differences between dyads with an addressee  
that was not able to talk and dyads with an addressee that was able to talk..................73
Figure 6: Significant interactions in process data measures between type of pictures and role of  
the addressee..................................................................................................................74
Figure 7: Mean number of correctly placed pictures (left) and mean time needed to complete  
the task (right) by type of pictures................................................................................75
Figure 8: Mean number of correctly placed pictures (left) and mean time needed to complete  
the task (right) by role of the addressee.........................................................................76
Figure 9: Type of use of the pictorial shared application (per cent) in concrete and abstract  
tasks by role of addressee.............................................................................................77
Figure 10: Type of use of the textual shared application (per cent) in concrete and abstract  
tasks by role of addressee.............................................................................................78
Figure 11: Process data measures with significant differences between ‘bad’, ‘medium’ and  
‘good’ performing dyads..............................................................................................82
Figure 12: Significant interactions in the process data measures between the performance  
group and the type of pictures.......................................................................................82
Figure 13: The nine target pictures on the speaker’s display for the four tasks used in study 2.  
..............................................................................................................................................89
Figure 14: Ideal structure of the collaboration process for interactive and non-interactive  
modes of communication...............................................................................................91
Figure 15: Constitution of the support measures with elements prior to and during collaboration

Figure 16: Screenshot of a speaker’s on-screen video (condition ‘support level 1 + 2/ interactive mode of communication’)

Figure 17: Speaker’s collaboration script for the individual (above) and collaborative phase (below) (translated version; condition ‘support level 1 + 2/ interactive mode of communication’)

Figure 18: Overview of the procedure in the different support conditions

Figure 19: Significant effects of the type of pictures on the problems and errors during communication process

Figure 20: Significant effects of the amount of support on the problems and errors during communication process

Figure 21: Significant effects of the mode of communication on the problems and errors during communication process

Figure 22: Significant interactions between type of pictures and mode of communication

Figure 23: Significant effects of the type of pictures (left), the amount of support (center), and the mode of communication (right) on the individual and grounding time

Figure 24: Significant effect of type of pictures on number of correctly placed pictures (left), of amount of support on time needed for description and positioning (center), and of communication mode on number of correctly placed pictures (right)

Figure 25: Significant interaction between the type of pictures and the mode of communication

Figure 26: Number of times the addressee rewound the audio recording by amount of support

Figure 27: (a) Number of pictures placed in the correct position in relation to the time needed for the ‘cat’ task for each of the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures in the correct position

Figure 28: (a) Number of pictures placed in the correct position in relation to the time needed for the ‘kites’ task for the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures on the correct position

Figure 29: (a) Number of pictures placed in the correct position in relation to the time needed for the ‘crochet’ task for the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures on the correct position
List of Figures and Tables

Figure 30: (a) Number of pictures placed in the correct position in relation to the time needed in the ‘circles’ task for the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures on the correct position. ...........................................118
Figure 31: Time needed for description and positioning of each single picture for both dyads (above ‘with support’, below ‘without support’). ..............................................................119
Figure 32: Schematic representation of the collaboration process with support (left) and without support (right). ...........................................................................................................121
Figure 33: Task 1 ‘kites’ with marked feature differences (red arrows). ........................................152
Figure 34: Task 2 ‘radar’ with marked feature differences (red circles and arrows).................153
Figure 35: Task ‘cats’ with marked feature differences (red circles and arrows). ......................153
Figure 36: Task 4 ‘flowers’ with marked feature differences (red arrows). ...............................154
Figure 37: Examples for the two types of use of the pictorial shared application....................155
Figure 38: Examples for the two types of use of the textual shared application.......................156
Figure 39: Task ‘cat’ with marked feature differences (red arrows and circles)......................157
Figure 40: Task ‘kite’ with marked feature differences (red arrows and circles).....................158
Figure 41: Task ‘crochet’ with marked feature differences (red arrows and circles)...............159
Figure 42: Task ‘circles’ with marked feature differences (red arrows and circles).................160
Figure 43: Collaboration script for speaker and addressee in interactive mode of communication/ Level 1 + 2 support for a concrete task. ....................................................161
Figure 44: Hardware configuration in study 1 and 2 .................................................................162

List of Tables

Table 1: Characteristics of concrete and abstract pictures..............................................................63
Table 2: Design of Study 1. ...............................................................................................................66
Table 3: Coding scheme for the analysis of the communication process. ....................................68
Table 4: ICC_{fixed} of the process variables for each of the four tasks in study 1.........................69
Table 5: Means and standard deviations (in parentheses) of the process data. ..........................72
Table 6: Means and standard deviations (in parentheses) of the performance measures.........75
Table 7: Rate of use (per cent) of the pictorial shared application by role of the addressee.....76
Table 8: Rate of use (per cent) of the textual shared application by role of the addressee......77
Table 9: Summary of the results (process data, performance measures, data of shared application). .........................................................................................................................79
List of Figures and Tables

Table 10: Means and standard deviations (in parentheses) of the process data for the performance groups and the type of pictures.................................................................81
Table 11: Design of study 2.................................................................................................................95
Table 12: The four realized sequences of tasks in study 2..............................................................96
Table 13: Coding scheme for the problems and errors occurring during communication........99
Table 14: ICC\textsubscript{just, fixed} of the problem and error categories for each of the four tasks in study 2. ..............................................................................................................................................100
Table 15: Overview of all dependent variables in study 2. ...........................................................103
Table 16: Means and standard deviations (in parentheses) of the process data included in MANOVA. ..............................................................................................................................................104
Table 17: Means and standard deviations (in parentheses) of the performance measures. .....109
Table 18: Summary of the results of study 2 (MANOVA including process data and performance measures). .....................................................................................................112
Table 19: Number of tasks checked again in each condition.......................................................113
Abstract

The aim of this dissertation was to develop effective instructional support for a net-based collaborative problem-solving task. The collaborative task used in the studies was similar to the referential communication task (Krauss & Weinheimer, 1964, 1966) that has often been used to study communication. However, it contained additional individual cognitive demands and was thus more comparable to realistic collaborative tasks: Two persons sitting each in a different room had to solve a picture sorting task together. One participant took the role of the speaker and the other one the role of the addressee. Both participants saw a number of pictures differing only in terms of minor details. The order of the pictures was different for speaker and addressee. The speaker had to describe the pictures and the order on his display to the addressee; the addressee had to arrange the pictures on his display according to the speaker’s description.

First, an overview of theories on the demands of collaboration, in particular theories of communication, and of relevant research on the support of net-based collaboration is presented. Specific research questions were identified and approached in two studies. In both studies two different sets of pictures (concrete and abstract), which differed according to their demands (concrete: more individual cognitive demands and need for mutual spatial perspective; abstract: more demands due to interaction), were used.

The research questions addressed by the first study were: (1) did the postulated individual cognitive demands and the demands due to interaction really have an impact on process and outcome of the collaboration? (2) Were the provided pictorial and textual shared applications useful and what demand did they mainly support? (3) How did the collaboration process of well and badly performing dyads differ? A 2x3 factor design (role of the addressee: able to talk/ not able to talk; availability of shared application: pictorial/ textual/ no shared
Abstract

application) with an additional within-subject factor (type of pictures: concrete/abstract) was implemented. 120 students participated in the study. The analysis of the process as well as the performance measures indeed showed differences depending on the types of pictures. As expected, the process measures showed more individual cognitive demands in concrete tasks, whereas there were no consistent results regarding the demands due to interaction. The process measures reflecting the demands due to interaction showed somewhat higher values in concrete tasks. The overall higher degree of difficulty in concrete tasks was interpreted as an illusion of simplicity (Nickerson, 1999). The shared applications were not perceived as being very useful as not all dyads that had a shared application available used it. However, the shared applications were more used to support the demands due to interaction. The comparison of well and badly performing dyads actually revealed differences in the process: Speakers in dyads with bad performance committed more errors related to both kinds of demands.

For the second study, instructional support measures were developed based on the results of the first study and following the promising approach of Rummel and Spada (2005a, 2005b). The instructional support measures combined a model collaboration provided as on-screen video as well as a collaboration script and consisted of two levels: The first level supported the individual cognitive demands (including hints for feature search and an individual picture editor to mark the features) and the second level supported the demands due to interaction (including hints for the labeling of the features and an individual text editor to note the labels).

The research questions addressed in the second study were: (1) did the instructional support measures including two levels of support have an impact on collaboration process and outcome? (2) Did non-interactive collaboration also benefit from the instructional support measures? A 2x3 factor design (amount of support: level 1 + 2/ level 1/ no support; mode of communication: interactive/ non-interactive) with an additional within-subject factor (type of pictures: concrete/abstract) was realized. 96 students participated in the study. The analysis of process data indeed showed an impact of the support measures. As expected, each of the two supported demands showed an effect on the process measures respectively. However, the measure’s impact did not reflect on the performance measures unanimously, as there were no
significant differences between the numbers of correctly placed pictures for the three support conditions. The impact of the training was not different for interactive and non-interactive conditions, but improved the collaboration process in both conditions.

To summarize: The studies helped to get a better understanding of the demands of collaboration by identifying individual cognitive demands as well as demands due to interaction as relevant factors. Furthermore, a combination of a model collaboration and a collaboration script showed to be a promising approach to foster net-based collaboration.
Zusammenfassung


Es wurden spezifische Forschungsfragen abgeleitet, die durch zwei Experimente beantwortet werden sollten. Diese werden nach einem Überblick über die relevanten Theorien bezüglich der Anforderungen von Kooperation dargestellt. Im theoretischen Teil liegt der Fokus insbesondere auf der Darstellung prominenter Kommunikationstheorien sowie einer Präsentation relevanter Forschungsarbeiten zur Unterstützung netzbasierter Kooperation.

In beiden Experimenten dieser Arbeit wurden zwei verschiedene Arten von Bildern (konkret und abstrakt) verwendet, welche sich in ihren Anforderungen voneinander unterschieden (konkret: mehr individuelle kognitive Anforderungen und Notwendigkeit zur Perspektivenübereinstimmung; abstrakt: mehr kommunikative Anforderungen).
Zusammenfassung


Zusammenfassung


Insgesamt trugen die Experimente zu einem besseren Verständnis der Anforderungen von Kooperation bei, indem sowohl individuelle kognitive als auch kommunikative Anforderungen als relevante Einflussfaktoren identifiziert werden konnten. Darüber hinaus erwies sich eine Kombination aus Modellkooperation und Kooperationsskript als viel versprechender Ansatz zur Unterstützung von Kooperation.
Introduction and Overview

Through the use of new technologies it becomes possible for groups of persons that are spatially distributed to learn or work together or to jointly solve problems. In the last decade, the use of remote technology for collaboration has increased substantially in different working contexts (Olson & Olson, 2000; Whittaker, 1995), as has the amount of research conducted in the field of remote collaboration. A variety of collaborative tasks have been implemented in both field and laboratory studies, ranging from real business tasks with remote teams (see Olson & Olson, 2000), to realistic tasks in laboratory settings such as remote dyads repairing a bicycle (Fussell, Kraut & Siegel, 2000) or solving a clinical case (Rummel & Spada, 2005a, 2005b).

Collaboration – especially in remote settings – poses various challenges and requires different kinds of skills: the challenges of collaboration are added to the individual cognitive demands inherent in a specific task. Furthermore, communication with computers offers opportunities and, at the same time, leads to new difficulties. It is therefore necessary to support remote collaboration.

A promising approach to foster collaboration is instructing collaborating partners on how to solve a remote collaborative task and improving the skills required (the collaborators learn how to collaborate). With this approach, a sustainable enhancement of the collaborative process and results can be achieved (Rummel & Spada, 2005a, 2005b). However, in order to design instructional measures, we need to understand in detail what skills are needed to solve a specific task. Studies of collaboration mostly use realistic and therefore quite complex settings. As the role and functioning of communication is in the center of interest of this dissertation, a remote collaborative problem-solving task was used that is more restricted and enables a more detailed analysis of the processes involved.
Introduction and Overview

One purpose of communication is the establishment of referential identity (Clark & Brennan, 1991). Many collaborative tasks, either in face-to-face or remote settings, require referential communication: If the conversation focuses on specific entities, the persons communicating need to have a mutual understanding of the referring expressions. This is the case, for example, if experts teach novices how to build something or when lawyers and witnesses try to reconstruct a crime in court. Referential communication takes also place if an architect discusses the plan of a new building with his client or when a specialist advises other surgeons how to proceed during remote surgery. In research on referential communication, a so-called referential communication task (Krauss & Weinheimer, 1964, 1966) has been used a great deal. In this task, one person – the speaker – describes a set of different objects or pictures and their order to another person. This other person – the addressee – sees the objects or pictures in a random order and has to arrange them according to the speaker’s description. The task used in this dissertation is similar to the referential communication task, but contains additional higher individual cognitive demands. Through these additional individual cognitive demands the task is more comparable to realistic collaborative tasks.

A first aim of this dissertation was to contribute to a better understanding of the demands and processes underlying collaboration by concentrating on aspects of communication as part of the collaborative demands. Differing from the linguistic methodology of Clark and colleagues, a more cognitive approach was chosen to understand communication processes in more detail. In addition, the focus lay also on individual cognitive skills needed to solve the task. Therefore a first study was conducted to test the impact of the postulated demands on collaboration process and outcome.

A second aim was to design instructional support measures using the results of the first study and following the promising approach of Rummel and Spada (2005a, 2005b). A model collaboration provided as on-screen video combined with a collaboration script were implemented as two-level support measures. This support was evaluated in a second study to identify whether the collaborators learned to collaborate, and more precisely, how they learn
to communicate under conditions of additional cognitive load posed by individual cognitive demands.

This dissertation consists of the following chapters:

**Chapter 1 – Theoretical Background.** The first chapter provides the theoretical background for the two studies included in this dissertation. In the first part, the demands of collaboration in net-based settings are illuminated, with a specific focus on the collaboration theories by Clark and colleagues (e.g. Clark, 1996), Keysar and colleagues (e.g. Keysar, 1998), and Nickerson (1999). A second part describes approaches to support net-based collaboration.

**Chapter 2 – The Collaborative Task.** The second chapter presents the collaborative task used in this dissertation. It describes the specific demands of this task based on the background of the first chapter and the support realized in the two studies.

**Chapter 3 – Study 1: Test for the Influence of Individual and Collaborative Demands on Process and Outcome.** The third chapter describes the method and results of the first study, which aimed at validating the postulated demand model and providing clues for the design of support measures through a detailed analysis of the collaboration process respectively. Furthermore, a first technical support measure – the availability of a shared application – was tested.

**Chapter 4 – Study 2: Instructional Support for Individual and Collaborative Demands.** The fourth chapter first describes the instructional support measures based on the results of the first study and on the promising approach of Rummel and Spada (2005a, 2005b). The bulk of this chapter covers the method and results of the second study, conducted to evaluate the developed instructional support measures.

**Chapter 5 – Overall Discussion.** The last chapter contains an interpretation and discussion of the results of study 1 and 2 in light of theories of communication and net-based collaboration.
Chapter 1

Theoretical Background

This first chapter provides the theoretical background for the two studies included in this dissertation. It contains two main parts:

- **sections 1.1 to 1.3**: the demands of collaboration in net-based settings and
- **section 1.4**: approaches to support net-based collaboration.

Collaboration is a broadly used term and research on collaboration can be found in numerous domains, each focusing on slightly different aspects. For example, the aspects of coordinating work and organising collaboration are salient in work and organizational psychology research, the role of groups and group behaviour are at the centre of interest in social psychology research on collaboration, and processes of collaborative learning are major issues from the viewpoint of educational sciences. Depending on the focus, the definition and intended meanings of collaboration differ. Teasley and Roschelle (1993) define collaboration as “[…] a coordinated, interactive activity that is the result of a continued attempt to construct and maintain a shared conception of a problem” (p. 235).

This definition emphasizes the roles of both coordination and, to a greater degree, communication. Mutual understanding and mutual problem representation are crucial for collaboration. Furthermore, the definition implies a distinction between collaboration and cooperation that has been drawn by several authors (see Dillenbourg, 1999). Cooperation is seen as work accomplished by different persons who solve sub-tasks individually whereas collaboration refers to the mutual engagement of different persons coordinating their efforts to solve the task together. Related to this distinction are the notions of interactive and non-interactive activity: Interactive activity is work done together at the same time and is often related to collaboration. Non-interactive activity is work done at different points in time and is
encapsulated by the term cooperation. In the following chapters, the distinction between collaboration and cooperation is not drawn and the two terms are used interchangeably. For the purpose of this dissertation, Teasley and Roschelle’s definition has been adapted:

Collaboration is a coordinated activity. It is the result of a continued attempt to construct and maintain a shared conception of a problem.

Work carried out in the field of collaborative learning and collaborative work has shown that collaboration is not an easy undertaking but poses several challenges. To collaborate successfully, different kinds of skills are required. In both individual and collaborative working and learning situations, individual actions and individual cognitive processes have to take place (Dillenbourg, 1999). Due to the need to interact, additional activities and cognitive processes are required. The additional demands of collaboration can lead to benefits on learning or problem-solving (e.g. because of the need for meta-cognitive activities). However, collaboration does also bring with it certain challenges, including coordinating the collaboration (Barron, 2000) or pooling unshared knowledge (Stasser & Titus, 1985). A important part of the additional demands on the collaborators arises through the need to communicate. Communication requires the establishment of common ground in order to achieve mutual understanding (e.g. Clark, 1996).

Most work done in the fields of collaborative learning or problem solving does not stress the difference between purely individual cognitive and collaborative activities, but rather distinguishes joint task related activities and communication (see Dillenbourg, Baker, Blaye & O’Malley, 1996). For example, the problem-solving activities themselves and the construction of a representation of the problem (Teasley & Roschelle, 1993) or the construction of a content space and a relational space (Barron, 2003). Van Bruggen, Boshuizen and Kirschner (2003) do indeed group the collaborative problem-solving demands into cognitive and communicative demands. However, the ‘cognitive demands’ comprise individual as well as mutual activities. Despite their heterogeneity, all these taxonomies assign relevance to meta-cognitive activities as needed for constructing a shared representation of a task. To collaborate successfully, individual and mutual task-related as well as meta-cognitive activities have to be
carried out. This is indeed demanding for the collaborators and without adequate support, successful problem-solving and learning cannot be guaranteed (e.g. Barron, 2000; Hogan, Nastasi, & Pressley, 2000).

Through the use of new technologies, it becomes possible to learn collaboratively (e.g. distant learning) or to jointly solve problems (e.g. remote teams in organizations) while the collaborators are located in different places. A variety of collaborative learning settings and collaborative tasks have been implemented, but a common challenge of most remote collaborative settings is that of communicating without a physically shared environment (for an overview see Kraut, Fussell, Brennan, & Siegel, 2002), on top of all the usual challenges of face-to-face collaboration.

The following sections (1.1 to 1.3) describe the three groups of demands in net-based collaboration in more detail.

1.1 Individual Cognitive Demands in Collaboration

The term **individual cognitive demands of a task** refers to the individual cognitive activities that are required to solve a collaborative task. The individual cognitive demands are extremely task-specific. Most tasks require quite a lot of complex cognitive processes on an individual level. Depending on the complexity of the task, they can be defined more or less clearly. In the following, some examples of individual cognitive demands among a variety of complex processes are described for prominent collaboration tasks. Of course, the mentioned cognitive activities are only a brief illustration and do not constitute an exhaustive list.

- In **collaborative learning** (e.g. Slavin, 1995), each collaborator has to read texts or to look at pictures on his own, he has to search for definitions, arguments, new findings etc.

- In **collaborative problem-solving tasks** (e.g. Teasley & Roschelle, 1993), each collaborator has to build an individual problem representation according to known rules, available actions and required goals.
In **collaborative decision making** (e.g. Stasser & Titus, 1985) or other discussion tasks, all information available and needed for the decision or discussion has to be acquired, processed, and evaluated.

In **referential communication tasks** (e.g. Krauss & Weinheimer, 1964, 1966; for more detail see section 1.2.2), where a speaker describes a set of objects to an addressee who has to identify the correct objects, both participants have to conduct a visual search among different objects for the target.

### 1.2 Demands due to Interaction: Coordination and Communication

Two main collaborative demands are the coordination of time processes and activities as well as the communication between collaborators. In this dissertation, the emphasis lies on the communication processes in collaboration. Therefore, this section first provides a brief overview of coordination research followed by a more detailed presentation of theories and research on communication.

#### 1.2.1 Coordination

Malone and Crowston (1990, 1994) developed a **coordination theory** integrating findings from multiple disciplines such as computer sciences, economics, organizational theory and psychology. The authors propose various definitions of coordination, each concentrating on specific aspects of coordination. Here are but a few of the proposed definitions:

Coordination is…

- “The additional information processing performed when multiple, connected actors pursue goals that a single actor pursuing the same goals would not perform” (Malone & Crowston, 1994, p. 112).
- “[…] managing dependencies between activities” (Malone & Crowston, 1994, p. 90).
This variety of definitions shows the complexity of the concept and the numerous facets of coordination. The first definition focuses on coordination as additional demand arising from collaboration, whereas the second definition encloses the whole domain of collaboration. The third definition stresses the fact that the need for coordination arises through the existence of dependencies. In contrast to the first and second definition, it includes also situations, in which one single actor has to coordinate different dependent activities. The term *dependencies* is crucial for Malone and Crowston’s coordination theory, which includes a taxonomy of common dependencies between activities (Malone & Crowston, 1994). In the following, the taxonomy will briefly be presented:

- **Dependency due to shared resources**: If activities share some limited resources, as for example money or an actor’s time, a resource allocation process has to take place to manage the interdependencies between these activities.

- **Producer/ consumer dependency**: A situation where one activity produces something that is part of another activity leads to a producer/ consumer relationship among the activities. Problems of sequencing and tracking processes to assure availability of the needed things or information as well as the transfer of them are related to this dependency.

- **Dependency because of simultaneity constraints**: If activities need to occur at the same time or cannot take place at the same time, there is an interdependency of this third type.

- **Task/ subtask dependency**: If different activities are part of an overall goal, a definition of sub-goals, that is goal decomposition, has to take place in order to perform the task successfully.

These dependencies are characteristics of various collaborative situations. Johnson and Johnson (1992) underline the importance of being aware that coordination is fundamental for successful collaboration. The importance of being aware of the interdependencies was also found by Barron (2000). She identifies three forms of coordination in collaborative problem-
solving groups that could be observed in a well-performing group of students in contrast to a badly performing group:

- **Mutuality of exchanges**: Are the interactions balanced such that all collaborators can contribute?
- **Achievement of joint attentional engagement**: Do the group members share the focus of attention?
- **Alignment of goals**: Do the collaborators have a collaborative orientation to the problem-solving activity (vs. an independent orientation)?

These three forms of coordination correlated with the result of the collaboration, with well-performing groups showing these coordination activities. Another aspect of coordinating collaborative work that proved to have an impact on the result is balancing individual and joint phases in collaboration (Hermann, Rummel & Spada, 2001; Rummel & Spada, 2005a, 2005b). Collaborators tend to ‘do everything together’, but phases of individual work, used to read information, to think about a problem etc., have shown to be important for good collaboration. This is especially important, if the collaborators have complementary expertise in different domains.

### 1.2.2 Communication

Many models of communication concentrate on language production and understanding as autonomous, *unidirectional processes* (see e.g. Ferreira, 2000; Searle, 1992; Levelt, 1989; Marslen-Wilson, 1987). But most models focusing on dialog do describe communication as a collaborative, *bidirectional process* (see e.g. Adler & Rodman, 2003; Clark & Krych, 2004; Horton & Keysar, 1996; Pickering & Garrod, 2004). According to these models, language production by a speaker and language understanding by an addressee are not explicitly separated processes. Both communication partners are sending and receiving massages continuously and alternately take the roles of speaker and addressee. What one person says is influenced by the other person’s utterances. This perspective on communication does not
solely concentrate on one individual but on the interaction between two persons: Communication is understood as the process between both.

The bidirectional models may furthermore be distinguished according to their concept of language design (see Figure 1): Do speakers design their utterances for an audience or do they design it egocentrically? The theory of communication by Clark and colleagues (e.g. Clark & Wilkes-Gibbs, 1986; Clark & Schaefer, 1987; Clark & Brennan, 1991; Clark, 1996) takes the position of audience design (Clark & Carlson, 1982), whereas the monitoring-and-adjustment model of Keysar and colleagues (e.g. Horton & Keysar, 1996; Keysar, 1998) takes the position of egocentric design of utterances. Both models will be described in the following, before the process model of communication by Nickerson (1999) is outlined. The model by Clark has a rather psycholinguistic background, whereas the other two models take more cognitive aspects into account.

**The Theory of Communication by Clark and colleagues.** Wherever two people do something together, common ground is needed to coordinate what they do. Clark (1996) defines common ground as:

“Two people’s common ground is, in effect, the sum of their mutual, common, or joint knowledge, beliefs, and suppositions” (p. 93).
If two persons do something together, their common ground changes: The common experiences add to their common ground (e.g. Clark & Marshall, 1981; Clark & Brennan, 1991). As mentioned above, Clark and colleagues view communication as a collaborative process and therefore common ground is a crucial precondition for communication. To coordinate their communication process, the common ground of two persons has to be updated continuously; otherwise problems in communication may occur. This update process is called **grounding** (Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986; Clark & Brennan, 1991). Grounding is a process by which the persons communicating try to reach the mutual belief of having understood what the other person said. This understanding is not always perfect but depends on the **grounding criterion**.

“The contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for current purposes” (Clark & Schaefer, 1989; p.262).

Grounding occurs as part of the collaborative process of **contributing to conversation** which can be divided into two phases (Clark & Schaefer, 1989; Clark & Brennan, 1991): A **presentation phase** in which the speaker presents an utterance and an **acceptance phase** in which the addressee accepts the utterance, that is, he gives evidence to the speaker that he believes he understands what the speaker has said. Both phases are needed for the contribution to be complete. It is during the acceptance phase that most complications arise and where grounding becomes most evident. Two cases are considered by the authors: the addressee signalizes understanding or he signalizes trouble in understanding. The main types of evidence showing that the addressee has understood are (Clark & Schaefer, 1989):

- **Continued attention:** The addressee continues listening.
- **Initiation of the relevant next contribution:** The addressee initiates the next utterance on the same level as the first one.
- **Acknowledgement:** E.g. the addressee nods, says ‘yes’ or ‘mmh’.
- **Demonstration:** The addressee demonstrates what he has understood.
• **Display:** The addressee displays verbatim the content of the utterance that he has understood.

Of course, the precise way in which grounding takes place depends on the situation, but in general, speaker and addressee tend to minimize the common effort, following the **Principle of Least Collaborative Effort** (Clark & Wilkes-Gibbs, 1986; Clark & Brennan, 1991):

“In conversation, the participants try to minimize their collaborative effort – the work that both do from the initiation of each contribution to its mutual acceptance” (Clark & Brennan, 1991; p. 135).

On the one hand this implies that speakers want to keep their explanations as short as possible, but on the other hand they must make sure that their utterances are understandable; hence a trade off occurs. The goal is to keep the **costs of grounding** as low as possible. These costs depend on and change with the communication medium (for more details see section 1.3). Clark and Brennan (1991) describe the following eleven costs:

- **Formulation costs:** It costs more time and effort for the speaker to retrieve uncommon words, to describe unfamiliar objects, to formulate perfect and complicated utterances.
- **Production costs:** It takes more effort for the speaker to type on a computer keyboard than to speak or point to something.
- **Reception costs:** Depending on the grade of abstraction it may be easier for addressees to listen to utterances or to read them.
- **Understanding costs:** It takes more effort for the addressee to understand certain words or phrases than others.
- **Start-up costs:** These costs arise when starting up a new conversation.
- **Delay costs:** These costs arise in order to plan or revise an utterance.
- **Asynchrony costs:** Costs that arise if the medium of communication is non-interactive.
Chapter 1: Theoretical Background

- **Speaker change costs:** Costs of turn-taking.

- **Display costs:** Costs that are connected with the possibility to gesture in order to display a meaning.

- **Fault costs:** These costs are associated with mistakes. These costs trade off with formulation costs.

- **Repair costs:** The amount of time or effort needed to make a repair.

There are different possibilities for dealing with problems of understanding, related to the fault costs. If the speaker’s meaning and the addressee’s understanding of an utterance do not correspond, then either the speaker himself may notice the problem or the addressee may signal the misunderstanding so that the speaker can repair the utterance (Clark, 1994). But following Schegloff, Jefferson and Sacks (1977) speakers prefer to repair their own utterances and to initiate these repairs themselves.

To further reduce the effort of collaboration in conversation, participants make use of heuristics. Clark and Marshall (1981) distinguish three kinds of heuristics:

- **Community membership heuristic:** If two persons are part of the same community (e.g. cohorts, roles, soccer team, nationality), they share certain beliefs, knowledge or even expertise. A speaker believing an addressee to be a member of the same community, may use words or phrases that are known to this community and do not have to be explained in great detail.

- **Physical copresence heuristic:** If two persons share the same physical environment they can refer easily to objects or events happening in this shared environment.

- **Linguistic copresence heuristic:** If two persons are in a conversation with each other, they can refer to things that have just been mentioned in their conversation.

These heuristics are used to make the grounding process smooth and effective. The nature of grounding depends on perceived group membership, physical environment and the time already spent for the conversation.
Grounding techniques also depend on the purpose and medium of communication. The constraints of different media on the characteristics and techniques of grounding will be described in section 1.3. Clark and Brennan (1991) describe two purposes of grounding that imply different grounding techniques: grounding references and grounding verbatim content. If it is important to remember the verbatim content of an utterance (e.g., to memorize a telephone number, register a name or spell a word) specialized grounding techniques are used.

If references are the content of conversation, that is, if they focus on objects or parts of objects, the participants have all to assign the same meaning to a word and to match this word both to the same object. They must make sure their references on objects are mutually understood: They have to install referential identity. These referential conversations arise for example if an expert is describing to a novice how to build things, when lawyers and witnesses try to reconstruct a crime in court, in a conversation between an architect and his client about a new building, or in remote surgery when a specialist advises other surgeons how to proceed. Generally speaking, it occurs in cases where people have to arrange objects or identify a certain object out of a number of objects. Referential identity may be established for example by using alternative descriptions or indicative gestures. An addressee can show he has correctly identified an object by offering an alternative description of this object. He can also point on the object, touch it, or look at it.

To study this referential communication in an experimental setting, a referential communication task has been used a lot, originally designed by Krauss and Weinheimer (1964, 1966). In this task, one person (a speaker) has to describe objects or pictures to another person (an addressee) in a way that the addressee can identify the described objects or pictures out of a number of objects. Typically, the participants have to arrange the objects several times. Thus, changes in the quality of references can be studied. Krauss and Weinheimer (1964, 1966) found that the mean length of the reference phrase diminished according to the frequency to which a figure was referred. If persons refer to the same objects repeatedly, they tend to use the same terms again. This phenomenon has become known as lexical entrainment (Garrod & Anderson, 1987) or lexical alignment (Pickering & Garrod, 2004).
Alignment has also been found on the level of syntax (speaker and addressee using the same syntactical structures; Branigan, Pickering & Cleland, 2000) or articulation (e.g. Bard et al., 2000) and not only in ‘normal’ face-to-face conversations, but also in written computer-mediated communication (Bromme, Jucks & Wagner, 2005).

According to Brennan and Clark (1996) the reason behind the lexical alignment is that speaker and addressee create a **conceptual pact** when grounding a reference. A conceptual pact is “[...] a temporary agreement about how the referent is to be conceptualized” (Brennan & Clark, 1996; p. 1484) and it leads for example to shorter expressions for particular objects. Once introduced, the participants continue to use the expressions during their conversation because they are aware of the conceptualization they had chosen together. This conceptual pact is established between two specific persons; if one participant of the communication changes, a new conceptual pact has to be established.

**Communication as ‘Monitoring and Adjustment’ (Keysar and colleagues).**

In contrast to Clark and colleagues, Keysar and colleagues claim that language use is egocentric. Children who have not developed a theory of mind yet (including a theory of communication) are known to have difficulties in understanding other’s perspectives (e.g. Wellman, Cross & Watson, 2001; Astington, Harris & Olson, 1988). Sometimes, however, adults also have problems understanding others’ perspectives. Keysar (1994) conducted experiments in which adults had to evaluate the perspective of other persons. He could show that it was very difficult for the participants to distinguish between their own privileged knowledge and the other’s knowledge. Similar to the reaction of children (Olson & Torrance, 1987), they attributed their privileged, egocentric knowledge of a speaker’s actual intention to an uninformed addressee. The author called this phenomenon **the illusory transparency of intention** “[...] because it suggests that once people know the intention behind an ambiguous utterance, it appears to be more transparent that it actually is” (Keysar, Barr & Horton, 1998; p. 47). This phenomenon is of particular relevance in situations where experts and laypersons communicate, since experts do have specific privileged knowledge but are quite often not aware of it (e.g. Bromme, Rambow & Nückles, 2001; Jucks, 2001). It also appears in
communication between native and non-native speakers, as native speakers perceive idiomatic expressions as more transparent as they actually are (Keysar & Bly, 1995).

Because this illusion of transparency is opposed to the predominant concept of audience design, Keysar and colleagues conducted experiments to investigate the cognitive processes during language use in more detail (for an overview see Keysar, 1998).

Horton and Keysar (1996) state that speakers take the addressee’s perspective only as part of a correction mechanism, instead of designing utterances according to the addressee’s perspective from the very beginning. They describe a monitoring-and-adjustment model. According to this model, speakers design their utterances using all information available to them without regard to the addressee’s perspective. This first plan of utterances may rely on non-shared information or not take sufficiently into account the common ground. But speakers monitor their plans and try to detect egocentric plans in order to revise them. According to the model, common ground is part of the correction mechanism.

Under normal circumstances, the egocentric design of utterances does not lead to problems and may not be distinguished from audience design, because of the monitoring-and-adjustment procedure. Nevertheless, egocentric interpretations as in Keysar’s experiments (1994) do occur. According to the author, this is due to an incomplete adjustment for perspective as it has been demonstrated for judgment tasks (Tversky & Kahneman, 1974). This failure of adjustment are expected e.g. under time pressure, when speakers do not have enough time and resources left to monitor and correct their utterances. In this condition, they fall back on their egocentric plans that did not take the common ground into account (Horton & Keysar, 1996; Keysar, 1998).

Keysar and colleagues (Keysar, Barr, Balin, & Brauner, 2000; Keysar, Barr, Balin, & Peak, 1998) also found evidence for addressees interpreting utterances egocentrically. To measure comprehension, the eye movements of two persons solving a referential communication task were recorded (Keysar et al., 2000). Some of the objects the speaker had to describe were visible for both speaker and addressee and some were only visible for the addressee. The addressee knew what objects were only in his but not in the speaker’s view. The eye
movement data of the addressees showed delays in fixations in trials where the critical object could either be an object seen by both or an object only visible for him alone. The authors interpreted the delay as showing uncertainty in the decision to select an object. Even if the addressee had physical evidence about what the speaker saw he still interpreted utterances egocentrically.

**Audience design or egocentric design?** The two positions have been discussed in extensive debate on the research by Clark and Keysar (e.g. Keysar, 1994; Gerrig, Ohaeri & Brennan, 2000; Keysar, 2000; Gerrig, Brennan & Ohaeri, 2001). There is evidence supporting Clark’s audience design hypothesis, like the lexical entrainment effect, but there is also evidence to support Keysar’s egocentric design hypothesis (e.g. Keysar et al., 2000). How is one to make sense of this discrepancy?

Keysar (1998) emphasizes that the main difference between the monitoring-and-adjustment model and Clark’s audience design model is the *level of analysis*. He claims that his own model describes cognitive processes whereas the audience design model describes higher level, interpersonal moves.

The next section describes the process model of communication by Nickerson (1999, 2001) which can cautiously be seen as a kind of synthesis integrating the two hypotheses because it contains both aspects.

**A Process Model of Communication by Nickerson.** Like Keysar and colleagues, Nickerson developed a model of communication containing an anchoring and adjustment procedure (Tversky & Kahneman, 1974). According to Nickerson (1999, 2001), communicating effectively with someone requires having accurate ideas about his knowledge. This is an easy undertaking if the addressee is well known (Fussell & Krauss, 1989). In contrast, having accurate ideas about what a stranger knows is more demanding. To develop an accurate model of knowledge among those one does not know, different steps have to be taken:
(1) The basis for developing a default model of what a random person knows is the $\textit{model of one's own knowledge}$. The unusual aspects of the own knowledge (e.g. due to education, political affiliation, special interests) help to adjust the model and to transform it into

(2) a $\textit{default model of random other's knowledge}$. Taking into account information of the specific person (derived from knowledge of category-membership and prior experience with this person), this model is the basis for constructing a more person-specific model:

(3) the $\textit{initial model of specific other's knowledge}$. This model is continually refined and updated as new information related to the specific other is obtained and results in

(4) a $\textit{working model of specific other's knowledge}$. 

Although the model of one’s own knowledge may be a helpful basis for a first assumption about the knowledge of a random other person, it often leads to problems. Persons have the tendency to assume that others have knowledge they do not actually have (e.g. Nickerson, Baddeley & Freeman, 1987; Fussell & Krauss, 1991, 1992). Nickerson (1999, 2001) describes several complications of knowledge imputations, the most important of them are:

- The $\textit{false-consensus effect}$ refers to the tendency to perceive oneself as more representative of others than one really is.

- The $\textit{curse of expertise}$ means that although experts know they have specific knowledge in a domain they still overestimate what others know.

- The $\textit{illusion of simplicity}$ refers to the impression that something is easy to understand or simple to solve just because one is familiar with it.

Nickerson (1999) suggests several possibilities to cope with the complications and to improve the accuracy of one’s conceptions about what other people know. He mentions for instance the need to be reflective about the assumptions made, to become more sensitive to the privacy of subjective experience and to uncertainty in general. He also claims the need for effective methods to help correcting the hindsight bias.
In research of communication between experts and laypersons, an area where the need for support is strong, there are attempts to support the construction of correct models of the other’s knowledge. For example Nückles, Wittwer and Renkl (2003) developed an assessment tool that helps to improve the expert’s initial model of the layperson’s knowledge by providing more information to them about a layperson’s knowledge. The results showed that the use of this assessment tool did in fact improve communication.

The Special Case of Monolog. Dialog is often viewed as the main setting for language use (see e.g. Pickering & Garrod, 2004; Clark, 1996), therefore most research on communication of the prominent research groups mentioned in this chapter deals with language processing and understanding in dialogs, but there are also some studies providing information about the underlying processes in monolog. Settings that involve forms of monologs are for example lectures or speeches in general (see Clark, 1996) and of course traditional written communication.

Monolog differs from dialog in important ways, as the speaker does not receive feedback and cannot be sure to be understood by the addressee. Speaking in terms of Clark and colleagues, there is no grounding possible in monologs due to the fact that the acceptance phase is missing and the grounding criterion is not definable (Clark & Schaefer, 1989; Schober & Clark, 1989). Lexical entrainment or alignment (Pickering & Garrod, 2004) is also not possible in monologs; speakers tend to produce longer descriptions if receiving no feedback from an addressee (Krauss & Weinheimer, 1966). Speakers need to be more careful in designing their utterances in non-interactive conditions and therefore are likely to use less egocentric perspectives (Schober, 1993).

The comprehension of monolog is also very demanding: Without the possibility to ask for clarifications, the addressee has to resolve occurring ambiguities on his own. The understanding of utterances is more difficult if a person does only listen to a conversation (‘overhearer’) instead of taking an active role (Schober & Clark, 1989) and even more difficult if only listening to a monolog instead of listening to a dialog (Murfitt & McAllister, 2001).
When solving a collaborative task together, non-interactive conditions lead to more errors and increased time required to complete the task (Clark & Krych, 2004).

To conclude, the special case of monolog is a very difficult situation, or as Pickering and Garrod (2004) put it: “Without training in monologue, people are very likely to go off track during comprehension and production” (p. 218).

### 1.3 Additional Demands of Net-Based Settings

Net-based communication poses additional demands on the collaborators because of the restricted possibilities for communication. Clark and Brennan (1991) propose a framework for systematization; it will be used in section 2.1.3 to describe the setting used in this dissertation. The authors list eight characteristics of communication media. The characteristics 1 to 6 are all attributes of face-to-face settings, whereas they are often missing in net-based communication settings.

1. **Copresence**: Do the people communicating share the same physical environment?
2. **Visibility**: Do they see each other?
3. **Audibility**: Do they hear each other?
4. **Cotemporality**: Are messages cotemporarily produced and received?
5. **Simultaneity**: Can messages be sent and received at the same time?
6. **Sequentiality**: Do the turns of the persons communicating stay in a fixed sequence?
7. **Reviewability**: May messages be reviewed later on? (This is not possible in classical face-to-face settings.)
8. **Revisability**: Is it possible to revise a message before the addressee receives it? (This is also not possible in classical face-to-face settings.)

Net-based settings generally impose additional challenges on the collaborators for example because the participants do not share the same physical space. According to Kraut, Fussell and Siegel (2003) physical copresence provides four types of visual information:
(1) Participants’ heads and faces

(2) Participants’ bodies and actions

(3) Task objects

(4) Work context.

Without this information communication is impeded. Compared to face-to-face communication, net-based communication was found, for example, to lead to an increase in the time required to complete tasks (e.g. Dubrovsky, Kiesler & Sethna, 1991), and decreases in group effectiveness (e.g. McLeod, Baron, Marti & Yoon, 1997) and in member satisfaction (e.g. Carey & Kacmar, 1997). Furthermore, it was found to be more difficult to achieve a group decision in net-based communication (Hiltz, Johnson & Turoff, 1986).

However, missing characteristics do not impose necessarily more demands on the communicators: In a study with children who had to solve a referential communication task, dyads in audio-only conditions outperformed dyads in face-to-face conditions (Doherty-Sneddon, McAuley & Bruce, 2000). The authors interpreted this result in the light of theories of memory; visual signals in face-to-face communication may interfere with task performance because they also employ the visual sketch pad (Baddeley, 1986). A distraction effect of high-quality videoconference systems has also been reported by Matarazzo and Sellen (2000). Though, remote communication technologies offer also chances for collaboration. A second framework for systematization is proposed by Rummel and Spada (2005a). It can be used to demonstrate the constraints and opportunities of different communication settings along three dimensions: (1) Communication channels (auditory, visual, text-based), (2) the mode of collaboration (synchronous, asynchronous), and (3) the way relevant material can be made available to the collaboration partner (exchanging documents, application sharing).

The authors describe several impacts these dimensions have on communication. For example, seeing the collaboration partner in addition to hearing him provides non-verbal information and supports turn-taking. In contrast to asynchronous communication, synchronous communication has the advantages of allowing immediate and spontaneous interactions, but
may lead to time pressure and cognitive overload. Shared workspaces allow for reflection on the collaboration process and the outcome; on the other hand they impose additional coordination demands.

Nevertheless, the additional challenges of collaborating through computers make it especially necessary to support remote collaboration. The next section gives an overview of possibilities for support.

1.4 Supporting Collaboration

Extensive research on collaboration has shown that successful collaboration and good results do not arise without adequate support (e.g. Diehl & Stroebe, 1991; Johnson & Johnson, 1992, Slavin, 1995) and that computer-mediated collaboration is particularly challenging (see Bromme, Hesse & Spada, 2005). Three main approaches to foster net-based collaboration are summarized in the following sections: (1) Improving the technical environment for net-based collaboration, (2) structuring the interaction process in order to achieve better collaboration process and outcomes, and (3) enhancing the skills needed to collaborate successfully.

1.4.1 Improving the Technical Environment

The first approach takes the technical environment as a starting point. By improving the media that were not originally designed for remote collaboration and adapting them to the needs of persons collaborating in net-based settings, better results may be achieved (Weinberger & Mandl, 2003).

As one major disadvantage of remote settings is the missing physical copresence, much work has been done in order to compensate for the consequences.

Technologies designed to improve remote collaboration can be classified with regard to their provision of type of visual information identified by Kraut et al. (2003): (1) Participants’ heads and faces, (2) participants’ bodies and actions, (3) task objects, and (4) work context (see section 1.3). The technical environment for remote collaboration can be enhanced, for example, through the use of video conferencing systems, by providing some electronic shared
environment like shared workspaces, external representations or applications, by enabling the visibility of the collaborators’ workspaces, or by supplying technology for remote gesturing.

**Video Conferencing Systems.** Participants’ heads and faces are provided through video conferencing systems. Most studies failed to show benefits of these kinds of ‘talking heads’ video systems (for a review see Whittaker & O’Connaill, 1997).

**Shared Workspaces** provide visual information of task objects and if they are interactive tools also of participants’ actions. But what are shared workspaces? Greenberg (1998) describes non-electronic versions as follows:

“[…] shared physical workspaces (such as whiteboards, control panels, and tabletops) and the artifacts they contain (sketches, controls, documents, structured drawings) act as a stage and offer props for rich person-to-person interaction” (p. 243).

The importance of shared workspaces lies in the facilitation of interaction over the content. Collaborators can gesture, can focus the other’s attention to certain objects, and they can see what the others are doing. Tang (1991) analyzed the process of problem-solving in small teams using large paper sheets as shared physical workspaces and identified three functions of the shared physical workspaces: (1) storing information, (2) expressing ideas, and (3) mediating interaction. Interestingly, expressing ideas and mediating interaction were found to constitute three quarters of all activities. Among the actions accomplished by the collaborators, gesturing played a prominent role (35 % of all actions).

The simplest version of electronic shared workspaces is shared whiteboards on which remote collaborators can draw or write (e.g. Tang, 1991; Whittaker, Geelhoed & Robinson, 1993). Yet, this simple form allows gesturing and enables an easy identification of objects and locations as well as depiction of spatial relations (Whittaker, Geelhoed & Robinson, 1993).

To facilitate net-based collaborative learning a number of tools have been developed; while many are sophisticated shared representations or workspaces, only a few of them will be described in the following short overview.
Bell (1997) distinguishes discussion-based tools (e.g. CSILE, Scardamalia & Bereiter, 1994) and knowledge representation tools (e.g. SenseMaker, Bell, 1997; Belvédère, Suthers & Weiner, 1995) as two variants of argument representation. The Belvédère system for example was designed to support collaborative scientific inquiry, where learners are confronted with challenging problems. Beside other tools, it contains a shared visual workspace where the learners can construct scientific explanations of problems in ‘evidence maps’ (Suthers, Toth & Weiner, 1997). To create such an evidence map, learners have to specify whether their contribution is a principle, a hypothesis, data or unspecified. Furthermore, they can link the contributions (e.g. ‘for’, ‘against’). In the context of research on Belvédère, Suthers (2001) developed the concept of representational guidance. It emphasizes the central and guiding influence of shared workspaces (or ‘representational tools’) on the process of collaborative learning.

Different studies compared the impact of structured, domain-specific shared workspaces with unstructured, domain-unspecific shared whiteboards (e.g. Ertl, 2003; Bruhn, 2000; Fischer, Bruhn, Gräsel & Mandl, 2000). Only one (Ertl, 2003) of the three studies mentioned above could reveal additional effects of the structured workspaces. This demonstrates that situation-specific, structured tools are not, per se, better than classical, unstructured workspaces.

**Video-as-data.** In contrast to classical video conferencing systems, which focus on the collaborators’ heads, the “[...] video image can be used to transmit real-time information about the work objects themselves, and this can then be used to coordinate conversational content among distributed teams [...]” (Whittaker, 1995). Video-as-data systems do exist in a broad variety. In the taxonomy of Kraut et al. (2003) they range from systems providing only visual information of the task objects (e.g. Gergle, Kraut & Fussell, 2004), to systems that additionally provide visual information of the participants’ bodies and actions and the work area (e.g. Nardi, Kuchinsky, Whittaker, Leichner & Schwarz, 1997). According to a number of studies, providing a view of the work area leads to better performance than audio communication alone (e.g. Fussell, Setlock & Kraut, 2003; Gergle et al., 2004). The use of
video-as-data is particularly of relevance for all tasks that need to be solved away from the computer, as for example neurosurgery (Nardi et al., 1997).

**Technology for Remote Gesturing.** Gesturing has been found critical for establishing common ground (Fussell, Setlock, Yang, Ou, Mauer & Kramer, 2004; Clark & Krych, 2004; Tang, 1991). Systems allowing remote gesturing can provide visual information of the types 2, 3 and 4 in the taxonomy described above; but they may differ according to the types of gestures they support. For example, Fussell et al. (2004) distinguish pointing gestures and representational gestures; the first being used to locate objects and the second to represent the form of objects. Remote gesturing has been realized in a variety of ways, for example through robots (GestureMan; Kuzuoka, Oyama, Suzuki, Yamazaki & Mitsuishi, 2000), video-based representations of hands (Agora; Kuzuoka, Yamashita, Yamazaki & Yamazaki, 1999), or video-based sketching (DOVE; Ou, Fussell, Chen, Setlock & Yang, 2003). The robot GestureMan (Kuzuoka et al., 2000) for example allows only pointing gestures with a laser pointer, whereas representation of hands or video-based sketching enables more types of gestures. One problem that has been identified with remote gesturing technologies is the emergence of so-called *fractured ecologies* (Luff, Heath, Kuzuoka, Hindmarsh, Yamazaki & Oyama, 2003): If the remote collaborators have different views on the task, if the surroundings or ‘ecologies’ are very distinct or separated locally from one another, problems of understanding are likely to occur. For example, collaborators in the DOVE system (Ou et al., 2003) need to connect the information of the other’s work presented on a separate monitor with their own workspace. One attempt to overcome these problems is the Gesture Projection System developed by Kirk and Fraser (2005), in which the gestures of one person are directly projected onto the workspace of the other person. Kirk and Fraser (2005) used this system for a task requiring remote instruction of how to construct LEGO© objects and could show long-term benefits of remote gesturing in instruction compared to audio-only instruction.
1.4.2 Structuring the Interaction Process: Collaboration Scripts

A second approach to foster net-based collaboration supports the interaction process as it occurs through providing a structure. **Collaboration scripts** are a common method for structuring interaction, and exist for a broad variety of purposes and in manifold forms. They help to reduce the demands of coordination between collaborators. As the three following definitions of collaboration scripts show, it is a rather general concept with a huge number of possible applications:

“A script is a story or scenario that the students and tutors have to play as actors play a movie script. Most scripts are sequential: students go through a linear sequence of phases” (Dillenbourg, 2002; p.71).

“[…] in the broadest sense we refer to them as an instructional means that directly affects collaborative learning processes by inducing certain activities rather than indirectly by manipulating context variables of collaborative learning” (Kollar, Fischer & Hesse, in press).

“The central idea of implementing a collaboration script is to foster fruitful collaboration by externally structuring the interaction process. Collaborating partners are guided through a sequence of interaction phases. […] The script is expected to prompt cognitive and social processes by participants that might otherwise not occur” (Rummel & Spada, in press).

Before describing exemplary collaboration scripts from different research areas, two classification systems for collaboration scripts are presented: one by Kollar, Fischer and Hesse (in press) and one by Dillenbourg and Jermann (in press). However, both classification systems do hold some problems: The classification by Kollar and colleagues is restricted to collaboration scripts used for collaborative learning scenarios and the classification by Dillenbourg and Jermann confounds components of the script itself and the persons using it. However, they are both presented to shed light on different aspects of collaboration script.
Kollar et al.: Components of Scripts. The classification of Kollar and colleagues is restricted to collaboration scripts used for collaborative learning scenarios. They identify six central components of those scripts:

- **The learning objectives:** Scripts are used to reach specific learning objectives. These objectives may vary between tasks and may include cognitive, metacognitive, motivational, or emotional changes.

- **The types of induced learning activities:** The concrete instructions provided by scripts are expected to trigger understanding and help reach the learning objectives. Furthermore, they may facilitate later recall and the use of information.

- **The sequencing procedures:** Scripts split activities into subtasks and define what subtask has to be carried out by whom at what point in time.

- **The role distribution:** The script assigns roles to the collaborators and defines the activities each collaborator has to engage in.

- **The type of representation:** The script may be presented in written, graphical or oral form to the collaborators.

- **The locus of representation:** Scripts may be presented externally e.g. in the learning environment or be merely internally represented in the learner’s cognitive system.

Dillenbourg and Jermann: Diversity of Scripts. Dillenbourg (2002) provides attributes and dimensions of CSCL (Computer Supported Collaborative Learning) scripts forming the ‘syntax and semantics’ of scripts (for a discussion of them see Rummel & Spada, in press). In a recent paper, Dillenbourg and Jermann (in press) describe on a more abstract level dimensions from which scripts differ from each other:

- **Cultural versus didactic scripts:** Cultural scripts are implicit scripts that have been acquired in the cultural environment someone lives in. The well-known restaurant-script of Schank and Abelson (1977) is an example of such a cultural script. Collaboration rules set up by teachers are examples of didactic scripts. Both kinds of scripts often interact and mutually influence the behavior of collaborators.
Chapter 1: Theoretical Background

- **Ideal, mental or actual scripts:** The actual collaboration is influenced by these three types of scripts. The ideal script is prescribed by teacher or environment. The mental script is the mental representation built from this prescription. The actual script contains the interactions that the collaborators actually realize. According to Dillenbourg and Jermann, the gap between ideal and actual scripts depends on several script attributes:
  - **The degree of coercion:** A script can contain just some initial instructions or can require the collaborators to perform one step after another. Dillenbourg (2002) defined five levels of coercion. High coercion scripts reduce the distance between ideal and actual scripts but may lead to motivational losses caused by overscripting.
  - **The intelligibility of the script:** If the script is too complex, a clear mental script cannot be constructed.
  - **The granularity of the script:** The duration of each phase and the grain size of subtask definition in the script have to fit to the level of granularity in naturally solved tasks in order for the actual script to reach the level of the ideal script.
  - **The fitness:** The roles distributed by a script have to match the group members’ skills and knowledge to get an actual script close to the ideal script.

- **Macro and micro-scripting:** According to the authors, micro-scripts act on the process of interaction and reflect a more psychological perspective. In contrast to macro-scripts, they are expected to be learned by the collaborators. Macro-scripts influence the process more indirectly and are merely a method to be followed by the collaborators. They reflect an educational perspective.

- **Objective versus method:** This dimension is closely intertwined with the macro-/micro-scripting dimension. If a script is meant to be learned by the collaborators in order to reuse it in other situations, the objective is to internalize a correct script. If a
script is designed only for one situation and is not expected to be learnt by the collaborators, it is a method.

- **Pre-scripting versus post-scripting**: Most scripts define a priori how collaboration should take place. In order to reflect on the collaboration and to internalize a script, the collaborators may also analyze recordings of their collaboration using a script (Harrer, Bollen & Hoppe, 2004).

Additionally, the authors mention the risk of **overscripting** (Dillenbourg, 2002): If natural interactions or problem-solving processes are disturbed by the script (the script can interfere with the learning procedure or trigger ‘goalless’ interactions) motivational losses can be the result (e.g. Bruhn, 2000; Kollar, 2001).

The dimension ‘ideal, mental or actual script’ refers to the effectiveness of a collaboration script and tries to systematize the components influencing the actual behavior. As mentioned above, this systematization slightly falls short. To describe the mechanisms influencing effective learning from scripts (Salomon, 1993), the component processes of observational learning by Bandura (e.g. 1977, 1986) are presented in the following section and applied to collaboration scripts.

**Processes influencing the effectiveness of scripts.** Bandura (e.g. 1977, 1986) identified four component processes influencing observational learning. These components will here be applied onto collaboration scripts in order to find factors influencing their effectiveness.

- **Attentional processes**: They determine what the collaborators selectively perceive and what they extract from the collaboration script.

- **Cognitive representational processes**: To remember the instructions of the script, an active process of restructuring information and transforming it into memory codes is necessary.

- **Behavioral production processes**: The symbolic memory representations of the script elements have to be translated into action. This takes place through a
conception-matching process which compares the adequacy of the produced actions to the conceptual model – in this case the representation of the collaboration script. If mismatches are detected, the actual behavior is corrected toward the represented conceptual model.

- **Motivational processes**: The performance of the behavior intended by collaboration scripts is influenced by three types of motivators: direct, vicarious, and self-produced. The likelihood of the intended behavior being shown depends on the benefit and costs the collaborators anticipate for demonstrating the intended behavior.

In the following, some examples of different collaboration scripts will be described in order to illustrate the various possibilities of using scripts. Most of the examples are scripts for collaborative learning; however, there are also examples for the use of collaboration scripts beyond learning contexts (Härder, 2003; Nückles & Ertelt, in press).

The dimensions of Dillenbourg and Jermann (in press) will then be used to classify the approaches. For a detailed discussion of the different scripts in the light of the components see Kollar et al. (in press). Both classifications will be used in chapter 2 to describe the script used in study 2 of this dissertation.

**Dansereau and O’Donnell: Scripted Cooperation.** The prominent approach by O’Donnell and Dansereau contains the so-called MURDER script (Mood-Understand-Recall-Detect-Elaborate-Review; Dansereau, 1988; O’Donnell & Dansereau, 1992; O’Donnell, 1999). It aims at helping dyads of students working face-to-face to learn from written texts. The text is broken into sections and the steps of the script have to be accomplished for each of these steps. The roles of the learners (e.g. recaller and listener) switch from section to section. After individually reading the text (U) one learner recalls the content (R) and the other monitors the explanation in order to detect errors (D). Now, they elaborate the content together (E). These steps are repeated for each section. At the end they review the whole text together (R).

Other well-known scripting approaches for face-to-face collaboration are the *reciprocal teaching technique* by Palincsar and Brown (1984), also designed to support students in text
processing, or the \textit{guided peer questioning} approach by King (e.g. 1999), which aims at supporting discussion of learning material.

As collaboration scripts have shown themselves to be a very effective means to foster collaboration, they have also been transferred to computer settings. Typically, they are embedded into computer-based learning environments. The following examples of scripts have all been developed to support computer-mediated collaboration.

\textbf{Scripts Structuring Dialogs.} Hron, Hesse, Reinhard and Picard (1997) implemented a collaboration script that structured a dyad’s interaction during a text-based problem-solving task. The task consisted of correcting graphical diagrams that contained errors. The script led through a dialogue consisting of correction propositions from the collaborators and their approval or disapproval of each other’s contributions. Manipulation on the diagram could only be made if the dyad had agreed on a correction.

Baker and Lund (1997) developed a collaboration script supporting two learners in communicating effectively in order to solve physical construction tasks together. They communicated via a chat-tool and had a shared diagram available. The script provided the learners with a set of buttons, so-called communicative act buttons, each related to a specific act: (1) Some buttons were directly connected to the construction area, (2) others helped the learners in reaching a consensus, (3) managing their interactions, or (4) doing something else than communicating (e.g. ‘read the handout’). The buttons either sent complete sentences to the chat tool (‘Where do we start?’) or just starters for contributions (‘I propose to…’).

Pfister and Mühlpfordt (2002) also designed a tool to support chat-based learning: Their \textit{learning protocol} forced the learners to make explicit what previous message they were referring to, provided them with a list of contribution types, such as questions, explanations or comments, and regulated the sequence of contributions.

\textbf{Interaction-related and Content-related Structuring Tools.} Weinberger and colleagues (e.g. Weinberger, Fischer & Mandl, 2002, 2003) as well as Ertl and colleagues (e.g. Ertl, Reiserer & Mandl, 2002, 2005; Ertl, Kopp & Mandl, in press) compared interaction-
related scripts with content-related structuring tools in numerous studies. The first type of script is equivalent to many of the scripting approaches (e.g. those described above), while the second type supports strategies for specific learning domains. These two types of tools were analysed in different settings (text-based communication, e.g. Weinberger et al., 2002, 2003; video-conferencing scenarios, e.g. Ertl, Reiserer & Mandl, 2002) as well as with different types of tasks (collaborative teaching, Ertl, Reiserer & Mandl, 2002; collaborative problem-solving, Kopp, Ertl & Mandl, 2004).

**External Scripts and Internal Scripts.** Kollar, Fischer and Slotta (2005) investigated the interaction of high versus low structured external scripts with the learner’s internal scripts on argumentation. The external scripts were embedded within a net-based collaborative inquiry learning scenario. The low structured version consisted only in asking the learners to discuss different topics. The high structured version defined roles, e.g. which learner first had to produce a new argument and provided sentence starters. The authors examined the role of the learner’s prior knowledge on argumentation (internal script), on the acquisition of new knowledge in a specific domain (biology/science), or on argumentation.

Scripts may not only be used for collaborative learning as in the examples described above, but also for a broad range of collaborative tasks. The two following approaches used scripts to support knowledge exchange (Härder, 2003) and expert-layperson communication (Nückles & Ertelt, in press). Scripts have also been used to enhance net-based collaboration between experts from different domains (Rummel & Spada, 2005a, 2005b). This approach will be described in more detail in the next section (1.4.3).

**Scripts Providing Rules for Information-pooling.** The collaboration script developed by Härder (2003) was meant to support remote dyads communicating via a desktop videoconferencing system in exchanging knowledge, drawing inferences and constructing an appropriate solution in a hidden-profile task. The script consisted of four general principles: (1) Exchanging information, (2) pooling unshared knowledge and drawing inferences, (3) recapitulating important information and (4) justifying the solution. This collaboration script
had a relatively low degree of coercion (Dillenbourg, 2002) because the participants were not forced to follow the principles but received them merely as hints.

**Scripts Supporting Expert-Layperson Communication.** Nückles and Ertelt (in press) designed a problem formulation script in order to support laypersons in describing their computer problems to experts in Internet-based helpdesks. The script consisted of four prompts: (1) asking the layperson to provide an explicit and detailed problem description; (2) prompting the layperson to explain the goal of their computer use; (3) asking for the layperson’s previous actions and for what can be seen on the monitor currently; and (4) encouraging a speculation about the probable cause of the problem. Two versions of the script were implemented: One sequenced version that requires answering each of the four prompts separately and successively and one non-sequenced version showing all four prompts in a list.

The approaches described so far can be classified in the dimensions of Dillenbourg and Jermann (in press) as follows: All scripts are in principle didactic scripts but contain, to some extent, culturally shared knowledge or refer to some cultural script. The script used by Härder (2003), for example, refers to some culturally acquired script of how to exchange information. The authors of the problem formulation script explicitly highlight the relation of their script to some culturally acquired knowledge of how to communicate with experts: “[…] this ‘problem formulation script’ is based on the idea that – despite the lack of domain-specific knowledge – laypersons can draw on familiar general heuristics from everyday problem solving […]” (Nückles & Ertelt, in press).

The attribute **degree of coercion** plays a crucial role concerning the gap between ideal and actual script. For the scripts structuring dialogs they can be ranked as follows: The script by Baker and Lund (1997) holds the least degree of coercion, followed by the script of Pfister and Mühlpfordt (2002), and finally the script of Hron and colleagues (1997), which contains the highest degree of coercion. The low and high-structured scripts employed by Kollar et al. (2005) as well as the sequenced and non-sequenced scripts of Nückles and Ertelt (in press) differ according to their degree of coercion.
With regard to the category *pre-scripting versus post-scripting* all scripts described here are pre-scripting approaches. Furthermore, they all fit into the class of micro-scripting and thus mostly have the objective to provide new strategies for collaboration in more than the scripted situation alone. Scripting as a method (versus an objective) may be the case for the content-related structuring tool of Weinberger and colleagues (e.g. Weinberger et al., 2002, 2003; Ertl et al., 2002, 2005), but most approaches use scripting as objective. Interestingly, the long-term effects of scripting, that is, if the collaborators do indeed learn how to collaborate, have not been studied by most of these approaches. The following section describes such an instructional approach: Rummel and Spada (e.g. 2005a, 2005b) used the skills as a starting point to foster net-based collaboration.

### 1.4.3 Enhancing the Skills: Model collaboration.

In order to promote the skills needed for collaboration, instructional measures that are presented to the collaborators *prior* to the collaboration itself (versus classical scripting approaches that support collaborators *during* collaboration) may be used. Rummel and Spada (2005a, 2005b) provided such an instructional support for net-based collaboration with a desktop video-conferencing system in order to enhance skills for collaboration. In their studies, two experts from different domains (one medical and one psychology student) had to solve clinical cases together. Three different learning conditions were implemented: (1) A model condition promoting observational learning from a worked-out example of computer-mediated collaboration, (2) a script condition enabling learning from scripted computer-mediated collaborative problem-solving, and (3) an unscripted condition were ‘learning-by-doing’ could occur. The studies consisted of two phases: A learning phase where the three learning conditions were implemented and an application phase in which the collaboration took place. The model as well as the script contained instructions regarding the ‘macro level’ of collaboration, e.g., time management and coordination of work, the ‘micro-level’ of collaboration (i.e. communication), e.g., giving feedback and turn-taking procedures, and domain specific instructions.
The idea behind the model condition was the following: While observing a model person solving the task, people should reflect upon the solution steps and engage in meta-cognitive activities that promote learning (e.g. Renkl, 1997; VanLehn, 1996). The model collaboration was realized as a multimedia presentation (audio recording and animations) on the participant’s monitor. Furthermore, it was accompanied by instructional explanations in text boxes as well as prompts for self-explanations that appeared on the screen. The script was provided as paper sheets and contained the same sequence of phases as the model condition. In contrast to the scripting approaches cited above, the script was not presented during the actual collaboration, but in the learning phase prior to collaboration. The third learning condition received no instructional support; the dyads could collaborate freely during the learning phase. A forth condition was implemented as control condition: The dyads had no learning phase at all, but collaborated only during the application phase.

As dependent variables three types of data were used: the collaborative problem-solving process, the joint solution, and the performance on a knowledge post-test containing domain-specific aspects of the collaboration as well as metaknowledge on characteristics of good collaboration. Overall, both instructional support measures led to better results in terms of the dependent variables than the control condition. With regard to the learning objective, the instructional measures could not only be shown to promote collaboration in a specific situation, but also to increase the metaknowledge on good collaboration. This raises the hope that future collaboration may also benefit from the instructional support received.
Chapter 2

The Collaborative Task

This chapter presents the collaborative task used in this work. It describes the specific demands of the task and the support realized in the two studies.

Due to the complexity of realistic settings, in studies of remote collaboration it is often difficult to conduct a detailed analysis of the processes involved. For example, the instructional measures used by Rummel and Spada (2005a, 2005b) concentrated mostly on the coordination aspects, e.g. balancing phases of individual and collaborative work, and showed significant effects on the outcome of the collaboration. However, due to the complexity of the task, it was not possible to conduct an in-depth analysis of the task-specific processes involved.

In this dissertation, a remote collaborative problem-solving task was used that is more restricted and enables a more detailed analysis of the demands involved. In the broadest sense, the task can be classified into a “general class of ‘mentoring’ collaborative physical tasks” (Kraut et al., 2003; p. 16) because one person manipulates objects – in this case pictures – and receives instructions from another person about where to place the pictures. The structure of these tasks is similar to the referential communication task (Krauss & Weinheimer, 1966), which has often been used to study communication (see section 1.2.2). In addition, though, realistic collaborative physical tasks also hold individual cognitive demands such as searching for mistakes or solving problems. The task used here therefore contained both higher individual cognitive demands and demands due to interaction.

In the two studies included in this dissertation, two persons had to jointly solve a picture-sorting task while being spatially distributed. One of the participants assumed the role of speaker and the other took on the role of addressee. The task was presented on two displays.
(see Figure 2), and oral communication between speaker and addressee was possible via an audio link.

Figure 2: The speaker’s and the addressee’s display (here one task from study 2).\textsuperscript{1}

On the speaker’s display, a number of pictures (9 in study 1 and 16 in study 2) were presented that differed only in terms of minor details. The speaker had to describe the pictures and their location to the addressee. The addressee saw the single pictures in a random order and had to arrange them according to the speaker’s description. The addressee could rearrange the pictures on the target area by using his mouse (drag and drop). Because the differences between the pictures were very small, the participants first had to search for them. This

\textsuperscript{1} The pictures have been developed based on pictures of the children’s game “Differix” from RAVENSBURGER.
component of feature detection constitutes the main difference from the classical referential communication task, in which the task demand consists only in the verbal description of clearly different pictures or objects. This additional individual cognitive demand makes the task more comparable to realistic collaborative tasks in which communication often has to take place in parallel to individual cognitive processes.

The task is furthermore comparable to collaborative settings with distributed resources or distributed skills, because the speaker has information about the required target-order of the pictures that the addressee has not. Therefore it is not adequate to facilitate the apparently easiest way to support the collaboration: making it possible for the addressee to see the target order of the pictures.

2.1 The Task Demands

As proposed above, the basic task demands can be grouped into individual cognitive demands and demands due to interaction. The first group of demands includes feature search and identification. These demands have to be dealt with individually; they would also exist if the sorting of the pictures had to be done by an individual instead of being part of a collaborative task. The original children’s game Differix by Ravensburger©, from which the pictures were taken, requires each player to arrange nine pictures on a template faster than one to five other players. The second group of demands contains the additional challenges of collaboration due to the need to communicate. Both demands have to be faced by speaker and addressee, albeit in a slightly different way.

In section 4.2.4 the demands arising at each step undertaken to solve the task are exemplarily shown for two dyads.

In the rest of this chapter, the specific demands arising in the collaborative task used here will be outlined and the kind of support assessed will be presented.
2.1.1 Individual Cognitive Demands: Visual Search and Object Recognition

The individual cognitive demands of the collaborative task used here consist in detecting the features that differ between the pictures and locating these features in a picture. These individual cognitive demands have to be faced by both speaker and addressee. Theories of object recognition such as the recognition-by-components theory of Biederman (e.g. 2000) as well as theories of visual search such as the feature integration theory of Treisman and colleagues (e.g. Treisman & Sato, 1990) can help to explain these kinds of demands.

The recognition-by-components theory allows the prediction that features of an object can be identified faster if they represent basic components of the object compared to features that are only a part of a component. In addition, research in the domain of object recognition showed that known objects can be identified more quickly than unfamiliar ones but details can be overlooked more easily in known objects (compare the influence of top-down processes in object recognition, e.g. Marr, 1982).

According to theories of visual search, the time needed to detect a feature increases with the number of objects in a picture if the objects differ from each other in more than one dimension (e.g. color, size, form). In this case, attention has to be allocated and serial processes are involved. Serial search, as opposed to parallel search, means that each single item or part of an object in a picture has to be examined in succession and at a rate of not more than one item per time unit. The search for and among unfamiliar objects in particular requires more attention as such objects capture the attention and cause longer fixation durations (Wang, Cavanagh & Green, 1994).

Taking together the issues of object recognition and visual search it can be concluded that a serial (vs. parallel) search has to take place in order for the relevant features in the pictures to be identified. This requires the allocation of attention and is a time-consuming and error-prone process (Treisman & Schmidt, 1982). The time needed to detect the relevant features depends on their size (whether they are a basic component or just a part of it) and on the familiarity of the objects. With familiar objects, details can be overlooked more easily.
2.1.2 Demands due to Interaction: Installing Referential Identity

The speaker has to transform the detected features that allow a differentiation between the single pictures into spoken language and the addressee has to understand these descriptions and match them to the objects and their features. The verbal descriptions given by the speaker therefore need to enable an unambiguous identification of each single picture. The purpose of the communication between the speaker and the addressee is to establish referential identity (Clark & Brennan, 1991; see section 1.2.1). In the task used here, referential identity is installed if an addressee has correctly identified the object the speaker described. This is only possible if both assign the same meaning to an expression and both match the expression to the same referent. For the collaborative task used here, referential identity has to be assured not only for the concept and name of a feature but also for the verbal description of the feature’s spatial position. The frame chosen by the speaker for the spatial descriptions may be either viewer-centered (deictic) or object-centered (intrinsic) (Miller & Johnson-Laird, 1976; Levelt, 1989; Taylor & Tversky, 1996). To enable a correct identification of the pictures, the communication partners must have a mutual frame of reference. The viewer-centered perspective is preferred by most speakers (e.g. Levelt, 1989; Miller & Johnson-Laird, 1976) unless the scene or picture does not influence them to use the other frame of reference. Miller and Johnson-Laird (1976) could show that addressees tend to understand descriptions from an intrinsic (object-centered) perspective if the picture contains certain landmarks that imply an intrinsic usage.

2.1.3 Additional Demands of Remote Collaboration via Audio-link

The net-based setting used in this project, in which two persons remotely collaborate and communicate via an audio-link, differs from face-to-face settings in several aspects. As mentioned above, the characteristics of net-based settings hold several challenges but also some new possibilities for collaboration. The criterions listed by Clark and Brennan (1991; see section 1.3) will now be used to characterize the setting used in this work:
The collaborators did not share the same environment as they were located in different rooms. They were not physically copresent and hence could not support their descriptions through gestures, they could not see what the other was looking at, nor could they guide the other’s attention through deictic gestures. Assuring referential identity was therefore more complicated than in face-to-face settings (see also Stahl, 2003).

Additionally, the collaborators could not see each other and were accordingly unable to transmit nonverbal information. This again impedes the grounding process.

The setting used in this dissertation enabled audibility: speaker and addressee communicated through an audio link.

In the first study, there was no delay between production and reception of the descriptions and messages could be sent and received at the same time, but in the second study there were conditions that actually imply such a delay and a lack of simultaneity (non-interactive conditions; see Chapter 4). The effects on grounding will be discussed in Chapter 4.

In both studies, the turns of speaker and addressee could not get out of sequence.

Looking at study 1 and 2, only four conditions in study 1 allowed reviewability: Either a pictorial or a textual shared application was provided, in which part of the communication and problem solving process may be reviewed later on (see Chapter 3).

Unlike in ‘normal’ face-to-face conversations, descriptions could not be revised in any of the conditions.

### 2.2 Support Measures Implemented

As described in section 1.4 there are different ways to enhance net-based collaboration. The aim of this dissertation is the development and evaluation of instructional support measures for the net-based collaborative problem-solving task described in section 2.1. One initial support measure has been tested in study 1: the use of shared applications (see Chapter 3). Furthermore, study 1 aimed at gaining information as a basis to develop effective instructional
support measures including learning from observation of a collaboration model and scripted collaboration. These instructional support measures were then tested in study 2 (see Chapter 4).

2.2.1 Study 1: Shared Applications

In the first study, the availability of shared applications was varied (Bertholet & Spada, 2005). Either the dyads had a pictorial or a textual shared application or they had no shared application available. The pictorial shared application consisted of a picture that was presented on a Netmeeting-Whiteboard. The picture presented in the Whiteboard corresponded to the pictures of a task. However, this picture had another feature combination in order to differ from the pictures that had to be described and arranged. With this tool, the features differing between the pictures can be marked and kept available for description and identification, and, furthermore, deictic gestures are possible that enable an easy identification of the pictures described.

The textual shared application was realized through a simple text editor that could be seen on both displays. Information entered into this text editor could be referred to at any time. In terms of Clark and Brennan (1991), it allowed for review.

These shared applications could be used in two different ways: either supporting the individual cognitive demands by marking or noting the relevant features in general (‘What are the differing features amongst the pictures?’), or supporting the demands due to interaction (‘What are the feature values of picture 1?’). In the latter case the dyad could use the shared application for the description of each single picture, e.g. by using the deictic function of the application to support the establishment of referential identity.

2.2.2 Study 2: Model for Observational Learning from Model Collaboration and Collaboration Script

In the second study, instructional support measures that were developed using the results of study 1 were evaluated. Following the approach of Rummel and Spada (2005a, 2005b), a model collaboration was developed. The dyads received this model collaboration as on-screen
video with audio instructions prior to collaboration (Hansen & Spada, 2006). Furthermore, a collaboration script reminding the collaborators of what they had just learned before was provided during collaboration. The model instruction as well as the script contained two levels of support, each corresponding to one of the two demands. Both levels contained hints for dealing with the respective demand in the best possible way and for each level one subtask was introduced (search for features and marking them in an individual picture editor as well as labeling of features and writing them down in an individual text editor). As the task contained individual cognitive demands in addition to those due to interaction, it can be divided into individual and collaborative phases. The collaborators first had to search for the feature differences on their own and then the speaker had to find names for the features before engaging in the description and positioning of the pictures. By making this division explicit, the instructional support measures were supposed to help overcome the collaborators’ tendency to solely engage in joint activities and forget about individual work phases (Hermann, Rummel & Spada, 2001).

The two classification systems for collaboration scripts described above will now be used to further describe the instructional measures (model and script) used in study 2. In terms of Kollar and colleagues (in press) the following components can be identified:

**The learning objectives:** The instructional support measures aim at improving the skills for collaboration. The collaborators should learn to deal with the individual cognitive demands of searching for features and the demands due to interaction, namely the establishment of referential identity.

**The types of induced learning activities:** The script reminds the collaborators of the different phases and subtasks introduced in the model and includes some crucial hints. Through the instructional measures, the collaborators are expected to reflect on the solution steps and to internalize them in order to better deal with the demands of the task.

**The sequencing procedure:** The instructional support measures include two phases (an individual and a collaborative phase; see Rummel & Spada, 2005a, 2005b) and three subtasks (search for features, labeling of features, grounding and description/positioning of pictures).
Yet, rather than strictly controlling the accomplishment of each subtask and imposing specific fixed actions, the benefits of proceeding according to the given instructions are explained to the collaborators, leaving them the choice to follow the instructions or not. Additionally, the subtasks and hints for solving them are presented in the upper part of the correspondent display.

**The role distribution:** The roles of speaker and addressee are randomly assigned to the two collaborators and are fixed. There is one script for each role.

**The type and locus of representation:** The instructions (division of task into subtasks, hints) are provided as part of the collaboration model incorporated in the multimedia presentation (on-screen video and audio). The script during collaboration is presented to the collaborators in a written form on their displays.

In the dimensions of Dillenbourg and Jermann (in press), the instructional measures of study 2 can be characterized as follows:

It is a *didactic script* with a rather *low degree of coercion*. The subtasks (individual search for features differing between the pictures and individual labeling of the features as well as the common part of picture description and positioning) should be performed in a given order but the completion of the subtasks is not controlled. During collaboration, the sequencing of the task is realized through two different screens with or without individual editors. Additionally, hints and instructions are displayed as short phrases on the top of the displays to remind the dyad what has been included in the on-screen video model. Due to this reminder function, the task is *not very complex* and allows for easy construction of a mental script. The *granularity of the script fits* the granularity of the task to be solved naturally quite well because it does not make any restrictions on the time spent on each subtask. Concerning the *fitness of the roles*, it cannot be assured that each participant gets the role he best suits because a random assignment of roles takes place.
Chapter 2: The collaborative task

The script is expected to be learned by the collaborators (script as *objective*) and is therefore a *micro-scripting* approach. Furthermore, it defines a priori and during collaboration what should happen (*pre-scripting*).

The following Chapters 3 and 4 present the two studies conducted, followed by an overall discussion of the results in chapter 5.
Chapter 3

Study 1: Test for the Influence of Individual and Collaborative Demands on Process and Outcome

The first study aimed at providing clues for the design of support measures through a detailed analysis of the collaboration process. It should have tested if the postulated demands reflected on the collaboration process and outcome (see Bertholet & Spada, 2005). Furthermore, a first – technical – support measure (the availability of a shared application) was tested. The following research questions were addressed:

- Do the collaboration process and the result of the collaboration differ with concrete and abstract types of pictures (both varying according to the individual cognitive demands and those due to interaction)?

- How are the shared applications used? What demand do they mainly support? Does the use of shared applications have an impact on the result?

- How is the collaboration process of well performing and badly performing dyads characterized? What are the main differences between well and badly performing dyads?

As the task materials were taken from the children’s game Differix from Ravensburger © and not primarily designed to be part of scientific research, the discussion of the first study also provides information on how to improve the material.

3.1 Method of Study 1

The study aimed at gaining a better understanding of what skills are needed to solve an experimental remote collaborative task successfully. The study examined the impact of
individual cognitive demands as well as demands due to interaction on process and performance measures of collaboration. For this purpose, different sets of pictures were used that differed according to their demands: concrete and abstract sets of pictures (see Figure 3). Two concrete and two abstract sets of pictures were used in the study, with each dyad required to solve all four tasks. The concrete sets contained familiar or known objects (kites in task 1 and cats in task 3), whereas the abstract sets depicted geometrical figures (‘radar’ in task 2) or patterns of figures (‘flowers’ in task 4). Each task consisted of nine pictures that differed only in terms of minor details. The features that differed between the pictures are marked and explained in Appendix A.

Figure 3: The speaker’s display in four different tasks.²

² The pictures have been taken from the children’s game “Differix” from RAVENSBURGER. (Mit freundlicher Genehmigung des Ravensburger Spielverlags.)
The difficulties that arose were different for the two groups of tasks: The concrete sets were expected to hold more challenges with regard to the individual cognitive demands and the abstract sets were assumed to be more difficult in terms of the demands due to interaction. As the concrete sets contain familiar objects, the detection of the relevant features was expected to be more error-prone because the details can be overlooked more easily (e.g. Biederman, 2000; see section 2.1.1). On the other hand, it was less difficult to describe the concrete sets of pictures because it was not necessary to find new expressions for the objects, the relevant features, and their position. To describe the pictures in task 3, the speaker could use terms like ‘cat’, ‘head’, or ‘tail’, and did not need to specify, for example, that the head is the upper part of the cat. To describe the unfamiliar objects or patterns in the abstract tasks, on the other hand, the speaker needed to invent terms and introduce them to the addressee; in other words a common language first had to be developed in order to establish referential identity (Isaacs & Clark, 1987; Clark & Brennan, 1991; see section 2.1.2).

<table>
<thead>
<tr>
<th>Demands due to interaction</th>
<th>Concrete pictures</th>
<th>Abstract pictures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual cogn. d.</td>
<td>Details of known objects may be overlooked more easily (e.g. Biederman, 2000).</td>
<td>Search for details among new, unknown objects requires attention and takes more time (Wang, Cavanagh &amp; Green, 1994).</td>
</tr>
<tr>
<td>Description is easier, because names for objects and object parts do already exist and are commonly known.</td>
<td>Description is more difficult, because names for objects and object parts have to be invented and grounded (Isaacs &amp; Clark, 1987).</td>
<td></td>
</tr>
<tr>
<td>Different frames of reference (viewer-centered vs. object-centered; e.g. Levelt, 1989) are possible. This may cause problems of referential identity.</td>
<td>Frames of reference are not likely to differ.</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Characteristics of concrete and abstract pictures.
Chapter 3: Study 1 – Test for the influence of individual and collaborative demands

The dichotomy presented so far – more salient individual cognitive demands in concrete tasks and more salient demands due to interaction in abstract tasks – needs to be supplemented by one aspect: In the concrete tasks, different spatial frames of reference were possible. This is particularly important for pictures that represent persons or figures because it is possible to take on their perspective. For example, it was possible to describe the cat’s line of sight from the speaker’s (viewer-centered) perspective or from the cat’s (object-centered) perspective (e.g. Schober, 1993; Levelt, 1989; Taylor & Tversky, 1996). This complicated the establishment of referential identity and may cause problems of understanding. By using four different sets of pictures – two in each group – the influence of the postulated demands on problem-solving and communication process as well as on the outcome can be analyzed. Table 1 gives an overview on the characteristics of concrete and abstract pictures.

In a collaborative setting where two persons are communicating freely it is difficult to separate their individual achievements. This means in the context of the collaborative task used here that the number of relevant features detected cannot only be ascribed to the speaker. If both are communicating freely, the addressee is also able to detect features and to take part in the labeling of features. The interaction and feedback effects that occur complicate a distinct separation of the performance components associated with the different demands. This interaction can be impeded if only the speaker is allowed to speak. In the current tasks, the role of the addressee had therefore been varied as a between-subject factor: either he was able to talk, as in classical collaborative situations, or he was not allowed to speak. In the latter case he did, however, have the possibility to send an audio warning signal to the speaker if he had not understood what the speaker had said. Thereby the speaker was able to repeat or repair his description and the situation retained its collaborative character.

The net-based setting used here, with two persons sitting each in different rooms and communicating through an audio-link, differs considerably from face-to-face settings (see section 2.1.3). Communication is quite challenging, because the collaborators cannot see each other and do not have physical copresence (Clark & Brennan, 1991). Thus, they neither can use deictic gestures to support their verbal utterances nor do they see at what the other person
is looking. As it is not possible to guide the other’s attention by pointing at specific objects or parts of them, the establishment of referential identity (Clark & Brennan, 1991) is hampered (see Stahl, 2003). A facilitation of the communication may be achieved through providing shared applications to the collaborators. The availability of shared applications leads to a kind of copresence, as both collaborators have a common view of the applications. Shared applications containing pictorial information often have a pointing feature or the possibility to draw. With these deictic gestures are possible, thus the establishment of referential identity may improve (Clark & Brennan, 1991). Shared applications containing textual information do have the additional benefit of allowing review; utterances can be looked at again later.

In the first study, the availability of a shared application was varied as a second between-subject factor: either the dyads had a pictorial or textual shared application or they had no shared application available. The shared applications could be seen by both speaker and addressee, but only the speaker was able to draw or write. This constraint was necessary because of the first between-subject factor (role of the addressee). If the addressee was not able to talk but allowed to write or draw on shared applications, there would again have been an interaction between the speaker’s and addressee’s performances.

The pictorial shared application consisted of a picture that was presented on a Microsoft Netmeeting-Whiteboard. The picture corresponded to each task respectively, e.g. in task 1 a kite was presented on the whiteboard. However, this picture had another feature combination in order to differ from the nine pictures that had to be described and arranged. The speaker could draw marks and point to parts of the picture. On the addressee’s display the marks and pointing could be seen on the whiteboard without any time delay.

The textual shared application was realized through a simple text editor with which the speaker could enter information. Through Netmeeting, the entered text could also be seen on the addressee’s display, again without any time delay.
3.1.1 Design

A 2x3 factor design was implemented (see Table 2). As between-subject factors, the role of the addressee (able to talk/ not able to talk) and the availability of a shared application (pictorial/ textual/ no shared application) were varied. Additionally, the type of pictures was varied as a within-subject factor (concrete/ abstract), so that each dyad was required to solve two concrete and two abstract tasks (see Figure 3). Each dyad was assigned randomly to one of the six conditions.

<table>
<thead>
<tr>
<th>Role of the addressee</th>
<th>Able to talk</th>
<th>Not able to talk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pictorial</td>
<td>concrete/ abstract</td>
<td>concrete/ abstract</td>
</tr>
<tr>
<td>Textual</td>
<td>concrete/ abstract</td>
<td>concrete/ abstract</td>
</tr>
<tr>
<td>No shared application</td>
<td>concrete/ abstract</td>
<td>concrete/ abstract</td>
</tr>
</tbody>
</table>

Table 2: Design of Study 1.

3.1.2 Participants

120 students (60 dyads) from the University of Freiburg participated in the study; 62 of them were male and 58 female. Students of Psychology were excluded. The participants had an average age of 23.35 years (stddev = 3.7; range = 20 to 37 years). All participants were German native-speakers and had normal color-vision. Each person received eight euros for his participation and the study took 60 to 90 minutes. The participants were randomly grouped into dyads and assigned the role of speaker or addressee. Participants did not know each other prior to the study.
3.1.3 Procedure

Each participant received instructions in form of a manual explaining the use of the technical environment and the procedure. The dyad then performed a training task to familiarize themselves with the technical environment. They were instructed to complete the tasks as accurately and quickly as possible. During the experimental phase, each dyad was required to solve four tasks. All dyads received the tasks in the same sequence. The speaker described the pictures and their positions and the addressee placed the pictures on his display accordingly. Depending on the condition, the addressee was either able to talk or he was not. In the latter case he did, however, have the possibility to send an audio warning signal to the speaker if he did not understand what the speaker had said. In the conditions with a shared application, the speaker was able to use this shared application to support his explanations. However, the speakers were not obliged to use the shared application, and were merely informed about the additional possibility of using it. After completing each task, the participants received feedback about how many pictures they had placed in the correct position and what features had been relevant in the task.

3.1.4 Measures

Three sets of data were collected to examine the outcome and the process of collaboration:

- audio recordings of the verbal communication,
- the drawings or writings in the shared applications, and
- performance measures (quality of the joint solution).

To examine whether the postulated demands affect the communication process, a coding scheme (see Table 3) consisting of categories for the individual cognitive demands and the demands due to interaction was developed. At the center of interest are three types of errors: one related to the individual cognitive demands and two related to the demands due to interaction. The categories relating to the individual cognitive demands are based on the number of relevant features that were identified by the dyad. The first type of errors committed by the speaker in the communication process is the failure to mention a relevant
feature (*feature not mentioned*). The categories relating to the demands due to interaction are aimed to answer the following questions: How much collaborative effort does the speaker invest? How often does he need to make corrections? Was he able to establish referential identity? The two errors related to the demands due to interaction both reflect failures to establish referential identity. The error *name of feature* was counted each time a speaker gave ambiguous description regarding the feature’s name and the error *frame of reference* each time a speaker gave ambiguous description regarding the frame of reference for the spatial description of the pictures. It should be noticed that the three errors are not independent of each other. Of course, if a speaker did not describe a relevant feature he could not commit one of the other two errors in the description.

<table>
<thead>
<tr>
<th>Individual cognitive demands</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of not identified features</strong></td>
<td>How many relevant features were NOT identified?</td>
</tr>
<tr>
<td><strong>Number of “Feature not mentioned” errors</strong></td>
<td>How often did the speaker give descriptions in which relevant features were not mentioned?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demands due to interaction</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of “Name of feature” errors</strong></td>
<td>How often did the speaker give descriptions that were ambiguous regarding the feature’s name?</td>
</tr>
<tr>
<td><strong>Number of “Frame of reference” errors</strong></td>
<td>How often did the speaker give descriptions that were ambiguous regarding the frame of reference?</td>
</tr>
<tr>
<td><strong>Number of picture descriptions</strong></td>
<td>How often did the speaker describe one picture?</td>
</tr>
<tr>
<td><strong>Number of warning signals or clarification questions asked</strong></td>
<td>How often did the addressee send a warning signal or ask for clarification?</td>
</tr>
<tr>
<td><strong>Number of coordination activities</strong></td>
<td>How often did the speaker use ‘meta-language’, e.g. to talk about the way in which he would describe the features?</td>
</tr>
</tbody>
</table>

Table 3: Coding scheme for the analysis of the communication process.
The audio recordings of the verbal communication were analyzed using these categories. Ten percent of the verbal data was coded by a second rater to enable the calculation of inter-rater reliability. The consistency of the coding was medium to high, indicating that the coding scheme could easily be used. Table 4 shows the ICC_{just, fixed} (intra-class correlation; see Wirtz & Caspar, 2002) for each of the process variables and for each task separately.

<table>
<thead>
<tr>
<th>Number of not identified features</th>
<th>Task 1 ‘kites’</th>
<th>Task 2 ‘radar’</th>
<th>Task 3 ‘cats’</th>
<th>Task 4 ‘flowers’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of “Feature not mentioned” errors</td>
<td>.99</td>
<td>.89</td>
<td>.63</td>
<td>.76</td>
</tr>
<tr>
<td>Number of “Name of feature” errors</td>
<td>1.0</td>
<td>.83</td>
<td>.91</td>
<td>.98</td>
</tr>
<tr>
<td>Number of “Frame of reference” errors</td>
<td>1.0</td>
<td>1.0</td>
<td>.96</td>
<td>1.0</td>
</tr>
<tr>
<td>Number of picture descriptions</td>
<td>1.0</td>
<td>.74</td>
<td>1.0</td>
<td>.94</td>
</tr>
<tr>
<td>Number of warning signals/ questions</td>
<td>.83</td>
<td>.97</td>
<td>.76</td>
<td>.93</td>
</tr>
<tr>
<td>Number of coordination activities</td>
<td>.73</td>
<td>.83</td>
<td>.52</td>
<td>.83</td>
</tr>
</tbody>
</table>

Table 4: ICC_{just, fixed} of the process variables for each of the four tasks in study 1.

In terms of the data gathered from the shared applications, the number of dyads that made use of them was counted. This data was then either coded as supporting the individual cognitive demands, or as supporting the demands due to interaction. In the first case, the speaker had marked or written down all features at the beginning of a task to make clear, what features differed between the pictures (e.g. the mouse’s tail, the position of the cat…). Whereas in the second case, the speaker had used the shared application to mark or write down the feature values of each picture (e.g. ‘In picture 1, mouse’s tail in s-form, cat on the left…’) during the description of the nine pictures. Appendix A.2 shows examples for both types of usage in shared pictorial and textual applications.

The quality of the joint solution was measured by the number of pictures placed in the correct position and the time needed to complete the task.
Chapter 3: Study 1 – Test for the influence of individual and collaborative demands

3.1.5 Hypotheses

The dependent variables were collected to answer the research questions by investigating the following hypotheses:

Research question 1 (differences in process and result between concrete and abstract sets of pictures).

As the two sets of pictures should differ according to their demands, fewer features are expected to be identified, more ‘feature not mentioned’ errors and more ‘frame of reference’ errors will be committed in concrete tasks. In abstract tasks, more ‘name of feature’ errors will be committed, more picture descriptions made, more warning signals sent or questions asked, and more coordination activities (= ‘meta language’ to agree on the communication and problem-solving process) will be carried out.

As the feature search is expected to be more error-prone in concrete sets of pictures, fewer pictures will be placed in the correct position for concrete pictures. However, the solution will be reached in less time compared to tasks with abstract pictures, because the demands due to interaction are higher in abstract sets of pictures.

Research question 2 (Use of the shared applications).

In concrete tasks, the shared applications are expected to be used more to support individual cognitive demands. In abstract tasks, they are expected to be used more to support the demands due to interaction.

Research question 3 (Good/ bad performing dyads).

A clear relation between the measures of process and outcome (performance measures) is expected: Dyads showing high performance are expected to have fewer difficulties during the collaboration process.
3.2 Results of Study 1

The following sections report the results of the analyses of process data, the performance measures and the data from the shared applications.

A multivariate analysis of variance (MANOVA) with repeated measures (for the factor ‘type of pictures’) was conducted to test the influence of the three factors (type of pictures, role of the addressee, availability of shared application) on the seven process variables (see Table 3) and the two performance measures (number of pictures placed in correct position, time needed to complete the task). There was an effect of the type of pictures ($F[9, 46] = 21.3, p < .01, \eta^2 = .81$), an effect of the role of the addressee ($F[9, 46] = 12.3, p < .01, \eta^2 = .71$) but no effect of the availability of a shared application. Additionally, there was a significant interaction between type of pictures and role of the addressee ($F[9, 46] = 2.9, p < .01, \eta^2 = .37$). The results of calculation of ANOVAs for all variables will be reported separately for the process data (section 3.2.1) and the performance measures (section 3.2.2).

As for the drawing and writing in the shared applications, only descriptive statistics have been calculated. The results will be reported in section 3.2.3. A summary of these results can be found in section 3.2.4 (Table 9). Further analyses regarding the relation between process and performance data are reported in section 3.2.5.

3.2.1 Process Data

Table 5 contains means and standard deviations for the seven process variables with each column corresponding to one factor. Overall, the standard deviations are quite high for most of the variables.
### Chapter 3: Study 1 – Test for the influence of individual and collaborative demands

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Type of pictures</th>
<th>Role of addressee</th>
<th>Availability of shared application</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of not identified features</strong></td>
<td>concrete: .9 (.9)</td>
<td>able to talk: .6 (.9)</td>
<td>pictorial: 7 (1.0)</td>
</tr>
<tr>
<td></td>
<td>abstract: .8 (1.1)</td>
<td>not able to talk: 1.1 (1.1)</td>
<td>textural: 1.1 (1.1)</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td></td>
<td>no s.a.: .7 (.9) n.s.</td>
</tr>
<tr>
<td><strong>Number of “Feature not mentioned” errors</strong></td>
<td>concrete: 9.0 (7.3)</td>
<td>able to talk: 5.3 (4.4)</td>
<td>pictorial: 6.3 (5.7)</td>
</tr>
<tr>
<td></td>
<td>abstract: 3.4 (3.1)</td>
<td>not able to talk: 7.1 (5.7)</td>
<td>textural: 7.5 (5.6)</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td>no s.a.: 4.8 (3.8) n.s.</td>
</tr>
<tr>
<td><strong>Number of “Name of feature” errors</strong></td>
<td>concrete: .6 (1.1)</td>
<td>able to talk: 1.2 (1.5)</td>
<td>pictorial: 1.3 (1.5)</td>
</tr>
<tr>
<td></td>
<td>abstract: 2.4 (2.8)</td>
<td>not able to talk: 1.8 (2.3)</td>
<td>textural: 1.1 (1.4)</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td>no s.a.: 2.1 (2.6) n.s.</td>
</tr>
<tr>
<td><strong>Number of “Frame of reference” errors</strong></td>
<td>concrete: 1.2 (3.1)</td>
<td>able to talk: .6 (1.0)</td>
<td>pictorial: .4 (1.1)</td>
</tr>
<tr>
<td></td>
<td>abstract: .2 (1.2)</td>
<td>not able to talk: .8 (2.8)</td>
<td>textural: .7 (2.0)</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td></td>
<td>no s.a.: .9 (2.6) n.s.</td>
</tr>
<tr>
<td><strong>Number of picture descriptions</strong></td>
<td>concrete: 22.6 (6.1)</td>
<td>able to talk: 20.4 (4.6)</td>
<td>pictorial: 21.0 (4.3)</td>
</tr>
<tr>
<td></td>
<td>abstract: 19.7 (3.6)</td>
<td>not able to talk: 21.8</td>
<td>textural: 21.3 (5.8)</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td>no s.a.: 20.9 (4.1) n.s.</td>
</tr>
<tr>
<td><strong>Number of warning signals or clarification questions asked</strong></td>
<td>concrete: 5.2 (4.5)</td>
<td>able to talk: 9.1 (4.7)</td>
<td>pictorial: 7.2 (5.2)</td>
</tr>
<tr>
<td></td>
<td>abstract: 7.4 (5.5)</td>
<td>not able to talk: 3.5 (3.1)</td>
<td>textural: 4.8 (3.9)</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td>no s.a.: 6.9 (5.4) n.s.</td>
</tr>
<tr>
<td><strong>Number of coordination activities</strong></td>
<td>concrete: .6 (.9)</td>
<td>able to talk: 1.4 (1.3)</td>
<td>pictorial: .8 (0.9)</td>
</tr>
<tr>
<td></td>
<td>abstract: 1.1 (1.1)</td>
<td>not able to talk: 0.6 (0.6)</td>
<td>textural: .9 (1.1)</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td></td>
<td>no s.a.: 1.0 (1.1) n.s.</td>
</tr>
</tbody>
</table>

* **p < .01; *p < .05; n.s. no significant effect

Table 5: Means and standard deviations (in parentheses) of the process data.

A calculation of ANOVAs revealed significant effects for all factors but among different variables. Figures 4 to 6 show the variables with significant differences for each factor and interaction.

![Figure 4: Process data measures with significant differences between concrete and abstract sets of pictures.](image)

---

72
For the type of pictures (see Figure 4) there were significant differences in the number of feature not mentioned' errors \( (F[1, 54] = 46.3, p < .01, \eta^2 = .46) \), the number of ‘name of feature’ errors \( (F[1, 54] = 22.3, p < .01, \eta^2 = .29) \), the number of ‘frame of reference’ errors \( (F[1, 54] = 5.2, p < .05, \eta^2 = .09) \), the number of picture descriptions \( (F[1, 54] = 16.9, p < .01, \eta^2 = .24) \), the number of warning signals or questions asked \( (F[1, 54] = 13.6, p < .01, \eta^2 = .20) \), and the number of coordination activities \( (F[1, 54] = 13.8, p < .01, \eta^2 = .20) \). As expected, in tasks with concrete pictures there were more ‘feature not mentioned’ errors committed, fewer ‘name of feature’ errors, more ‘frame of reference’ errors, fewer warning signals sent or questions asked and fewer coordination activities carried out. However, more picture descriptions were made in concrete tasks. It can be estimated that this was related to the higher number of ‘feature not mentioned’ errors as well as the higher number of ‘frame of reference’ errors: Both the detection of overlooked features as well as the recognition of misunderstandings regarding the frame of reference may cause a higher number of picture descriptions.

For the role of the addressee (Figure 5), there were significant differences for the number of not identified features \( (F[1, 54] = 5.7, p < .05, \eta^2 = .10) \), the number of ‘name of feature’ errors \( (F[1, 54] = 3.1, p < .05, \eta^2 = .06) \), the number of warning signals or questions asked \( (F[1, 54] = 46.5, p < .01, \eta^2 = .46) \), and the number of coordination activities \( (F[1, 54] = 9.9, p < .01, \eta^2 = .16) \). In dyads with the addressee being able to talk there were more features identified, fewer ‘name of feature’ errors committed, more questions asked and more coordination activities carried out.

![Figure 5: Process data measures with significant differences between dyads with an addressee that was not able to talk and dyads with an addressee that was able to talk.](image)
Chapter 3: Study 1 – Test for the influence of individual and collaborative demands

There were significant interactions between the type of pictures and the role of the addressee (Figure 6) for the number of ‘feature not mentioned’ errors \(F[1, 54] = 8.1, p < .01, \eta^2 = .13\), the number of picture descriptions \(F[1, 54] = 5.6, p < .05, \eta^2 = .09\) and the number of warning signals sent or questions asked \(F[1, 54] = 8.8, p < .01, \eta^2 = .14\). In concrete tasks the number of ‘feature not mentioned’ errors was higher overall than in abstract tasks and especially high in dyads where the addressee was not able to talk. For the number of picture descriptions there was a difference in the factor ‘role of the addressee’ in concrete tasks, with more picture descriptions in dyads where the addressee was not able to talk. The number of warning signals sent did not differ significantly between concrete and abstract tasks for dyads with the addressee being not able to talk. However, addressees who were able to talk asked more questions in abstract tasks.

![Figure 6: Significant interactions in process data measures between type of pictures and role of the addressee.](image)

### 3.2.2 Performance Measures.

Table 6 contains means and standard deviations for the performance measures (number of pictures placed in correct position and time needed to complete the task). The mean of correctly placed pictures has to be related to a maximum number of 18 correctly placed pictures: For concrete and abstract types of pictures, the maximum number of correctly placed pictures is 18 respectively (2 tasks x 9 pictures). To keep the mean values comparable, the means of the other two factors were divided by 2.
Chapter 3: Study 1 – Test for the influence of individual and collaborative demands

Table 6: Means and standard deviations (in parentheses) of the performance measures.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Type of pictures</th>
<th>Role of addressee</th>
<th>Availability of shared application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pictures placed in correct position (max. 18)</td>
<td>concrete: 11.2 (4.8) abstract: 15.5 (2.6)**</td>
<td>able to talk: 14.3 (3.1) not able to talk: 12.4 (4.0)**</td>
<td>pictorial: 13.9 (4.0) textual: 12.4 (2.9) no s.a.: 13.9 (3.8) n.s.</td>
</tr>
<tr>
<td>Time needed to complete the task (in sec.)</td>
<td>concrete: 600.3 (200) abstract: 972.2 (295.2)**</td>
<td>able to talk: 737.7 (250.9) not able to talk: 834.9 (236.9)*</td>
<td>pictorial: 763.5 (220.6) textual: 795.2 (250.0) no s.a.: 800.3 (190.4) n.s.</td>
</tr>
</tbody>
</table>

** p < .01; * p < .05; n.s. no significant effect

As shown in Figure 7, fewer pictures were placed in the correct position in tasks with concrete pictures ($F[1, 54] = 63.5, p < .01, \eta^2 = .54$), but, as expected, it took less time to complete them ($F[1, 54] = 105.8, p < 01, \eta^2 = .66$).

![Figure 7: Mean number of correctly placed pictures (left) and mean time needed to complete the task (right) by type of pictures.](image)

The number of correctly placed pictures differed significantly between the conditions in which the addressee was able to talk and those in which he was not ($F[1, 54] = 3.1, p < .05, \eta^2 = .11$), as did the time needed to complete the task ($F[1, 54] = 6.8, p < .01, \eta^2 = .06$). As expected, dyads in which the addressee was able to talk placed approximately one picture more in the correct position. In addition, they took roughly 50 seconds less to complete the task (see Figure 8).
Figure 8: Mean number of correctly placed pictures (left) and mean time needed to complete the task (right) by role of the addressee.

There was no significant effect for the availability of a shared application and no significant interactions.

3.2.3 Use of the Shared Application

As mentioned above, the data of the shared applications were either coded as supporting the individual cognitive demands (‘what are the features differing between the pictures’), or as supporting the demands due to interaction (‘what are the features values of picture 1?’). The use of the shared application – if available – was not obligatory. As the number of dyads having used the shared application in one task is quite small, the following closer look at the rate of use and type of use consists only in descriptive statistics.

**Pictorial Shared Application:** First, the number of dyads that used the pictorial shared application available to them was counted. Overall, 80% of the dyads used it for at least one task, but the number of dyads using the pictorial shared application differed notably from task to task (see Table 7). Overall, it was used less in concrete than in abstract tasks.

<table>
<thead>
<tr>
<th>Role of the Addressee</th>
<th>Concrete tasks</th>
<th>Abstract tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'Kites'</td>
<td>'Cats'</td>
</tr>
<tr>
<td>Not able to talk</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Able to talk</td>
<td>10%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Table 7: Rate of use (per cent) of the pictorial shared application by role of the addressee.
As Figure 9 shows, the pictorial shared applications were mostly used to support the demands due to interaction. This was especially the case in dyads where the addressee could talk.

### Textual Shared Application

Overall, 50% of the dyads that had a textual shared application available used it in at least one task. Again, the number of dyads using it differed from task to task (see Table 8). However, the rate of usage was higher in concrete tasks.

<table>
<thead>
<tr>
<th>Role of the Addressee</th>
<th>Concrete tasks</th>
<th>Abstract tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>‘Kites’</td>
<td>‘Cats’</td>
</tr>
<tr>
<td>Not able to talk</td>
<td>70%</td>
<td>50%</td>
</tr>
<tr>
<td>Able to talk</td>
<td>70%</td>
<td>30%</td>
</tr>
</tbody>
</table>

Table 8: Rate of use (per cent) of the textual shared application by role of the addressee.

Figure 10 shows how the textual shared applications were used. In contrast to the hypotheses, they were again in both sets of pictures mostly used to support the demands due to interaction. This was even more the case in dyads where the addressee could talk.
Figure 10: Type of use of the textual shared application (per cent) in concrete and abstract tasks by role of addressee.
### 3.2.4 Summary of the Results (Study 1)

<table>
<thead>
<tr>
<th>Type of Pictures</th>
<th>Process Data (MANOVA)</th>
<th>Performance Measures (MANOVA)</th>
<th>Data of Shared Application (Descriptive Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>concrete tasks:</td>
<td>- more ‘feature not mentioned’ errors</td>
<td>- fewer pictures placed in the correct position</td>
<td>- pictorial s.a. was less often used</td>
</tr>
<tr>
<td></td>
<td>- fewer ‘name of features’ errors</td>
<td>- less time needed to complete them</td>
<td>- textual s.a. was more often used</td>
</tr>
<tr>
<td></td>
<td>- more ‘frame of reference’ errors</td>
<td></td>
<td>no difference between concrete and abstract tasks in terms of type of usage (overall s.a. mostly used to support demands due to interaction)</td>
</tr>
<tr>
<td></td>
<td>- more picture descriptions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- fewer warning signals/questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- fewer coordination activities</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Role of the Addressee</th>
<th>Process Data (MANOVA)</th>
<th>Performance Measures (MANOVA)</th>
<th>Data of Shared Application (Descriptive Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dyads where the addressee was able to talk:</td>
<td>- more features identified</td>
<td>- more pictures placed in the correct position</td>
<td>- used even more to support the demands due to interaction</td>
</tr>
<tr>
<td></td>
<td>- fewer ‘name of features’ errors</td>
<td>- less time needed to complete the task</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- more questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- more coordination activities</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Availability of Shared Application</th>
<th>Process Data (MANOVA)</th>
<th>Performance Measures (MANOVA)</th>
<th>Data of Shared Application (Descriptive Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- no significant effects -</td>
<td>- no significant effects -</td>
<td></td>
<td>- used it more often than dyads with textual s.a.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interactions</th>
<th>Process Data (MANOVA)</th>
<th>Performance Measures (MANOVA)</th>
<th>Data of Shared Application (Descriptive Statistics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>if the addressee was not able to talk:</td>
<td>- more ‘feature not mentioned’ errors in concrete tasks</td>
<td>- no significant effects -</td>
<td>- pictorial s.a. more often used in abstract tasks</td>
</tr>
<tr>
<td></td>
<td>- more picture descriptions in concrete tasks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>if the addressee was able to talk:</td>
<td>- more questions in abstract tasks</td>
<td></td>
<td>- textual s.a. used nearly equally to support both demands</td>
</tr>
</tbody>
</table>

Table 9: Summary of the results (process data, performance measures, data of shared application).
3.2.5 Relation between Process and Performance Data

To develop instructional support measures, further information about the relation between process and performance data is required. The most important part of the ‘work’ is done by the speaker because it is his role to describe the pictures. As mentioned above, interaction and feedback effects, found in situations in which both participants are able to talk freely, influence the communication process. Therefore only the dyads in which the addressee was not able to talk were included in the following analysis that examines the relation between the way in which the speakers dealt with the demands (process data) and the dyads’ performance.

To obtain one single indicator of the dyads’ performance, a quotient was calculated indicating the number of correctly placed pictures per second (number of correctly placed pictures / time needed to complete the task).

Based on this value, the dyads were then divided on the 33rd and 66th percentile according to their performance. Three groups resulted, which can be labeled as ‘very well’, ‘well’, and ‘badly’ performing dyads.

A multivariate analysis of variance (MANOVA) was conducted to test the influence of the dyads’ performance on all seven process measures listed in the coding scheme (Table 3). In addition to the performance group factor, the type of pictures was included as second factor (within-subject) to check for possible interactions between a dyad’s/the dyads’ performance and the different demands arising in concrete and abstract tasks.

Table 10 contains means and standard deviations for the seven process variables with each column corresponding to one factor.
### Table 10: Means and standard deviations (in parentheses) of the process data for the performance groups and the type of pictures.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Performance group</th>
<th>Type of pictures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of not identified features</td>
<td>bad: 1.5 (1.2) medium: 1.0 (1.1) good: .8 (1.0) n.s.</td>
<td>concrete: 1.1 (1.9) abstract: 1.0 (1.2) n.s.</td>
</tr>
<tr>
<td>Number of “Feature not mentioned” errors</td>
<td>bad: 10.5 (6.5) medium: 6.9 (10.3) good: 4.0 (2.4) **</td>
<td>concrete: 11.1 (8.0) abstract: 3.1 (3.3) **</td>
</tr>
<tr>
<td>Number of “Name of feature” errors</td>
<td>bad: 2.9 (3.0) medium: 1.4 (1.7) good: 1.2 (1.5) *</td>
<td>concrete: .6 (1.1) abstract: 3.0 (3.4) **</td>
</tr>
<tr>
<td>Number of “Frame of reference” errors</td>
<td>bad: .5 (1.4) medium: 1.9 (3.2) good: 0 (0) n.s.</td>
<td>concrete: 1.3 (3.4) abstract: .3 (1.6) n.s.</td>
</tr>
<tr>
<td>Number of picture descriptions</td>
<td>bad: 23.4 (2.6) medium: 22.2 (5.0) good: 19.9 (2.6) n.s.</td>
<td>concrete: 24.1 (6.4) abstract: 19.5 (3.2) **</td>
</tr>
<tr>
<td>Number of warning signals</td>
<td>bad: 5.1 (4.7) medium: 2.4 (1.6) good: 3.0 (2.2) *</td>
<td>concrete: 3.3 (3.1) abstract: 3.7 (3.4) n.s.</td>
</tr>
<tr>
<td>Number of coordination activities</td>
<td>bad: .6 (.7) medium: .7 (.6) good: .4 (.6) n.s.</td>
<td>concrete: .4 (.5) abstract: .7 (.7) **</td>
</tr>
</tbody>
</table>

** p < .01; * p < .05; n.s. no significant effect

There was a significant effect for the performance group \((F[14, 44] = 2.2, p < .05, \eta^2 = .42)\) as well as for the type of pictures \((F[7, 21] = 15.7, p < .01, \eta^2 = .84)\). Additionally, there was a significant interaction between the performance group and the type of pictures \((F[14, 44] = 2.0, p < .05, \eta^2 = .39)\).

A calculation of ANOVAs revealed significant effects for both factors but in different variables. Figure 11 shows the variables with significant differences for the factor performance group. There were significant differences in the number of ‘feature not mentioned’ errors \((F[2, 27] = 5.7, p < .01, \eta^2 = .30)\), the number of ‘name of feature’ errors \((F[2, 27] = 3.0, p < .05, \eta^2 = .18)\), and the number of warning signals \((F[2, 27] = 13.9, p < .05, \eta^2 = .18)\). Dyads with bad performance had more errors of both types and the addressees sent more warning signals.
Chapter 3: Study 1 – Test for the influence of individual and collaborative demands

As the factor type of pictures is only included in this analysis to check for possible interactions with the performance groups, the ANOVA results for this factor are presented only briefly. For the type of pictures there were significant effects for the number of ‘feature not mentioned’ errors ($F[1, 27] = 53.3, p < .01, \eta^2 = .66$), the number of ‘name of feature’ errors ($F[1, 27] = 13.4, p < .01, \eta^2 = .33$), the number of picture descriptions ($F[1, 27] = 22.0, p < .01, \eta^2 = .45$), and the number of coordination activities ($F[1, 27] = 10.5, p < .01, \eta^2 = .28$).

Figure 12 shows the significant interactions between the performance group and the type of pictures.

There were significant interactions for the number of ‘feature not mentioned’ errors ($F[2, 27] = 6.8, p < .01, \eta^2 = .34$) and the number of ‘frame of reference’ errors ($F[2, 27] = 3.8, p$
<.05, $\eta^2 = .22$). The difference between bad, medium and good performing dyads in the number of ‘feature not mentioned’ errors was much more significant/notable in concrete than for abstract tasks. Regarding the ‘frame of reference’ errors the analysis revealed an interaction that is difficult to interpret: In concrete tasks only medium performing dyads committed ‘frame of reference’ errors, whereas in abstract tasks only bad performing dyads committed such errors. This may be due to the small total number of these errors and seems to be a quite accidental effect.

To conclude: Dyads with bad performance (represented by the number of pictures placed in the correct position per second) indeed experienced difficulties in the process of the experiment: Both kinds of demands (individual cognitive and those due to interaction) were sources of errors and the addressee had to send many warning signals. Especially in concrete tasks, speakers of bad performing dyads often did not mention relevant features.

### 3.3 Discussion of Study 1

The first study aimed at gaining a better understanding of what skills are needed to solve an experimental remote collaborative task successfully. The results should provide a basis to design instructional support measures for enhancing remote collaboration. The study examined the impact of individual cognitive demands as well as demands due to interaction on process and performance measures of collaboration. For this purpose, different sets of pictures were used that differed according to their demands.

The analysis of the process data showed clear differences depending on the type of pictures: In line with the hypotheses, more ‘feature not mentioned’ and ‘frame of reference’ errors were committed with concrete pictures. With abstract tasks, more errors ‘name of feature’ were committed, more warning signals were sent and more coordination activities took place. This indicates that concrete tasks present more ‘difficulties’ regarding the individual cognitive demands – the search for features that differ between the pictures is more error-prone for concrete tasks with familiar looking objects. Moreover, problems due to missing referential identity concerning the spatial frame of reference arise more often with concrete tasks. On the
other hand, abstract tasks seem to pose greater demands due to interaction (more ‘name of feature’ errors, more warning signals or questions, more coordination activities). But contrary to the hypotheses, the speaker described the pictures more often in concrete tasks. This is likely to be related with the higher number of ‘feature not mentioned’ and ‘frame of reference’ errors: Both the detection of overlooked features as well as the recognition of misunderstandings regarding the frame of reference may cause a higher number of picture descriptions. If speaker and addressee notice that they are missing feature differences or that they do not have a common point of view on the pictures, they need more picture descriptions to solve the task because they have to repeatedly clarify the description and positioning of the pictures. They will do this until they identify all relevant features and assure a mutual frame of reference.

Differences due to variant demands were also noted in the analysis of the performance measures: more pictures were placed in the correct position in tasks with abstract pictures compared to those with concrete pictures, but more time was required to complete these tasks. It was predicted that concrete tasks involve greater individual cognitive demands while abstract tasks are more demanding regarding interaction competencies. This result strongly points towards an illusion of simplicity (Nickerson, 1999), which occurs for the concrete sets of pictures: The familiarity of the concrete pictures leads to the deceptive impression that the differences were easy to find and the pictures easy to describe.

There was also a difference between concrete and abstract pictures regarding the rate of usage of the shared applications: The pictorial shared application had a higher rate of usage for abstract tasks whereas the textual shared application was used more often for concrete tasks. The dyads seem to have used the possibility of indicative gestures for establishing referential identity especially for the abstract tasks (see Tang, 1991). As it is easier to write down features with already existing and known names, the dyads used the textual shared application more often with concrete tasks. Regarding the type of usage, there was no difference between concrete and abstract tasks. In both tasks, they were used more to support the demands due to interaction, which means that the speaker used the shared application to show or describe
the feature values of each single picture instead of more generally showing or writing what feature differences exist in the task.

Not surprisingly, dyads where the addressee was able to talk had better results (more pictures placed in the correct position in less time) than dyads where the addressee was unable to talk. As the analysis of the process data showed, this was due to the fact that, working together, two people identified more features (more features were identified in dyads where the addressee was able to talk). Furthermore, they had the possibility of speaking about problems (more questions in dyads where the addressee was able to talk) and the planned procedure (more coordination activities). The shared applications were – if available – also used more in dyads where the addressee was able to talk.

A second goal of the study was to test a first (technical) support measure. There was either a pictorial, textual or no shared application available that the speakers could use to support their verbal descriptions. The calculation of MANOVA both for the process data and for the performance measures revealed no significant effect for the factor availability of a shared application. This is surely due to the fact that not all dyads of the respective conditions used them, as their usage was not obligatory. Depending on the condition and the task, the rate of use varied significantly. Overall, 80% of the dyads used the pictorial shared application at least once, whereas only 50% of the dyads used the textual shared application at least one time. For some tasks the shared applications were used only by 10, 20 or 30% of the dyads. It seems that both shared applications, but especially the textual one, were not perceived as being very helpful for solving the tasks.

In order to support the development of instructional support measures an analysis was conducted to gain more information and to scrutinize the relation between the dyads’ performance and the way in which the speaker dealt with the task demands in the collaboration process. For this purpose, we compared good, medium, and bad performing dyads with regard to the process measures. To focus on the speaker’s performance and to exclude possible interaction and feedback effects, only dyads with the addressee being unable to talk were included in the analysis.
The results showed that dyads with bad performance indeed had problems in the process: Speakers in dyads with bad performance committed more ‘feature not mentioned’ as well as ‘frame of reference’ errors and the addressees sent more warning signals. The number of ‘feature not mentioned’ errors for bad performing dyads was much higher in concrete than in abstract sets of pictures.

Beside the three main research questions that were addressed, study 1 should also provide information on how to improve the task material. The two sets of pictures that ought to differ according to their demands showed clear differences, as expected: Abstract pictures proved to be more difficult in terms of demands due to interaction, while concrete pictures presented more ‘challenges’ with regard to individual cognitive demands because the search for feature differences is more error-prone. Additionally, for concrete pictures it should be noted that all concrete tasks contain objects that allow for different spatial descriptions. Furthermore, all tasks should be comparably complex with regard to the individual cognitive demands, that is, the feature differences being more or less equally difficult to detect for each task, i.e. the pictures of one task should contain approximately as many elements as the pictures of the other tasks.

To conclude: What did we learn from study 1 that will help us develop instructional support measures? First, the results show that a dyad’s success clearly depends on the collaboration process: Difficulties in dealing with the two demands are related to poor performance. Therefore a sustainable support should aim at reducing the number of errors related to both demands. Furthermore, there should be different hints for concrete and abstract pictures. More specifically, for concrete pictures, it is important to carefully search for all feature differences and to establish a mutual frame of reference. For abstract pictures, the need to define understandable names for the features should be pointed out.
Chapter 4

**Study 2: Instructional Support for Individual and Collaborative Demands**

The results of the first study were used to design instructional support measures to foster collaboration in a net-based problem-solving task. These instructional measures included a model collaboration presented to the dyads prior to collaboration in the form of an on-screen video with audio instructions. Furthermore, they also received a collaboration script during collaboration to remind participants of what they had just learned. The instructional measures contained two levels of support, each corresponding to one of the two demands (individual cognitive and those due to interaction). The developed measures were evaluated in a second study. Again, a collaborative picture-sorting task was used.

As collaboration takes place not only under ‘optimal’ conditions (where the collaborators can interact freely and immediately), the effect of the support measures were additionally examined in non-interactive conditions. The second study addressed the following research questions:

- Do the instructional support measures have an impact on the collaboration process and outcome? Are there incremental effects of the two levels of support (level 1 – feature search vs. level 1 + 2 – features search and labeling of features)? Does the supported level have a specific impact on process and outcome of one specific type of task (concrete vs. abstract sets of pictures)?

- Does non-interactive collaboration also benefit from the instructional support measures? Does the impact of the demands (individual cognitive demands and demands due to interaction) differ in interactive and non-interactive conditions? Are some demands more important in interactive or non-interactive conditions?
4.1 Method of Study 2

Set of Pictures. Again, four different tasks were used (see Figure 13), two with concrete and two with abstract pictures, and each dyad was required to solve all four.

Also similarly, nine pictures had to be described and positioned. However, this time the addressee had 16 pictures per task on his display to choose from. The speaker still had 9 target pictures but saw 7 additional pictures on his display. Because there were more pictures that had to be described, the difficulty of choosing the right picture from the remaining ones thus still existed for the 8th and 9th picture (see Schober & Clark, 1989). In many studies it has not been considered that the participants can use processes of elimination to position the last objects or pictures if all presented objects or pictures have to be used (e.g. Krauss & Weinheimer, 1964, 1966; Clark & Wilkes-Gibbs, 1986; Isaacs & Clark, 1987; Wilkes-Gibbs & Clark, 1992; Murfitt & McAllister, 2001).

The pictures themselves were also modified. They differed from the pictures of study 1 in the following ways: First, the pictures were not colored but consisted only of grayscales. We did this to eliminate the possible influences of colors on the perception and attention processes. Furthermore, the different sets of pictures were of almost equal complexity with regard to the number of elements. All feature differences were dichotomous, i.e. had only two possible values.

As was confirmed in the first study, concrete sets of pictures hold more individual cognitive demands due to greater difficulties in finding the distinguishing features among the pictures. Additionally, the concrete sets of pictures hold challenges regarding the establishment of referential identity for the frame of reference because concrete pictures can be described from different point of views. Abstract sets of pictures are more demanding regarding the labeling of the features, that is establishing referential identity for the feature’s name.
Figure 13: The nine target pictures on the speaker's display for the four tasks used in study 2.  

**Amount of Support.** Following the approach of Rummel and Spada (2005a, 2005b; see section 1.4.3), a collaboration model was developed (Hansen & Spada, 2006). It was presented

---

3 The pictures “cat” and “kite” have been developed based on pictures of the children’s game “Differix” from RAVENSBURGER.
to the dyads as on-screen video with audio instructions prior to their collaboration. Furthermore, a collaboration script reminding the collaborators of what they had just learned was provided during collaboration. Taking into account the results of the first study, both demands – the individual cognitive demands and those due to interaction – should be supported. Thus, the model instruction as well as the script contained two levels of support, each corresponding to one of the two aforementioned demands. Both levels contained hints for dealing with the demands in the best possible way and, furthermore, were related to one subtask.

- **Level 1 – support of the individual cognitive demands**: In the first study, many errors occurred particularly in concrete tasks because the collaborators failed to identify all relevant features. Therefore, the hints of level 1 stressed the importance of a careful search for the relevant features, especially of the concrete pictures, which appear at a first glance to be less difficult to manage. Additionally, a subtask was introduced: The collaborators were asked to search for the features that differed between the pictures and to mark them in the picture editor, which contained one of the pictures from each task.

- **Level 2 – support of the demands due to interaction**: The second pitfall of the task lies in the establishment of referential identity. The hints of level 2 emphasized that referential identity for the feature’s name and position, as well as for the spatial frame of reference, are crucial for successful communication and should be established as early on as possible. The subtask of level 2 consisted of labeling the feature differences. The speaker was asked to enter the features’ names into an individual text editor. Of course, searching for feature’s names is normally part of the grounding process (e.g. Clark & Brennan, 1991) and thereby a joint action. However, the aim was to design one version of instructional support for both interactive and non-interactive conditions (see the next section ‘Mode of communication’ for more details). Therefore, searching for feature’s names was assigned to the speaker as an individual task.
As the task contains individual cognitive demands besides those due to interaction, it can be divided into an individual and a joint phase. In the individual phase, the collaborators first have to search for the feature differences on their own and additionally the speaker has to think of names for the features before engaging in the description and positioning of the pictures in the collaborative phase. By making this division into individual and collaborative phases explicit, the instructional support measures may help to overcome the collaborators’ tendency to solely engage in joint activities and forget about individual work phases (Hermann, Rummel & Spada, 2001).

The two phases and the three subtasks introduced in the support measures follow a structure considered as ideal because of the separation of individual and collaborative phases and the differentiation of individual subtasks. Figure 14 illustrates the structure of the task along the following subtasks: (1) an initial individual phase in order to search for the relevant features and (2) to name them, as well as (3) a collaborative phase to be used for grounding and for the description and positioning of the nine pictures. The first and second subtask are both part of the individual phase. The third phase can further be divided into the time for grounding of the features and the time needed for the description and positioning of each of the nine pictures.

Figure 14: Ideal structure of the collaboration process for interactive and non-interactive modes of communication.

Figure 15 provides an overview of the support measures. As mentioned above, the instructions prior to collaboration (= the collaboration model) were realized as on-screen videos (one for each role). Furthermore, the script support was realized during the collaboration.
Chapter 4: Study 2 – Instructional support for individual and collaborative demands

<table>
<thead>
<tr>
<th>Support elements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prior to collaboration</strong></td>
</tr>
<tr>
<td><strong>On-screen video</strong></td>
</tr>
<tr>
<td>• Hints for feature search (level 1) and naming of pictures (level 2)</td>
</tr>
<tr>
<td>• Structuring of the task, division into subtasks, definition of roles</td>
</tr>
<tr>
<td>• Technical directions</td>
</tr>
</tbody>
</table>

| **During collaboration** |
| **Cooperation script** |
| (1) Search for features (picture editor) - level 1 |
| (2) Naming of features (text editor) – level 2 |
| only speaker |
| (3) Grounding of features and names |
| (4) Description and positioning of pictures |

Figure 15: Constitution of the support measures with elements prior to and during collaboration.

The on-screen videos contained visual and auditory information. This is in line with Mayer’s (2001) modality effect, which claims that students learn more deeply from information presented in a mixed mode (e.g. animation and spoken text) compared to information presented in a single mode (e.g. animation and written text). The on-screen videos showed the screen of a speaker or an addressee solving a concrete task (‘clown’) similar to the ones participants would have to solve later. To guide the participant’s visual attention, the location of buttons or other important elements on the screen mentioned in the auditory instructions were highlighted (yellow squares). Figure 16 exemplarily shows a screenshot of a speaker’s video in one condition (support level 1 and 2/ interactive; see section 4.3).

Each video was accompanied by auditory information. Depending on the experimental condition, the audio-channel was used for instructional explanations containing hints for feature search and picture description, instructions concerning the structuring of the task into phases and subtasks, as well as information regarding the roles and technical directions. Important parts of the conversation, namely the utterances of speaker and addressee, were displayed in text-boxes to provide a model collaboration. Thus, the functions of auditory information and text-boxes were reversed compared to the model used by Rummel and Spada (2005a, 2005b). Their model included presenting audio recordings of two collaborators solving a task to the learners and, additionally, displaying instructional explanations in text-boxes. Due to the experimental variation of the amount of support, including conditions that should
receive no support at all (‘no support’) and conditions receiving only support of the individual cognitive demands (‘level 1’), it was not possible to follow this procedure. Providing the participants with an audio recording of a model dyad solving a task would have ruled out this variation of the amount of support, because it was impossible to create an audio model for example purposes only that included the feature search (individual cognitive demands) and not the labeling of the features (demands due to interaction).

Figure 16: Screenshot of a speaker’s on-screen video (condition ‘support level 1 + 2/interactive mode of communication’).

The collaboration script (see Figure 17) had the function of reminding the collaborators of what they just had learned in the on-screen video (i.e. what has to be done at what point in time) and contained further technical advice.

---

4 The pictures have been taken from the children’s game “Differix” from RAVENSBURGER.
### Task 1 – Search for features and labeling of features

1.) Open picture editor for task 1 and chose red color.
2.) Click on “Start search task 1”
3.) Search for features and mark them in picture editor
4.) Open text editor
5.) Name features and write them down in text editor
6.) Click on “go on”

**Please do NOT speak with your partner during this subtask „Search for features and labeling of features“!**

---

### Task 1 – Description of pictures

1.) Click on “Start task 1” and let your partner know that you have started
2.) Explain names for features and mention the point of view of your description
3.) Describe pictures one by one (start with position 1, then position 2,...)
4.) Click on “done”

**Please proceed as quickly and as accurately as possible!**

---

Figure 17: Speaker’s collaboration script for the individual (above) and collaborative phase (below) (translated version; condition ‘support level 1 + 2/ interactive mode of communication’).

To test if the two levels of support have incremental effects, the amount of support was varied as one between-subject factor. Either there was no support at all, or only the individual cognitive demands were supported (only level 1), or the dyad received the complete support (levels 1 and 2).

**Mode of Communication.** Research has shown that speakers need feedback from their addressees to design optimal and effective descriptions (e.g. Krauss & Weinheimer, 1966; Schober, 1993; Clark & Krych, 2004; see section 1.2.2). However, many communication settings allow feedback only with time delay. This is the case for asynchronous settings such as written communication via letters or electronic mail. Furthermore, some settings do not enable interaction at all. Such settings emerge for example in classical lecture style teaching or if following written instructions on how to assemble furniture from pieces of wood and screws. ‘Collaboration’ in non-interactive settings requires much more effort as the grounding procedure is impaired: The addressee cannot signalize his understanding or misunderstanding of utterances. To test if the speakers could be compensated for the constraints of non-interactive settings, the mode of communication was varied as a second between-subject factor, being either interactive or non-interactive. In interactive conditions, speaker and
addressees communicated through an audio link simultaneously. In non-interactive conditions, on the other hand, the speakers recorded their descriptions, and addressees later arranged the pictures from the recordings. In both modes the same instructional support measures (no support vs. level 1 vs. level 1 + 2) were implemented.

4.1.1 Design

A 2x3-factor design was used (see Table 11) with one additional within-subject factor.

<table>
<thead>
<tr>
<th>Mode of communication</th>
<th>Interactive</th>
<th>Non-interactive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1 + 2</td>
<td>concrete/ abstract</td>
<td>concrete/ abstract</td>
</tr>
<tr>
<td>(complete support)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>concrete/ abstract</td>
<td>concrete/ abstract</td>
</tr>
<tr>
<td>(only indiv. demands)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No support</td>
<td>concrete/ abstract</td>
<td>concrete/ abstract</td>
</tr>
</tbody>
</table>

Table 11: Design of study 2.

As one between-subject factor, the amount of support was varied (complete support - level 1 + 2 vs. only support of individual demands - level 1 vs. no support). As a second between-subject factor, the mode of communication was varied, being either interactive or non-interactive. As a within-subject factor, the set of pictures was varied so that each dyad had to solve all four tasks. Two tasks contained concrete pictures and two contained abstract pictures (see Figure 13). To control for sequence effects, the tasks had to be solved in one out of four possible sequences (see Table 12). Each task sequence was realized twice per condition. The participants were assigned randomly to one of the six conditions.
Sequence n° | Task 1 | Task 2 | Task 3 | Task 4
--- | --- | --- | --- | ---
1 | Kite | Crochet | Cat | Circles
2 | Cat | Circles | Kite | Crochet
3 | Crochet | Cat | Circles | Kite
4 | Circles | Kite | Crochet | Cat

Table 12: The four realized sequences of tasks in study 2.

4.1.2 Participants

96 students (48 dyads) of the University of Freiburg participated in the study. Students of Psychology were excluded. 36 of the participants were male and 60 were female. The participants had an average age of 24.15 years (stddev = 4.4, range = 18 to 48). All participants were German native-speakers. The study took 90 to 120 minutes and each participant received 15 Euros for his participation. The participants were randomly grouped into dyads and assigned the role of speaker and addressee. Participants did not know each other prior to the study.

4.1.3 Procedure

Each participant received instructions in the form of an on-screen video. To keep the procedure and time-on-task comparable, the dyads without support were also shown an on-screen video prior collaboration that just contained some technical instructions. The on-screen videos of both support conditions introduced the two subtasks, defined the roles, and contained information regarding the structuring of the task as well as hints for feature search. For the complete support condition there were also hints for the labeling of the pictures (see Figure 17). After watching the on-screen video, the dyads performed a training task (including description and positioning of three pictures out of sixteen) in order to familiarize themselves with the technical environment and the subtasks. They were instructed to complete the tasks as accurately and quickly as possible. During the experimental phase, each dyad was required
to solve four tasks. The tasks had to be solved in one out of four possible sequences (see Table 12).

Each task involved the use of two different screens (related to one individual and one collaborative phase respectively): Speakers and addressees in all conditions received the same first screen (see Figure 18). It contained all 16 pictures in a random order but different to the target order.

Figure 18: Overview of the procedure in the different support conditions.
In dyads with complete support (level 1 + 2), speaker and addressee used this first screen to individually search for the feature differences and mark them in individual picture editors; the speaker furthermore wrote down the names for the relevant features in an individual text editor. In dyads with support of the individual demands only (level 1), speaker and addressee also searched for the feature differences and marked them in individual picture editors. Dyads without support did also see this first screen but without instruction as to what to do with it.

The second screen of the speaker displayed the target order. The second screen of the addressee displayed the sixteen pictures in a random order. This time, the addressee could move and rearrange the pictures using his mouse (drag and drop). In all conditions, the speaker described the pictures and their positions and the addressee accordingly placed the pictures on his display.

In non-interactive conditions, speaker and addressee performed the tasks not at the same time but in succession. The speaker had a recording device on his display and could start and stop recording his explanations to the addressee as he liked. A microphone was positioned on his table next to the monitor. The description of each speaker was randomly assigned to one addressee who later arranged the pictures following the recorded descriptions. The addressee had an audio-player device on his display and could start, stop and rewind the recording of the speaker’s descriptions as often as he wanted to. That said, speakers and addressees were told, just as in interactive conditions, to proceed as accurately and quickly as possible.

### 4.1.4 Measures

Two sets of data were collected to examine the collaboration process as well as the outcome: audio recordings of the verbal communication and performance measures. The coding scheme for the verbal communication data emphasizes different kinds of problems and errors committed during collaboration (Table 13).

The categories refer to the two demands on the collaborators: the first two categories to the individual cognitive demands, and categories 3 to 7 to the demands due to interaction. The three errors coded in the verbal data analysis of study 1 as well as the ‘number of not
identified features’ category are included in the coding scheme of this second study. Furthermore, the ‘number of features not mentioned before description’, ‘number of irrelevant features’ as well as ‘number of complicated descriptions’ categories are included in the analysis of the second study’s process data. These three new categories replace three categories used in the first study (number of picture descriptions, number of warning signals or clarification questions asked, number of coordination activities) that could not be used to analyze the process of non-interactive conditions.

### Individual cognitive demands

- **Number of not identified features**
  - How many relevant features were NOT identified?

- **Number of “Feature not mentioned” errors**
  - How often did the speaker give descriptions in which relevant features were not mentioned?

### Demands due to interaction

- **Number of features not mentioned before description**
  - How many relevant features were NOT mentioned by the speaker before starting the description of the first picture?

- **Number of “Name of feature” errors**
  - How often did the speaker give descriptions that were ambiguous regarding the feature’s name?

- **Number of “Frame of reference” errors**
  - How often did the speaker give descriptions that were ambiguous regarding the frame of reference?

- **Number of irrelevant features**
  - How often did the speaker give description of irrelevant features?

- **Number of complicated descriptions**
  - How often did the speaker give complicated and pedestrian descriptions of a feature?

Table 13: Coding scheme for the problems and errors occurring during communication

The audio recordings of the communication were analyzed using these categories. To check for inter-rater reliability, ten percent of the verbal data was coded by a second rater. The consistency of the coding was medium to high, indicating that the coding scheme could easily
be used. Table 14 shows the $\text{ICC}_{\text{just, fixed}}$ (intra-class correlation; see Wirtz & Caspar, 2002) for each of the process variables and for each task separately.

<table>
<thead>
<tr>
<th>Number of not identified features</th>
<th>‘cat’</th>
<th>‘kite’</th>
<th>‘crochet’</th>
<th>‘circle’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.0</td>
<td>.97</td>
<td>.76</td>
<td>.97</td>
</tr>
<tr>
<td>Number of “Feature not mentioned” errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.83</td>
<td>.97</td>
<td>.76</td>
<td>1.0</td>
</tr>
<tr>
<td>Number of features not mentioned before description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>1.0</td>
<td>.99</td>
<td>.97</td>
</tr>
<tr>
<td>Number of “Name of feature” errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.86</td>
<td>1.0</td>
<td>.94</td>
<td>.97</td>
</tr>
<tr>
<td>Number of “Frame of reference” errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.97</td>
<td>.86</td>
<td>.86</td>
<td>.76</td>
</tr>
<tr>
<td>Number of irrelevant features</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.86</td>
<td>.94</td>
<td>.76</td>
<td>.97</td>
</tr>
<tr>
<td>Number of complicated descriptions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.76</td>
<td>.66</td>
<td>.73</td>
<td>.83</td>
</tr>
</tbody>
</table>

Table 14: $\text{ICC}_{\text{just, fixed}}$ of the problem and error categories for each of the four tasks in study 2.

Additionally, the amount of time used before starting the description of the pictures (‘individual’ and ‘grounding time’) was measured. Two time periods were registered: The first corresponds to the individual phase of the task where speaker and addressee searched for features or labeled the features individually. Additionally, a grounding period was measured. It represents the period during which the grounding of the features took place or during which the speaker introduced the features and their labels to the addressee in case of the non-interactive conditions. Unfortunately, the possibility to compare the impact of the support measures in interactive and non-interactive conditions implicates problems of comparability. It is important to note that these measures are not completely comparable for the two modes of communication, as the ‘grounding phase’ in case of the non-interactive conditions consisted only in the presentation phase (Clark & Schaefer, 1989).

Furthermore, one nominal scaled variable coded from the audio recordings (‘need for security’), assessed if the dyad invested additional time into controlling the work done or if they did not (check all/ only some of the pictures again).

Moreover, in the non-interactive conditions it was counted how often an addressee rewound the audio recording to listen to parts of the description once more (‘amount of rewinding’).
As in the first study, the performance measures also included the number of pictures placed in the correct position at the end of one task and the time needed for description and positioning of the pictures.

### 4.1.5 Hypotheses

The dependent variables were collected to answer the research questions by investigating the following hypotheses:

**Research question 1 (Impact of the instructional support measures).**

Dyads without support are expected to show more problems and produce more errors (all problems and errors listed in Table 13). As dyads with level 1 support are only provided with support of the individual cognitive demands, more problems and errors referring to the communicative demands (number of features not mentioned before description, number of ‘name of feature’ errors, number of ‘frame of reference’ errors, number of irrelevant features, and number of complicated descriptions) are predicted compared to the dyads with complete (level 1 + 2) support.

As the complete support contains instructions to search for the feature differences as well as the features’ names individually and the level 1 support contains only instructions to search for the former, more ‘individual time’ will be used in complete support (level 1 + 2) conditions, followed by conditions with support of the individual demands only (level 1). More ‘grounding time’ will be used by dyads with complete support (level 1 + 2), because the complete support contains instructions for grounding.

Dyads in conditions with complete support (followed by dyads with support of individual demands only) will show better performance with more pictures placed in the correct position and less time used to complete the tasks.

Considering abstract tasks present greater demands due to interaction and concrete tasks more individual cognitive demands, fewer problems and errors referring to the individual cognitive demands but more problems and errors referring to the communicative demands (see Table
are expected to occur in abstract tasks. Furthermore, more ‘grounding time’ will be needed by the dyad in abstract tasks.

Dyads with complete support (level 1 + 2) are expected to perform better (more pictures placed in the correct position and less time needed to complete the task) in abstract tasks than dyads without support or with level 1 support.

**Research question 2 (Interactive/ non-interactive conditions).**

In non-interactive conditions, more problems and errors (see Table 13) as well as a higher amount of ‘individual time’ and less ‘grounding time’ are expected, because the speaker receives no feedback of his addressee, speaker and addressee cannot search for feature differences together, and grounding is not possible.

The number of correctly placed pictures will be higher in interactive conditions, as will be the amount of time needed to complete the description and positioning of the pictures.

In concrete tasks, the number of problems and errors related to the individual cognitive demands are expected to be particularly high for non-interactive conditions. Comparatively, the number of problems and errors related to the communicative demands (see Table 13) in abstract tasks are expected to be particularly high for non-interactive conditions because the addressee cannot ‘help’ search for features. In non-interactive conditions, the time needed to describe and position the pictures is expected to be higher for abstract tasks as the speaker will use fewer shortened expressions.

**4.2 Results of Study 2**

The following sections report the results of the process analyses and the performance measures. Table 15 gives an overview on the dependent variables included in the MANOVA and those excluded from the MANOVA for various reasons (nominal scaled or only available for a subgroup).
Dependent variables included in MANOVA

- Problems and errors during communication process (7)
- ‘Individual’ and ‘grounding time’ (2)
- Performance measures (2)

Further dependent variables

- ‘Need for security’ (nominal scaled)
- Amount of rewinding (only for non-interactive conditions)

Table 15: Overview of all dependent variables in study 2.

As an ANOVA revealed no effect of the task order, the sequence will not be taken into account in the following analyses.

A multivariate analysis of variances (MANOVA) with repeated measures (for the factor ‘type of pictures’) was conducted to test the influence of the three factors on process variables and performance measures (see Table 12). There was an effect of the type of pictures ($F[11, 32] = 16.4, p < .01, \eta^2 = .85$), an effect of the amount of support ($F[22, 66] = 2.8, p < .01, \eta^2 = .48$) and an effect of the mode of communication ($F[11, 32] = 9.6, p < .01, \eta^2 = .77$).

Additionally, there was a significant interaction between the type of pictures and the mode of communication ($F[11, 32] = 2.7, p > .01, \eta^2 = .48$).

ANOVA were calculated for all variables. The results will be reported separately for the process data (section 4.2.1) and the performance measures (section 4.2.2). Section 4.2.3 summarizes the results of the MANOVA and section 4.2.4 contains the results of further analyses.

### 4.2.1 Process Data

Table 16 contains means and standard deviations for the process variables included in MANOVA, with each column corresponding to one factor. Overall, the standard deviations are quite high for most of the variables. The results of the calculations of ANOVAs for each variable will first be described for the problems encountered and the errors committed during
the communication process and subsequently for the time used before starting the
descriptions (‘individual’ and ‘grounding time’).

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Type of pictures</th>
<th>Amount of support</th>
<th>Mode of communication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of not identified features</strong></td>
<td>concrete: 2.1 (1.8) abstract: 0.7 (1.3) **</td>
<td>complete: 1.4 (1.3) indiv: 1.2 (1.3) no train.: 1.8 (1.9) n.s.</td>
<td>interactive: 0.5 (0.7) non-interactive: 2.4 (1.5) **</td>
</tr>
<tr>
<td><strong>Number of ‘feature not mentioned’ errors</strong></td>
<td>concrete: 27.1 (14.3) abstract: 9.2 (12.6) **</td>
<td>complete: 14.4 (11.5) indiv: 15.9 (10.3) no train.: 24.5 (14.6)</td>
<td>interactive: 12.6 (10.5) non-interactive: 23.7 (13.7) **</td>
</tr>
<tr>
<td><strong>Number of features not mentioned before description</strong></td>
<td>concrete: 6.9 (3.2) abstract: 4.4 (3.7) **</td>
<td>complete: 3 (2.7) indiv: 6.2 (1.3) no train.: 7.7 (2.8) n.s.</td>
<td>interactive: 4.6 (3.5) non-interactive: 6.7 (3.1) **</td>
</tr>
<tr>
<td><strong>Number of ‘name of feature’ errors</strong></td>
<td>concrete: 0 (0) abstract: 1.2 (1.8) n.s.</td>
<td>complete: 0.5 (0.8) indiv: 0.7 (1.2) no train.: 0.7 (0.8) n.s.</td>
<td>interactive: 0.4 (0.7) non-interactive: 0.8 (1.0) n.s.</td>
</tr>
<tr>
<td><strong>Number of ‘frame of reference’ errors</strong></td>
<td>concrete: 2.7 (4.8) abstract: 0 (0) n.s.</td>
<td>complete: 0.5 (0.8) indiv: 1.4 (3.2) no train.: 2.1 (2.4) n.s.</td>
<td>interactive: 1.4 (2.3) non-interactive: 1.3 (2.2) n.s.</td>
</tr>
<tr>
<td><strong>Number of irrelevant features</strong></td>
<td>concrete: 13.7 (13) abstract: 3.1 (7) **</td>
<td>complete: 6.9 (9.4) indiv: 8.3 (9.4) no train.: 9.9 (6.1) n.s.</td>
<td>interactive: 8 (9.9) non-interactive: 8.7 (10.2) n.s.</td>
</tr>
<tr>
<td><strong>Number of complicated descriptions</strong></td>
<td>concrete: 0.4 (1.4) abstract: 2 (3.2) **</td>
<td>complete: 0.6 (1.3) indiv: 1.2 (1.6) no train.: 1.8 (2.5) n.s.</td>
<td>interactive: 0.8 (1.8) non-interactive: 1.6 (2.8) n.s.</td>
</tr>
<tr>
<td>‘Individual time’ (in seconds)</td>
<td>concrete: 357.7 (333.5) abstract: 549.9 (428.8) **</td>
<td>complete: 705.3 (397.7) indiv: 410.9 (201.4) no train.: 154.6 (149.1) **</td>
<td>interactive: 346.6 (278.1) non-interactive: 560.7 (432.4) **</td>
</tr>
<tr>
<td>‘Grounding time’ (in seconds)</td>
<td>concrete: 72.5 (89.2) abstract: 103 (114.7) **</td>
<td>complete: 156.3 (101.6) indiv: 54.7 (72.8) no train.: 52.3 (96.4) n.s.</td>
<td>interactive: 118 (104.3) non-interactive: 57.5 (91.5) n.s.</td>
</tr>
</tbody>
</table>

** p <.01; * p < .05; n.s. no significant effect

Table 16: Means and standard deviations (in parentheses) of the process data included in MANOVA.

104
Chapter 4: Study 2 – Instructional support for individual and collaborative demands

Problems and errors during communication process. Figures 19 to 22 show the variables with significant differences for each factor and for the interaction.

Figure 19: Significant effects of the type of pictures on the problems and errors during communication process.

For the type of pictures (see Figure 19) there were significant differences in all seven problem and error categories: As predicted, the number of not identified features \((F[1, 42] = 49, p < .01, \eta^2 = .54)\), the number of ‘feature not mentioned’ errors \((F[1, 42] = 80.6, p < .01, \eta^2 = .66)\), the number of features not mentioned before description \((F[1, 42] = 80.5, p < .01, \eta^2 = .66)\), as well as the number of ‘frame of reference’ errors \((F[1, 42] = 14.7, p < .01, \eta^2 = .26)\), occurred more often in concrete tasks. In contrast to the predictions, the number of irrelevant features \((F[1, 42] = 28.4, p < .05, \eta^2 = .40)\) occurred also more often in concrete tasks. However, the number of ‘name of feature’ errors \((F[1, 42] = 22.2, p < .01, \eta^2 = .35)\) as well as the number of complicated descriptions \((F[1, 42] = 21.8, p < .01, \eta^2 = .34)\) arose more often in abstract tasks, as expected.

As can be seen in Table 16, the means of problems and errors during communication differed mostly in the expected way between the conditions with complete support, support of individual demands only and without support. In general, fewer problems and errors occurred in conditions with support. Nevertheless, there were significant differences only in the
number of ‘feature not mentioned’ errors ($F[2, 42] = 5.2, p < .05, \eta^2 = .20$) and in the number of features not mentioned before description ($F[2, 42] = 14.1, p < .01, \eta^2 = .40$).

As expected, both errors were mostly committed by dyads in conditions without support (see Figure 20).

![Figure 20: Significant effects of the amount of support on the problems and errors during communication process.](image)

Again, for the mode of communication the means of problems and errors during communication differed mostly in the expected way: Fewer problems and errors occurred in interactive conditions. However, there were only significant differences in the number of not identified features ($F[1, 42] = 52.1, p < .01, \eta^2 = .55$), in the number of ‘feature not mentioned’ errors ($F[1, 42] = 17.9, p < .01, \eta^2 = .30$), and in the number of features not mentioned before description ($F[1, 42] = 8.4, p < .01, \eta^2 = .17$). As predicted, all three errors were more often committed in dyads with non-interactive mode of communication (see Figure 21).

![Figure 21: Significant effects of the mode of communication on the problems and errors during communication process.](image)

Between the type of pictures and the mode of communication (see Figure 22) there were significant interactions for the number of not identified features ($F[1, 42] = 8.3, p < .01, \eta^2$
Chapter 4: Study 2 – Instructional support for individual and collaborative demands

= .17), the number of features not mentioned before description ($F[1, 42] = 8.1, p < .01, \eta^2 = .16$), and the number of complicated descriptions ($F[1, 42] = 4.5, p < .05, \eta^2 = .10$).

![Graphs showing significant interactions between type of pictures and mode of communication.]

Figure 22: Significant interactions between type of pictures and mode of communication.

As expected, the number of problems and errors related to the individual demands in concrete tasks and those related to the demands due to interaction in abstract tasks were higher for non-interactive conditions: In concrete tasks, fewer features were identified in non-interactive conditions. The difference in the mean number of not identified features between interactive and non-interactive conditions was smaller in abstract tasks. The number of features not mentioned before starting the description differed more between interactive and non-interactive conditions in abstract tasks. In non-interactive conditions, the difference between concrete and abstract tasks in the number of complicated descriptions was more important than in interactive conditions with more complicated descriptions in abstract tasks.
‘Individual’ and ‘Grounding time’. Figure 23 shows for each factor the amount of individual and grounding time.

![Figure 23: Significant effects of the type of pictures (left), the amount of support (center), and the mode of communication (right) on the individual and grounding time.](image)

There was more individual time ($F[1, 42] = 21.7, p < .01, \eta^2 = .34$) used in abstract tasks. As expected, more grounding time ($F[1, 42] = 7.7, p < .01, \eta^2 = .16$) was also used in abstract tasks.

As expected for the amount of support, the dyads with complete support took the highest amount of individual time followed by dyads with support of the individual cognitive demands and finally the dyads without support ($F[2, 42] = 41.1, p < .01, \eta^2 = .66$). Dyads with complete support took three times as much time for grounding than dyads in the other two conditions ($F[2, 42] = 8.7, p < .01, \eta^2 = .29$).

Also in line with the expectations, dyads in non-interactive conditions took more individual time ($F[1, 42] = 13.5, p < .01, \eta^2 = .24$) but less grounding time than dyads in interactive conditions ($F[1, 42] = 6.8, p < .05, \eta^2 = .14$).
4.2.2 Performance Measures

Table 17 shows means and standard deviations for the performance measures. The mean of correctly placed pictures has to be related to a maximum number of 18 correctly placed pictures.\(^5\) Again, each column corresponds to one factor.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Type of pictures</th>
<th>Amount of support</th>
<th>Mode of communication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pictures placed in correct position (max. 18)</td>
<td>concrete: 5.9 (7.5)</td>
<td>complete: 9.3 (7.2)</td>
<td>interactive: 14.4 (3.9)</td>
</tr>
<tr>
<td></td>
<td>abstract: 13 (7.2)**</td>
<td>indiv: 10.3 (6.5)</td>
<td>non-interactive: 4.6 (5.2)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>no train.: 8.9 (8.5)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Time needed for description and positioning of the pictures (in sec.)</td>
<td>concrete: 1063.5 (499.4)</td>
<td>complete: 896.2 (371.2)</td>
<td>interactive: 1135.5 (483.1)</td>
</tr>
<tr>
<td></td>
<td>abstract: 1057.6 (471.7)</td>
<td>indiv: 1103.2 (509.2)</td>
<td>non-interactive: 985.7 (438.4)</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>no train.: 1182.3 (538.4)</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

\(** p < .01; * p < .05; n.s. no significant effect\)

Table 17: Means and standard deviations (in parentheses) of the performance measures.

Figure 24 displays for each factor and for the interaction the variable with significant differences.

Figure 24: Significant effect of type of pictures on number of correctly placed pictures (left), of amount of support on time needed for description and positioning (center), and of communication mode on number of correctly placed pictures (right).

For abstract pictures there were more **pictures placed in the correct position** than for concrete pictures \(F[1, 42] = 39.1, p < .01, \eta^2 = .48\). This result was in line with the results of

---

\(^5\) For concrete and abstract types of pictures, the maximum number of correctly placed pictures is 18 respectively (2 tasks \(\times\) 9 pictures). To keep the mean values comparable, the means of the other two factors were divided by 2.
study 1 and the expectations for study 2. However, there was no significant difference in the time needed for description and positioning of the pictures, in contrast to expectations.

As expected, dyads without support took more time for description and positioning of the pictures than dyads with support of the individual demands as well as dyads with complete support. Dyads with the complete support were faster than dyads in the other conditions ($F[2, 42] = 2.5, p < .05, \eta^2 = .11$). In contrast to the hypotheses, there was no significant difference in the number of correctly placed pictures between the conditions with different amounts of support. As predicted, dyads with interactive mode of communication placed more pictures in the correct position than dyads with non-interactive mode of communication ($F[1, 42] = 72.6, p < .01, \eta^2 = .64$). However, there was no significant difference in the time needed for description and positioning of the pictures between interactive and non-interactive conditions.

This is due to the significant interaction between the type of pictures and the mode of communication (see Figure 25) that occurred for the time needed for description and positioning of the pictures ($F[1, 42] = 17.4, p < .01, \eta^2 = .29$). In the interactive conditions, dyads took more time for concrete tasks. If necessary, dyads in interactive conditions can search for the feature differences together or take additional time for establishing a mutual frame of reference. In non-interactive conditions the dyads took more time for abstract tasks; accordingly, speakers take more time to describe unfamiliar objects if they do not get any feedback from their addressee. The next section contains a table summarizing the results of the MANOVA.

Figure 25: Significant interaction between the type of pictures and the mode of communication.
### 4.2.3 Summary of the Results (Study 2)

<table>
<thead>
<tr>
<th>Type of Pictures</th>
<th>Process Data – Problems and errors during communication process</th>
<th>Process Data – ‘individual’ and ‘grounding time’</th>
<th>Performance Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>concrete tasks:</td>
<td>- fewer features identified</td>
<td>concrete tasks:</td>
<td>concrete tasks:</td>
</tr>
<tr>
<td></td>
<td>- more ‘feature not mentioned’ errors</td>
<td>- less individual and less grounding time</td>
<td>- fewer pictures placed in the correct position</td>
</tr>
<tr>
<td></td>
<td>- fewer features mentioned before the description</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- fewer ‘name of feature’ errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- more ‘frame of reference’ errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- more irrelevant features described</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- fewer complicated descriptions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amount of support</th>
<th>non-support:</th>
<th>complete support:</th>
<th>complete support:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- more ‘feature not mentioned’ errors</td>
<td>- more individual and more grounding time</td>
<td>- less time needed to complete the task</td>
</tr>
<tr>
<td></td>
<td>- fewer features mentioned before the description</td>
<td>support of individual demands:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- more individual time than</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>without support</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mode of Communication</th>
<th>non-interactive:</th>
<th>non-interactive:</th>
<th>non-interactive:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- fewer features identified</td>
<td>- more individual but less grounding time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- more ‘feature not mentioned’ errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- fewer features mentioned before the description</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Interactions

<table>
<thead>
<tr>
<th>Process Data – Errors during communication process</th>
<th>Process Data – ‘individual’ and ‘grounding time’</th>
<th>Performance Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>if the mode of communication was non-interactive:</td>
<td>if the mode of communication was non-interactive:</td>
<td>if the mode of communication was non-interactive:</td>
</tr>
<tr>
<td><em>particularly high number of not identified features in concrete tasks</em></td>
<td><em>more time needed to complete the task in abstract tasks</em></td>
<td><em>more time needed to complete the task in concrete tasks</em></td>
</tr>
<tr>
<td><em>difference in number of features not mentioned before description between concrete and abstract tasks smaller than in interactive conditions</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>particularly high number of complicated descriptions in abstract tasks</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 18: Summary of the results of study 2 (MANOVA including process data and performance measures).
4.2.4 Further Analyses

This section contains further analyses of the process variables that have not been included in the MANOVA (see Table 15): ‘need for security’ and ‘amount of rewinding’ in the non-interactive conditions.

To get supplementary information on the impact of the support measures, the two performance measures were related for each of the three support conditions. The resulting graphs are presented in this section.

Additionally, the time needed for the description and positioning of each single picture have been looked at exemplarily for one dyad with complete support and one dyad without support (for more details see Hansen & Spada, 2006).

**Need for security:** Table 19 shows the number of tasks being checked again by the dyads in each condition. In interactive conditions without support the ‘need for security’ seams to have been particularly high, with about twice as many tasks checked again compared to the other interactive conditions. In non-interactive conditions, a comparable high number of tasks checked again emerged for abstract sets of pictures, regardless of the amount of support. However, a calculation of a chi-square test did not reveal significant differences.

<table>
<thead>
<tr>
<th>mode of communication</th>
<th>interactive</th>
<th>non-interactive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>concrete</td>
<td>abstract</td>
</tr>
<tr>
<td>complete support</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>only support of individual demands</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>no support</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 19: Number of tasks checked again in each condition.

**Amount of rewinding.** Figure 26 shows for conditions with complete support, support of individual demands only and no support the mean number of times an addressee rewound the audio recording to listen to parts of the description again.
Dyads with complete support had a smaller mean number of rewinding than dyads in the two other support conditions. However, an ANOVA calculated to test the influence of the type of pictures and the amount of support on the number of times the addressee rewound the audio recording did not reveal any significant main effects or interaction.

![Figure 26: Number of times the addressee rewound the audio recording by amount of support.](image)

**Relation of Number of Correctly Placed Pictures and Time Needed.** To gain further insight on the impact of the instructional support measures, the two performance measures ‘number of correctly placed pictures’ and ‘time needed for the description and positioning’ were related. Figures 27 to 30 show for each task separately how many pictures were placed in the correct position after a particular period of time. It is important to note, that the measures were related as follows: For each of the three support conditions the mean time needed to place one, two, three etc. pictures on the correct position was calculated. These means include a different number of dyads’ measures, dependent on the number of dyads having placed 1, 2, 3 etc. pictures on the correct position.
E.g. in the task ‘cat’ (Figure 27 a) five pictures were placed in the correct position after about 100 seconds in the level 1 + 2 support condition (complete). This mean time includes the measures of 11 dyads (Figure 27 b), because the other 5 dyads placed fewer pictures in the correct position. In the level 1 condition (individual) 5 pictures were placed in the correct position after about 110 seconds, again with 11 dyads included. Dyads in the conditions without support took about 170 seconds to place the same number of pictures in the correct position, again with 11 dyads included.

Figure 27: (a) Number of pictures placed in the correct position in relation to the time needed for the ‘cat’ task for each of the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures in the correct position.
In the task ‘kites’ (Figure 28 a) five pictures were placed in the correct position after again about 100 seconds in the level 1 + 2 support condition (complete), after about 250 seconds in the level 1 support condition (individual), and after about 180 seconds in the condition without support. These means are calculated from the measures of 12 dyads for the level 1 + 2 and the level 1 support conditions and 9 dyads for the condition without support (Figure 28 b).

Figure 28: (a) Number of pictures placed in the correct position in relation to the time needed for the ‘kites’ task for the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures on the correct position.
In the task ‘crochet’ (Figure 29 a) five pictures were correctly placed after about 180 seconds in the level 1 + 2 support condition (complete), after about 230 seconds in the level 1 support condition (individual), and about 260 seconds in the condition without support. These means are calculated from the measures of 14 dyads for the level 1 + 2 support condition, 16 dyads for the level 1 support condition, and 11 dyads for the condition without support (Figure 29 b).

![Figure 29: (a) Number of pictures placed in the correct position in relation to the time needed for the ‘crochet’ task for the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures on the correct position.](image)
In the task ‘circles’ (Figure 30 a) dyads with level 1 + 2 support (complete) took about 210 seconds, dyads with level 1 support (individual) about 310 seconds, and dyads without support about 330 seconds to place five pictures in the correct position. These means are calculated from the measures of 13 dyads for the level 1 + 2 support as well as the level 1 support conditions and 10 dyads for the condition without support (Figure 30 b).

Figure 30: (a) Number of pictures placed in the correct position in relation to the time needed in the ‘circles’ task for the three ‘amount of support’ conditions; (b) Number of dyads having placed 1, 2, 3 etc. pictures on the correct position.
Time Needed for Each Single Picture. The time needed for positioning each single picture will be presented by two exemplary dyads. One of the two dyads had the complete support (‘with support’) while the other was not supported at all (‘without support’). Both dyads performed in the interactive mode of communication and solved the four tasks in the same order (sequence 1: kite, crochet, cat, circles; see Table 12). These two dyads were chosen because they were prototypical for the two conditions.

Figure 31 depicts the time needed for the description and positioning of each single picture for each of the two dyads.

The dyad with support took a little more time for the first picture, but altogether the time did not differ a great deal for the different pictures. The dyads in the support conditions were
instructed to use individual time to search for the feature differences and to think of the features’ names before starting the description and positioning of the pictures.

For the dyad without support, the large amount of time for the first picture is striking, and is followed by a smaller peak around the fourth and fifth pictures. By looking at the verbal data, it becomes clear that the large amount of time for the first picture is due to the search for the feature differences and the grounding process. Even if long, the search and grounding were not done properly (especially in tasks kite, crochet and circles) and thus problems occurred after the description and positioning of some pictures: The peak around the fourth or fifth picture in the tasks ‘circles’, ‘kite’, and ‘cat’ is due to repair mechanisms as a restart of the feature search, repeated grounding, or the provision of an extremely detailed description of the picture containing numerous irrelevant features.

Figure 32 shows for the ideal collaboration (‘with support’) and for the unsupported collaboration (‘without support’) a schematic example of the collaboration process, including the demands arising at each step, undertaken to solve the task. The diagram on the left hand side shows the ideal collaboration process: a linear sequence of steps taken, whereas the diagram on the right hand side shows a recursive process typical in the collaboration of dyads without support.
Figure 32: Schematic representation of the collaboration process with support (left) and without support (right).

S = Speaker
A = Addressee

Chapter 4: Study 2 – Instructional support for individual and collaborative demands
4.3 Discussion of Study 2

The aim of the second study was to evaluate the impact of the instructional support measures on the collaboration process and outcome. As in the first study, two sets of pictures were used that differed according to their demands. Again, the concrete sets of pictures turned out to be more difficult, and, counter intuitively, the abstract sets of pictures were easier to manage. This may again be caused by an illusion of simplicity with the concrete sets of pictures. The *instructional support measures* included specific help for both kinds of demands realized in two levels. To evaluate the impact of these two levels on the collaboration process and outcome, the amount of support was varied in study 2. As it is impossible to give hints for the labeling of the features without emphasizing the need to carefully search for features, the two levels could not be realized independently. To investigate the impact of the two levels, the conditions with level 1 support only were compared with the conditions that received the complete support (level 1 + 2). The analysis of the verbal data indeed showed differences in the problems and errors depending on the amount of support: In line with the hypotheses, problems and errors related to the individual cognitive demands were higher for the conditions without support and relatively similar in the two support conditions that both received hints for feature search and a subtask with an individual picture editor. The number of problems and errors related to the communicative demands was also highest for the conditions without support but it were relatively close to the number of errors in the level 1 support conditions. The differences across the support conditions were significant only for the number of ‘feature not mentioned’ errors and the number of features not mentioned before description, but for the other problems and errors the means showed tendencies in the predicted way.

As expected, the amount of ‘individual’ and ‘grounding time’ also differed significantly for the factor ‘amount of support’: Dyads with complete support took the highest amount of individual time followed by dyads with level 1 support only and finally the dyads without
support. At the same time, dyads with complete support took three times as long for grounding as dyads in the other two conditions.

So far, the instructional support measures seem to have improved the collaboration process in the intended way. Unfortunately, the measures’ impact did not reflect on the performance measures unanimously: The complete support conditions indeed took less time for the description and positioning of the pictures, but there were no significant differences between the numbers of correctly placed pictures for the three support conditions. Why did the improved collaboration process have no impact on a dyad’s performance? To get further information on the effect of the support measures, the numbers of correctly placed pictures in relation to the time needed for description and positioning of the pictures were inspected. In the support conditions, less time was needed to place a greater number of pictures in a correct position than in conditions without support. It can be supposed, that the support could have shown effects on the performance measures in the future; i.e. if more pictures had to be described and positioned.

The support measures did have an impact. In addition to the process measures included in the MANOVA, this was evident from two further process measures: Dyads in interactive conditions without support felt less confident about their performance than dyads in interactive conditions with complete or only level 1 support. When collaborating interactively, the simple fact of having received hints (regardless whether for only one or for two demands) seems to have been enough to reduce the dyad’s need to double-check the pictures. Even if the support did not have the same effect on the perceived confidence in non-interactive conditions, the complete support had an impact on the extent to which the speaker’s descriptions were understandable: For dyads in non-interactive conditions the number of times an addressee rewound the speaker’s description was significantly lower with complete support. Furthermore, the comparison of two dyads regarding the time they needed for the single pictures revealed that the process was smoother with complete support.

A further goal of the study was to illuminate the processes in non-interactive collaboration and to test if the process and performance also of non-interactive dyads could be enhanced by
the designed instructional measures. As reported by different authors (e.g. Clark & Krych, 2004), the non-interactive conditions had more difficulties to solve the tasks in an appropriate way. As predicted, the numbers of all seven problems and errors were higher in non-interactive conditions (significant differences only for the number of not identified features, the number of ‘feature not mentioned’ errors, and the number of features not mentioned before description). The illusion of simplicity seems to have been higher in non-interactive conditions, since the number of features that the speaker did not identify at all was particularly high for concrete tasks. In non-interactive conditions, the ‘need for security’ was also higher in abstract tasks, which again gives rise to the supposition of an illusion of simplicity. One main difficulty of abstract tasks is to find appropriate expressions to describe the features. This was especially pronounced in non-interactive conditions, where a high number of complicated descriptions occurred in abstract tasks. In both types of tasks the number of pictures not mentioned before starting the description was equally high in contrast to interactive conditions, where it was quite low for abstract tasks.

Additionally, a higher amount of individual time and less grounding time were used by non-interactive dyads. The speakers wanted to be sure, not to overlook features and therefore took more time to individually search for them. Of course, the ‘grounding process’ consisted only in the presentation phase (Clark & Schaefer, 1989; Clark & Brennan, 1991) and was shorter than a ‘real’ grounding process in interactive communication.

The advantage of interactive communication was also evident from the performance measures: The interactive conditions outperformed the non-interactive conditions with a higher number of correctly placed pictures. There was no significant difference for the time needed for description and positioning of the pictures, but, as expected, the mean time was slightly longer in interactive conditions. An interesting interaction effect occurred for this second performance measure: Non-interactive dyads took more time to complete abstract tasks, while interactive dyads took more time to complete concrete tasks. This result points in the same direction as the interaction effects for errors coded in the collaboration process. The illusion of simplicity seems to be more prevalent in non-interactive conditions and at the same
time the speaker has more difficulties to find short and understandable expressions for the features in abstract tasks. In line with assumptions by Clark and colleagues (e.g., Clark, 1996; see section 1.2.2) unidirectional communication or the production and reception of monologues is not as efficient as the interactive and bidirectional communication process. The higher amount of time needed for concrete tasks in interactive conditions is presumably due to problems of establishing a mutual frame of reference (see results of study 1).

To summarize: The second study was designed to investigate whether, through the developed support, the collaborators could learn to overcome the problems that occurred during the collaborative task and, in particular, communicate effectively. The dyads without support faced various problems: Speaker and addressee did not structure the process or balance individual and collaborative work phases, were confused about how to deal with the different demands at the same time, and therein committed many errors. By contrast, the dyads with support seemed to have learned how to collaborate: The collaboration process was better organized, clearly structured and contained fewer communication errors. The combination of an on-screen video containing a model collaboration and a collaboration script to help structure the collaboration process therefore seems, so far, to be a promising approach to foster collaboration. Furthermore, the results suggest the need for both support levels. Especially with concrete tasks, the support of both demands is crucial and should contain instructions to assure a mutual frame of reference.
Chapter 5

*Overall Discussion*

This section includes a summary and discussion of the results of both studies followed by perspectives for further research and practical implications.

5.1 Summary and Discussion of Results

5.1.1 Types of Demands

One aim of this dissertation was to contribute to a better understanding of the demands and processes underlying collaboration. Following ideas of Dillenbourg (1999), Teasley and Roschelle (1993) or Van Bruggen et al. (2003), two basic types of demands on the collaborators were postulated occurring in addition to the demands of net-based communication: Individual cognitive demands and demands due to interaction. Because of his simplified and restricted character, the collaborative task used in this dissertation could clearly be analyzed according to the two demands. Processes of visual search and object recognition were considered as individual cognitive demands, whereas the installment of referential identity was the main marker of demands due to interaction.

The first study aimed to test the influence of these demands on process and outcome of collaboration. Therefore, two types of tasks were used: Concrete sets of pictures containing known objects and abstract sets of pictures containing unknown objects. The two types of tasks were expected to differ according to their demands: The concrete sets of pictures should hold more challenges with regard to the individual cognitive demands because of the more error-prone visual search processes. In contrast, the abstract sets were assumed to be more difficult in term of the demands due to interaction because of the difficulty to describe
unfamiliar objects. Additionally, the concrete sets of pictures allowed for different frames of references.

As expected, the results of the first study indeed showed differences in the collaboration process depending on the types of pictures with more ‘feature not mentioned’ and ‘frame of reference’ errors committed in concrete tasks and more ‘name of feature’ errors, more warning signals and more coordination activities occurred in abstract tasks. This indicates that concrete tasks hold more ‘difficulties’ regarding the individual cognitive demands – the search for features that differ between the pictures is more error-prone for concrete tasks with familiar looking objects. Moreover, problems due to missing referential identity concerning the spatial frame of reference arise more with concrete tasks. On the other hand, abstract tasks indeed appear to hold more challenges regarding the demands due to interaction (more ‘name of feature’ errors, more warning signals or questions, more coordination activities).

The performance measures also differed between the tasks, with better results for abstract tasks but more time needed to complete them. This result strongly points towards an illusion of simplicity (Nickerson, 1999) with the concrete sets of pictures: The familiarity of the concrete pictures led to the deceptive impression that the differences were easy to find and the pictures easy to describe. The results of the second study certainly reinforced this supposition: The collaboration process contained less errors and the performance was better in abstract tasks. Again, the abstract sets of pictures were easier to deal with.

Further evidences for the postulated ‘demand model’ could be found in the type of use of the shared applications in study 1: In abstract tasks, the pictorial shared application and in concrete tasks the textual shared application were used more often. The dyads seem to have used the pictorial shared application’s option for indicative gestures - especially for abstract tasks where speaker and addressee had to establish a conceptual pact (Brennan & Clark, 1996) over how to refer to the features. In more complicated situations, such as the description of unfamiliar pictures, the dyads relied on the visual modality to be sure of reaching a mutual understanding. This is in line with Clark and Kryches’ (2004) assumption: ‘In dialogue, the
participants use vocal and gestural modalities in parallel. [...] The visual modality is faster and more secure than the auditory modality for certain types of communication” (p. 78).

The dyads probably used the textual shared applications more often with concrete tasks, because it is easier to write down features if they have short and efficient ‘vocabulary’ to describe them. To describe abstract pictures, the speaker would have had the choice between a figurative (e.g. ‘the big eye’) and a literal (e.g. ‘the small circle inside of the big circle’) description perspective (Fussell & Krauss, 1989; Murfitt & McAllister, 2001). Figurative descriptions are short and efficient, but they require more coordination between speaker and addressee to achieve a mutual understanding and are therefore not likely to be chosen in the context of this task. The literal descriptions are longer and higher costs thus arise if writing them down in a textual shared application. Consequently, speakers preferred to use the verbal modality.

Finally, the analysis of the collaboration process in relation to the dyads’ performance (only for dyads where the addressee was not able to talk) in the first study also showed differences between the performances in concrete and abstract tasks: Poorly performing dyads showed more errors overall. However, the number of ‘feature not mentioned’ errors was for poorly performing dyads much higher in concrete tasks.

On a theoretical level the results of study 1 showed different domains of referential identity that each implied its own specific problems. In the collaborative task used in this dissertation, these domains were, besides the feature’s name and the feature’s position, the frame of reference used in the spatial description. In tasks with concrete pictures containing objects with which the speaker can identify, the problem of assuring a mutual frame of reference added to the individual cognitive demands in an important way.

5.1.2 Dialog vs. Monolog

In both studies, conditions enabling ‘normal’ bidirectional communication and interaction, i.e., dialog, were compared to conditions enabling no feedback or only restricted feedback from the addressee, i.e., monolog conditions. The implementation of dialog versus monolog
conditions allows inspection of the individual achievements without interaction and feedback effects. Furthermore, the ‘collaboration’ or problem-solving processes can be compared to examine if speakers in monolog conditions can compensate for the lack of feedback.

In the first study the addressee was either able to talk or he was not. In the latter case he had the possibility to send an audio warning signal to the speaker if he didn’t understand what the speaker had said. Thereby the speaker was able to repeat or repair his description and tailor it more or less to the needs of the addressee; the problem-solving situation still had a collaborative character. Nevertheless, the speaker had to find the features on his own, without help of the addressee, and he had to design his descriptions without relying on clearly grounded expressions. Speakers getting no feedback from addressees are known to design longer and more elaborated descriptions (Krauss & Weinheimer, 1966). At the same time, the addressee had to find the described picture without the possibility to ask for clarification in case of ambiguities or misunderstandings. Schober and Clark (1989) showed that overhearers (addressees who cannot interact) make more errors in understanding than do addressees with feedback possibilities.

As expected, dyads where the addressee was able to talk placed more pictures in the correct position in less time as dyads where the addressee was not able to talk. The analysis of the process data showed that this was due to the facts that two persons can together identify more features and because they have the possibility to talk about problems and plan the procedure together.

In the second study interactive and non-interactive conditions were compared. In interactive conditions speaker and addressee communicated through an audio-link at the same point in time. In non-interactive conditions, however, speakers recorded their descriptions and addressees later arranged the pictures based on the recordings. Here, the non-interactive conditions held even more difficulties for the participants than the first studies’ monolog condition because of the total lack of feedback: E.g. the addressee could not signal his understanding or misunderstanding of utterances and could not warn for overlooked features.
Chapter 5: Overall discussion

The results did again fit nicely into the models that interpret communication as a bidirectional process (e.g. Clark, 1996; Keysar, 1998; Pickering & Garrod, 2004): Again, there were more errors during the ‘collaboration’ process of non-interactive conditions. As predicted, the numbers of all seven problems and errors coded from the verbal data were higher in non-interactive conditions. The illusion of simplicity discussed above seems to have been even more important in non-interactive conditions, since the number of features that the speaker did not identify at all was especially high in concrete tasks. Speakers in non-interactive conditions searched longer for features and adequate descriptions than speakers in interactive conditions. They tried harder to carefully design the utterances and to avoid egocentric i.e., for the addressee not understandable descriptions (Schober, 1993). However, more errors occurred during the descriptions which led to fewer pictures placed in the correct position. Why do these failures in communication occur? Fussell and Krauss (1989) argue that the perspective-taking process required from the speaker in monolog conditions involves complicated inference processes. The speaker may simplify these processes by using the same sorts of heuristics applied in social inference tasks (e.g. Nisbett & Ross, 1980; Tversky & Kahneman, 1974) and may be prone to the same sorts of biases. The false-consensus effect described by Nickerson (1999; see section 1.2.2) may also be a bias accountable for the problems in monolog conditions. Or as Fussell and Krauss (1989) put it: “[…] the salience of their own view of a stimulus may lead speakers to overestimate the degree to which is it shared by others and to underestimate its applicability to other stimuli” (p.522). Of course, these biases may also occur in dialog, but can immediately be corrected by the addressee.

Nevertheless, professional writers as newspaper reporters or novelists seem to be understood by their audience without getting feedback. They must have gained expertise through training and experience how to write in a way to be understood with minimal difficulty (Traxler & Gernsbacher, 1992). In study 2 of this dissertation the designed instructional support measures were also provided to dyads in non-interactive conditions to examine if the support had an impact on their performance. These results will be discussed in the next section.
5.1.3 Support of Collaboration

The major aim of this dissertation was to develop instructional support measures to foster net-based collaboration.

In the first study, one initial technical support measure was implemented: This was either a pictorial, textual or no shared application available. No significant effect could be found both for the process data and the performance measures. This is caused by the small number of dyads using the shared application if they had one available. As discussed above (section 5.1.2), the pictorial shared application was used more often in abstract tasks, whereas the textual shared application was used more often in concrete tasks.

If the shared applications were used, they served two different purposes: Either to indicate all the features differing between the pictures (support of individual cognitive demands) or to visually support the verbal description of each single picture by showing the feature value of each picture (support of demands due to interaction). Mostly, they were used to serve the second purpose.

However, a possible effect of the shared application could have been shown if the participants had been more informed about the usefulness of the shared applications or usage had been mandatory.

Based on the results of the first study and following the approach of Rummel and Spada (2005a, 2005b), a collaboration model and an analogical collaboration script were developed. As the first study showed different impacts due to the individual cognitive demands and the interactional demands, two levels of support were implemented, each corresponding to one demand. The amount of support was varied to evaluate the effect of both levels on collaboration process and outcome. Additionally, the support measures were also tested in non-interactive collaboration.

The mean numbers of problems and errors differed in the predicted way between the three support conditions or, if no significant effects could be found, they showed at least tendencies to differ according to the hypotheses. The specific hints emphasizing a careful search for
features and the introduction of the picture editor in level 1 led to fewer errors related to the individual cognitive demands (significantly fewer ‘feature not mentioned’ errors and tendency for fewer features not identified). The additional hints for a clear and efficient description of the features in level 2 lead to fewer errors due to the interactional demands (significant difference only for number of features not mentioned before description). These results point towards an incremental benefit of the level 2 support.

However, it can be supposed that in non-interactive conditions the effect of the support may be different: Looking on a descriptive level at the number of times an addressee rewound the audio-recordings of the speaker’s description in non-interactive conditions, only the complete support had an impact on the extent to which the speaker’s descriptions were understandable. The addressee rewound the audio-recordings of the speaker’s description less often in complete support conditions than in the other conditions. However, regarding the perceived security of having solved the task properly (dependent variable ‘need for security’), there was no difference between conditions with the complete support and conditions with level 1 support only. The simple fact to have received support seems to have been enough to reduce the dyad’s need to check the positioning of the pictures again.

Unfortunately, the positive impact of the support measures on the collaboration process did not result in better performance as there were no differences in the number of correctly placed pictures for the three support conditions.

As the collaborative task is not very complex, one could guess that motivational problems caused by overscripting are the reason for the lack of improved performance (Dillenbourg, 2002). Even if the instructional measures do not impose a specific ‘hard-wired’ procedure, the script could still be perceived as being too coercive. However, this is not really probable because the participants in the support conditions did follow the proposed procedure.

Overall, the achievement was relatively low in each task: Only about half of the pictures were placed in the correct position. Was this due to the hampered material with smaller feature differences?
To get further information on the effect of the support on the performance of the dyads, the two performance measures (numbers of correctly placed pictures and time needed for description and positioning of the pictures) were related. The analysis of this relation gives rise to the supposition that effects on performance measures could have been shown ‘in the future’, i.e., if more pictures had to be described and positioned. Furthermore, a qualitative comparison of the time structure of the dyad with complete support and the dyad without support showed that the process was smoother with complete support. The dyad with complete support did not need to search for additional differences or to correct misunderstandings after the processing of three or four pictures, whereas the dyad without support did.

Interestingly, the impact of the support measures was not different for interactive and non-interactive conditions, as there was no significant interaction between the amount of support and the mode of communication.

All in all, the process data clearly indicates an improvement through the use of the instructional support measures. Furthermore, the results suggest the need for both support levels. The combination of an on-screen video containing a model collaboration and a collaboration script to help structure the collaboration process seems to be a promising approach to foster collaboration.

5.2 Perspectives for Further Research and Practical Implications

What has been learned from the two studies that used a referential communication task with additional individual cognitive demands?

First, it could be shown that the individual cognitive demands do play a role in the process and results of collaboration and they should be taken into account in the analysis and the support of more complex and realistic tasks. Of course, in realistic tasks it is not easy to differentiate between the different types of demands. Nevertheless, the attempt to include support for all the demands arising in a specific task should be made to provide adequate and
effective support measures. Some support measures developed for collaboration include advice concerning text reading and processing or supporting one person to remember important information (e.g. O’Donnell & Dansereau, 1992; Rummel & Spada, 2005a, 2005b). These elements to support individual cognitive demands in collaboration already exist.

Furthermore, preventions for errors such as the illusion of simplicity should be included in support of collaboration. Of course, this error may also appear in individual problem solving tasks but it is more salient in collaborative tasks due to the variety of demands in collaboration, and especially in net-based collaboration. In the instructional support measures tested in the second study of this dissertation there was, in fact, an attempt to correct for the illusion of simplicity. The on-screen video in both the level 1 and the complete support conditions included explanations regarding the error of perceiving the concrete tasks as being easier to solve. This instruction helped to improve the collaboration as there was no significant interaction between the type of task and the amount of support, but there still were many features not identified in concrete tasks. In further research, instructional support measures should try harder to compensate for this bias.

The comparison of monolog and dialog conditions or interactive and non-interactive modes of communication revealed results that are in line with results of Clark and colleagues (e.g. Clark & Krych, 2004; Schober & Clark, 1989). The difficulties arising in non-interactive conditions have to be taken into account in settings that involve forms of monologs or non-interactive communication. For example, lecture style teaching in a non-interactive setting potentially holds comparable difficulties of understanding, while courses in interactive style probably do not.

The initial attempt to foster collaboration through the availability of shared applications failed. The reason for this probably lays in the implementation of the technical support measure: The participants were not forced to use the shared applications and received no clear advice on how or when to use them. Consequently, the shared applications were not used by all participants who had one available. A descriptive analysis of the data in the shared application did show different purposes of use, signaling a possible confusion among the
participants on how to use the shared pictorial or textual editors. If shared applications are provided to support the collaboration, they should be accompanied by clear instructions on how and when to use them (see Fehse, 2001).

The *instructional support measures* developed in this dissertation included two levels of support and combined a model collaboration provided as on-screen video and a collaboration script. This showed to be a promising approach to foster collaboration as the analysis of the collaboration process revealed fewer errors in dyads with support. Nevertheless, the performance measures did not differ between conditions with and without support. To have chosen a collaborative task that was more restricted, so as to be more realistic, had many advantages, but the reduction of complexity in the task may on the other hand have been responsible for the lack of visible learning results. A good way to gain further insights into skill or knowledge acquisition would be the measurement of meta-knowledge of the task demands.

To conclude: The developed support measures can clearly not be used to support other net-based collaborative tasks directly but need to be arranged and tailored according to the particular demands that arise. That said, instructional support measures that combine a model collaboration and a collaboration script and take into account the different demands arising from the collaborative task can help to improve the collaboration process and in future hopefully also the results of collaboration.
References


References


References


References


References


Rummel, N., & Spada, H. (2005a). Instructional support for collaboration in desktop videoconference settings: How it can be achieved and assessed. In R. Bromme, F. W. Hesse, & H. Spada (Eds.), Barriers and biases in computer-mediated knowledge communication and how they may be overcome (pp. 59-88). New York: Springer.


References


Appendices

Appendix A: Study 1

A.1 Materials from Study 1

Figure 33: Task 1 ‘kites’ with marked feature differences (red arrows).\(^6\)

\(^6\) The pictures are taken from the children’s game “Differix” from RAVENSBURGER. (Mit freundlicher Genehmigung des Ravensburger Spieleverlags.)
Appendices

Figure 34: Task 2 ‘radar’ with marked feature differences (red circles and arrows).

Figure 35: Task ‘cats’ with marked feature differences (red circles and arrows).

7 The pictures are taken from the children’s game “Differix” from RAVENSBURGER. 
(Mit freundlicher Genehmigung des Ravensburger Spieleverlags.)
The pictures are taken from the children’s game “Differix” from RAVENSBURGER. (Mit freundlicher Genehmigung des Ravensburger Spieleverlags.)

8 The pictures are taken from the children’s game “Differix” from RAVENSBURGER. (Mit freundlicher Genehmigung des Ravensburger Spieleverlags.)
A.2 Examples for the two Types of Usage of the Shared Applications in Study 1

Support of individual cognitive demands:

e.g. “These are the features of this task....”

Support of demands due to interaction:

e.g. “In picture 1 the red square is here...”

Figure 37: Examples for the two types of use of the pictorial shared application.  

9 The pictures are taken from the children’s game “Differix” from RAVENSBURGER.  (Mit freundlicher Genehmigung des Ravensburger Spieleverlags.)
Support of individual cognitive demands:

e.g. “These are the features of this task....”

Support of demands due to interaction:

e.g. “In picture 1...”

Figure 38: Examples for the two types of use of the textual shared application.
Appendix B: Study 2

B.1 Materials from Study 2

Figure 39: Task ‘cat’ with marked feature differences (red arrows and circles).\textsuperscript{10}

\textsuperscript{10} The pictures have been developed based on pictures of the children’s game “Differix” from RAVENSBURGER.
Figure 40: Task ‘kite’ with marked feature differences (red arrows and circles).\textsuperscript{11}

\begin{itemize}
\item The pictures have been developed based on pictures of the children’s game “Differix” from RAVENSBURGER.
\end{itemize}
Figure 41: Task ‘crochet’ with marked feature differences (red arrows and circles).
Figure 42: Task ‘circles’ with marked feature differences (red arrows and circles).
B.2 Examples of Collaboration Script for Speaker and Addressee in Study 2

Concrete task – subtask ‘search for features and labeling of features’:

<table>
<thead>
<tr>
<th>Aufgabe 1 - Merkmale suchen und Bezeichnung finden</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.) Bildeditor Aufgabe 1 öffnen und rote Farbe auswählen</td>
</tr>
<tr>
<td>2.) &quot;Start Suche Aufgabe 1&quot; drücken</td>
</tr>
<tr>
<td>3.) Unterscheidungsmerkmale suchen und in Bildeditor eingraben</td>
</tr>
<tr>
<td>4.) Texteditor öffnen</td>
</tr>
<tr>
<td>5.) Bezeichnungen für die Merkmale finden und in Texteditor eingraben</td>
</tr>
<tr>
<td>6.) &quot;Weiter&quot; drücken</td>
</tr>
</tbody>
</table>

Bitte während dieser Teilaufgabe "Merkmale suchen und Bezeichnung finden" NICHT mit deinem Partner sprechen!

Concrete task – subtask ‘description and positioning of features’:

<table>
<thead>
<tr>
<th>Aufgabe 1 - Bilder beschreiben</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.) &quot;Start Aufgabe 1&quot; drücken und dies dem Partner mitteilen</td>
</tr>
<tr>
<td>2.) Bezeichnungen für Unterscheidungsmerkmale erklären und gewählte Perspektive für Beschreibung mitteilen</td>
</tr>
</tbody>
</table>
| 3.) Bilder Position für Position beschreiben (bei 1 beginnen, dann 2...)

4.) "Fertig" drücken.

== Arbeit mit bitte so schnell und gleichzeitig so genau wie möglich! ==

<table>
<thead>
<tr>
<th>Aufgabe 1 - Bilder anordnen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.) Auf Signal von Partner warten.</td>
</tr>
<tr>
<td>2.) Nach Signal von Partner &quot;Start Aufgabe 1&quot; anklicken.</td>
</tr>
<tr>
<td>3.) Bilder Position für Position entsprechend Beschreibung in Zielanordnung positionieren.</td>
</tr>
<tr>
<td>4.) &quot;Speichern&quot; drücken, &quot;Weiter&quot; drücken.</td>
</tr>
</tbody>
</table>

== Arbeit mit bitte so schnell und gleichzeitig so genau wie möglich! ==

Figure 43: Collaboration script for speaker and addressee in interactive mode of communication/ Level 1 + 2 support for a concrete task.
Appendix C: Hardware Configuration in Study 1 and 2

Figure 44: Hardware configuration in study 1 and 2.