Designing instructional support for individual and collaborative demands on net-based problem-solving in dyads

Miriam Hansen & Hans Spada
Department of Psychology, University of Freiburg, 79085 Freiburg, Germany
Email: hansen@psychologie.uni-freiburg.de, spada@psychologie.uni-freiburg.de

Abstract: Instructional measures were designed to foster collaboration in a net-based problem-solving task. They include support for communication and for the individual and collaborative problem-solving process. A "worked-out example" of collaboration realized as an on-screen video was combined with elements of cooperation scripts. In a 2x3 factorial design, the type of support and the type of communication were varied. Audio recordings and performance measures were taken. As first steps of the data analysis, the performance measures of 24 dyads were analyzed and two in-depth case studies were conducted in order to compare the collaboration with and without the support of the developed measures. The collaborators with instructional support made fewer communication errors, managed to better structure the collaborative problem-solving process, and solved the task faster – they appear to have learned to collaborate. Further quantitative analyses of the whole sample of 48 dyads are work in progress.

Introduction

Collaboration isn't easy! This simple phrase summarizes findings from work carried out in the field of collaborative learning and collaborative work. Collaboration poses several challenges. Different kinds of skills are required: the demands of collaboration are added to individual cognitive demands inherent to a specific task. In order to learn or to solve a problem, individual actions and individual cognitive processes have to take place (see Dillenbourg, 1999), and due to the need to interact, additional activities and cognitive processes are required. These additional demands offer opportunities because of the benefit they might have on learning or problem-solving. However, collaboration also brings with it certain challenges, including coordinating the collaboration (Barron, 2000) or pooling unshared knowledge (Stasser & Titus, 1985). An 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). 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. It is therefore necessary to support remote collaboration.

By instructing collaborating partners on how to solve a remote collaborative task and improving the skills required (the collaborators *learn to collaborate*), a sustainable enhancement of the collaborative process and results can be achieved (Rummel & Spada, 2005). 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 quite complex settings. As we aimed to examine the role and functioning of communication, we used a remote collaborative problem-solving task that is more restricted and enables a more detailed analysis. The task is similar to the Referential Communication Task (Krauss & Weinheimer, 1966), which has been used a great deal to study communication, but holds more individual cognitive demands. The task we used therefore contained both individual cognitive demands and demands due to interaction. In this paper, we introduce instructional measures designed to support individual cognitive as well as collaborative demands. These measures were designed based on results of a first study that revealed the task demands (Bertholet & Spada, 2005). We will present a validation study conducted to examine the developed support measures, and will report first quantitative results as well as two case studies conducted to examine the problems that occur during collaboration. Our aim was to identify whether the collaborators learn to collaborate, and more precisely, how they *learn to communicate* under conditions of additional cognitive load posed by individual cognitive demands.

The collaborative task

Two persons have to jointly solve a picture-sorting task while located in two different rooms. One of the participants assumes the role of speaker and the other takes on the role of addressee. The task is presented on two displays (see Figure 1), and oral communication between speaker and addressee is possible via an audio link. On the speaker's display, sixteen pictures are presented that differ only in terms of minor details. The speaker has to describe nine of the pictures and their order of presentation to the addressee. The addressee sees the single sixteen pictures in a random order and has to arrange nine of them according to the speaker's description. The addressee can rearrange the pictures on the target area by using his mouse (drag and drop). Because the differences between the pictures are very small, the participants have to first detect these differences and then the speaker has to design appropriate utterances that will enable the addressee to choose the correct pictures by considering the relevant features. This component of feature detection constitutes the main difference from the classic referential communication task, in which the task demands consist only in the verbal description of clearly different objects. This individual cognitive demand makes the task more comparable to realistic collaborative tasks in which communication often has to take place under additional cognitive load.

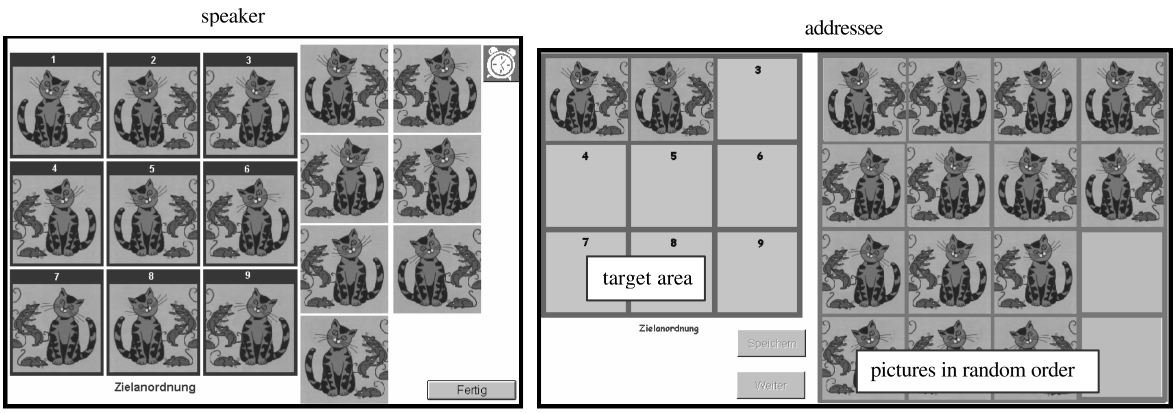

Figure 1. The pictures on the speaker's display (left) and the addressee's display (right)

Four different tasks were used. These four sets of pictures can be categorized as more "concrete", containing familiar or known objects (e.g. the pictures in Figure 1), or more "abstract", depicting unknown objects such as geometrical figures or patterns.

The demands of the task

The basic demands of the task can be grouped into "*individual cognitive demands*" and "*demands due to interaction*". The first group of demands includes feature search and identification. The collaborators have to detect the relevant features that differ between the pictures (e.g. the position of the cat, the cat's line of sight, the direction of the cat's whiskers). These demands have to be dealt with individually; they would also exist if the sorting of the pictures had to be carried out by an individual instead of being part of a collaborative task. (The original children's game "Differix" by Ravensburger©, from which we took some of the pictures, consists in arranging nine pictures on a template individually, competing against one to five other players.) The second kind of demands contains the additional challenges of collaboration due to the need to communicate. The speaker has to describe the relevant feature values (e.g. "the cat's whiskers are pointing upwards") of one picture in an appropriate way and the addressee has to understand the speaker's description and to match the description to the features in the picture.

In a previous study (Bertholet & Spada, 2005), it could be shown that the difficulties arising from the two demands are different for the two groups of pictures: The concrete sets hold more challenges with regard to the individual cognitive demands, and the abstract sets are more difficult in terms of the demands due to interaction. As the concrete sets contain familiar objects, the detection of the relevant features is more error-prone because the details can be overlooked quite easily. This result is in line with theories of object recognition (e.g., Biedermann, 2000). On the other hand, it is more difficult to describe the unfamiliar objects and their features in the abstract tasks because the speaker needs to invent terms and customize them for interaction. Speaker and addressee have to develop a common language in order to establish referential identity (Clark & Brennan, 1991). If they use different frames of reference, this results in specific errors in some concrete tasks (e.g. Taylor & Tversky, 1996). For example, the cat's line of sight can be described from the speaker's or from the cat's perspective. To enable a correct identification of the pictures, speaker and addressee must have a mutual frame of reference.

Designing instructional support measures

How is the support realized?

Several approaches attempt to foster collaboration by structuring the problem-solving and communication process. Scripted cooperation (O'Donnell & Dansereau, 1992) is a prominent technique. It aims at optimizing the interaction process by sequencing it into different phases, defining roles, and assigning them to the collaborative learners. As cooperation scripts have shown themselves to be a very effective means to foster collaboration, they have also been transferred to computer settings (e.g. Baker & Lund, 1997; Kollar, Fischer & Slotta, 2005). Typically, they are embedded into computer-based learning environments and guide the collaborators in a step-by-step fashion through different activities. Computer-based cooperation scripts may be used not only to support learners to acquire knowledge in a specific domain but also to support them to learn how to collaborate (Rummel & Spada, 2005). In order to promote the skills needed for collaboration, instructional measures may be used that are presented to the collaborators prior to the collaboration itself. A "worked-out example" of the collaborative task is such an instructional approach (Rummel & Spada, 2005): While observing a speaker or addressee 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).

In our instructional measures, the instructions prior to the collaboration were realized as on-screen videos (one for each role). Furthermore, during the collaboration, scripts were used. Figure 2 provides an overview of the support measures. The *on-screen videos* consisted of 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, for example, the screen of a speaker or an addressee solving a task. Each video was accompanied by auditory information containing hints for feature search and picture description, instructions concerning the structuring of the task as well as the roles and technical directions. The *cooperation scripts* were used to structure the collaboration process (i.e. what has do be done at what point in time?). As the task contains individual cognitive demands besides those due to interaction, it can be divided into an individual and a common part. The collaborators first have to search for the feature differences on their own before engaging in the description and positioning of the pictures. By making this division explicit, a script may help to overcome the collaborators' tendency to solely engage in joint activities and forget about individual work phases (Hermann, Rummel & Spada, 2001). Furthermore, we used the role-defining function of scripts (who does what?). In order to foster the communication process, the script prescribes, for example, that the speaker has to label the features and to propose the naming to the addressee. In contrast to the classic cooperation scripts (O'Donnell & Dansereau, 1992), the roles of speaker and addressee are not interchangeable but fixed.

		Training elements
Prior collaboration	On-screen video	• Hints for feature search (level 1) and naming of pictures (level 2) • Structuring of the task, division into subtasks, definition of roles • Technical directions
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 2. Procedure of the training with elements prior to and during collaboration.

What is supported?

Taking into account the results of the previous study, both demands – the individual cognitive demands and those due to interaction – should be supported. We developed two levels of support, each corresponding to one of the two demands. Both levels contained hints and, furthermore, individual editors as part of a subtask.

Level 1 – support of the individual cognitive demands

In the previous 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 differences, especially for the concrete sets of pictures, which appear at a first glance to be less difficult to deal with. Additionally, a subtask was introduced: The collaborators were asked to search the features that differed between the pictures and to mark them in the picture editor containing one of the 16 pictures of a 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 is crucial for successful communication and should be established as early on as possible. The subtask of level 2 consisted in the naming of the feature differences. The speaker was asked to enter the features' names into an individual text editor.

Validation of the resulting training

The instructional support measures that we designed result altogether in a training that fosters collaboration. The effects of the training were examined in an experimental validation study. In this paper, we present two case studies in order to illuminate the communication and problem-solving process of one dyad collaborating with the support of our training and one dyad collaborating without such support. As the collaboration of the two dyads was part of the validation study, we will begin by briefly describing the design of the study before reporting the results of the case studies.

Design of the study and procedure

A 2x3-factor design was used. As one between-subject factor, the amount of support was varied. Either there was no training at all, or only the individual cognitive demands were supported (only level 1), or the dyad received the complete training (levels 1 and 2). As a second between-subject factor, the mode of communication was varied, being either synchronous or asynchronous. A synchronous mode of communication means that speaker and addressee communicated through an audio link at the same point in time. The results section will include two in-depth case studies of dyads performing within the synchronous communication condition as well as a quantitative analysis of the performance of all 24 dyads in the synchronous condition. As 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. 96 students (48 dyads) of the University of Freiburg, Germany participated in the study. Figure 2 shows the procedure of the complete training. To keep the procedure and time-on-task comparable, the dyads without training were also shown an on-screen video prior to the collaboration, but just containing technical instructions. During collaboration, only the dyads with training were supported by a cooperation script; the dyads without training collaborated without support. All dyads had the possibility to first look individually at the 16 pictures of a task before starting the description and positioning phase. The dyads with training received instructions to use this time to search individually for the differences between the pictures.

Measures

Two sets of dependent measures were collected to examine the collaboration process as well as the outcome: audio recordings of the verbal communication and performance measures. We developed a coding scheme for the verbal communication data that emphasized the time structure and different kinds of errors committed during collaboration. Corresponding with the subtasks introduced in the on-screen video as well as the cooperation script (see Figure 2), the ideal *time structure of the collaboration* contains (1) an initial individual period to search for the features and to name them (= "individual time"), (2) a common period used for grounding (= "grounding time"), and (3) a common period for the description and positioning of the nine pictures (= "description/ positioning time"). The collaboration process was divided into these three periods. The third period can further be divided into the time needed for the description and positioning of each of the nine pictures. The *errors committed during collaboration* are: (1) number of not identified features in a task (error of speaker and addressee), (2) number of times the speaker failed to mention a relevant feature in the description of a picture, (3) number of times the speaker gave ambiguous descriptions with regard to the referential identity, and (4) number of irrelevant features mentioned by the speaker. The errors refer to one of the two demands on the collaborators, (1) and (2) to the individual cognitive demands and (3) and (4) to the demands due to interaction.

The performance measures included the number of pictures placed in the correct position at the end of one task, the time used to look at the pictures individually before starting the description/ positioning (separately for speaker and addressee), and the time needed to complete the task. These measures of the time spent are not equal to the time structure of the collaboration process described above, because they only consist in computer-based measurement and do not reveal what the collaborators exactly did during collaboration.

Results of the two case studies

One dyad had the support of the complete training ("with training") while the other was not supported at all ("without training"). Both dyads solved all four tasks in the same order (tasks 1 and 3 concrete, tasks 2 and 4 abstract). We will now take a closer look at the collaboration process. The *time structure* differed considerably with regard to the existence or duration of the periods: The dyad with training took between 140 and 200 seconds of individual time, between 50 and 120 seconds of grounding time, and between 290 and 370 seconds of description/ positioning time per task. The dyad without training, by contrast, took practically no individual time at all (between 7 and 24 seconds), and their collaboration did not include grounding time in tasks 1 to 3. Only in the last task were there 320 seconds of grounding time. Instead, the dyad without training took between 860 and 1200 seconds of description/ positioning time. Figure 3 depicts the time needed for the description and positioning of each single picture separately for the two dyads. The dyad with training took a little more time for the first picture, but altogether the time did not differ a great deal for the different pictures. For the dyad without training, the especially large amount of time for the first picture is striking, and is followed by a smaller peak around the fourth and fifth pictures. The large amount of time for the first picture is due to the search for the feature differences and the grounding process. Both were carried out only while the speaker was trying to describe the first picture. As the search and grounding were often not done properly (especially in tasks 1, 2 and 4), problems occurred after the description and positioning of some pictures: The peak around the fourth or fifth picture is due to repair mechanisms as a restart of feature search, repeated grounding or a too detailed description of the picture containing a lot of irrelevant features.

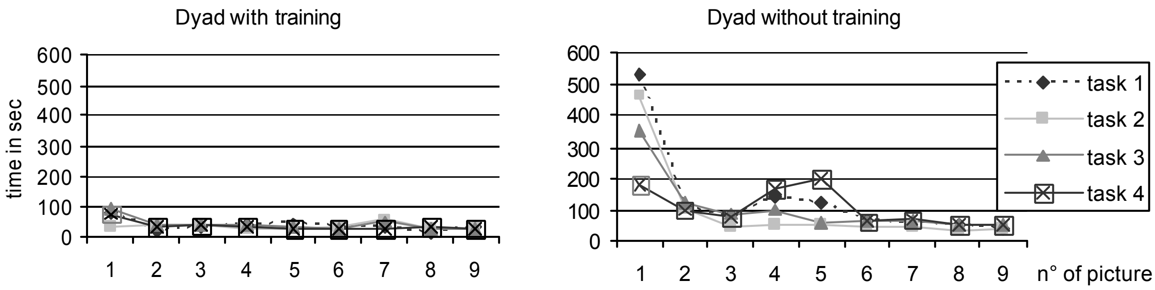

Figure 3. Time needed for description and positioning of each single picture for both dyads.

Regarding the *errors committed during collaboration*, the differences are even more striking. Table 1 shows the four types of errors for the dyad with and the dyad without training. The dyad without training showed a great number of errors of all four types, whereas only one error could be detected in the collaboration process of the dyad with training.

Table 1. Number of errors committed during collaboration.

	N not identified features	N speaker did not mention a feature	N speaker did not establish ref. id.	N irrelevant features
With training	0	1	0	0
Without training	2	31	12	68

The *performance measures* showing the result of the collaboration indicate a clear advantage for the dyad with training: In all four tasks, the dyad without training took longer to complete the task (mean with training = 640 sec and mean without training = 1210 sec) and, for the concrete tasks, arranged fewer pictures (only 6 of 9) in the correct position. For the abstract tasks, both dyads placed all pictures in the correct position.

Quantitative results of the performance measures – synchronous condition (24 dyads)

A multivariate analysis of variance (MANOVA) with repeated measures (for the within-subject factor) was conducted in order to test the influence of the type of support and the type of task on the performance measures. There were significant effects for both factors: The dyads with the complete training implemented the instructions, as they used more time to look at the pictures individually (speaker: $F[2] = 19$; $p<.000$; addressee: $F[2] = 16.5$; $p<.000$). Furthermore, the complete training led to better results. The number of correctly placed pictures did not differ between the support conditions, but with complete training less time was needed to complete the task. In line with the results of the first study, more pictures were placed in the correct position in abstract tasks ($F[1] = 9.5$; $p<.000$). The collaborators took more time to look at the pictures individually in abstract tasks (speaker: $F[1] = 13.2$; $p=.001$; addressee: $F[1] = 15.6$; $p<.000$) and needed less time to complete them ($F[1] = 5.9$; $p=.024$). Furthermore, we can assume from the results that problems with different frames of reference occurred if collaborating without the complete training.

Discussion

The aim of the case studies was to examine the problems that can occur during the collaboration process and to scrutinize whether, through the developed training, the collaborators could learn to overcome these problems and, in particular, communicate adequately. The dyad without support faced various problems: Speaker and addressee did not structure the process or balance individual and common work phases, were confused about how to deal with the different demands at the same time, and thereby committed many errors. The uncertainty of the speaker as to how to solve the task appropriately may, for example, be seen in the number of irrelevant features he included in his descriptions. By contrast, the dyad with training seemed to have learned how to collaborate: The collaboration process was better organized, clearly structured and contained nearly no communication errors. This led to results that were greatly superior to those of the dyad without training. In contrast to some evidence from scripted collaboration, we were not able to detect motivational problems in the dyad that collaborated with scripts.

The combination of an on-screen video containing a "worked-out example" of collaboration and a cooperation script to help structure the collaboration process therefore seems, so far, to be a promising approach to foster collaboration. The results of the quantitative analysis of the performance measures for the 24 dyads in the synchronous condition also indicate that the training measures help to improve the collaboration, as the dyads with training took less time to complete the tasks. Furthermore, the results suggest the need for both training levels. Especially with concrete tasks, the support of both demands is crucial and should contain instructions to assure a mutual frame of reference.

Further quantitative analyses of the whole sample of 48 dyads (including the synchronous as well as the asynchronous condition) are in progress. It is also planned to analyze the collaboration process of all 48 dyads in the same way as in the case studies. In a further step, the training could be transferred to other collaborative settings in order to support the various demands on the collaborators and to improve the results of the collaboration. In the future, the introductory statement could then perhaps be changed to: Collaboration may not be easy – but it is not too difficult with the right support!

References

Baker, M., & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning, 13*(3), 175-193.

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences, 9*, 403-436.

Bertholet, M., & Spada, H. (2005). Cognitive analysis of a remote collaborative problem-solving task and experimental validation of the resulting demand model. In K. Opwis & I.-K. Penner (Eds.), *Proceedings of the German Cognitive Science Conference 2005* (pp. 15-20) Basel: Schwabe.

Biederman, I. (2000). Recognition-by-components: A theory of human image understanding. In S. Yantis (Ed.), *Visual perception: Essential readings* (pp. 320-340). Philadelphia, PA: Psychology Press.

Clark, H. H. (1996). *Using language*. New York, NY: Cambridge University Press.
Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.
Dillenbourg, P. (1999). What do you mean by 'collaborative learning'? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-19). Oxford: Elsevier.
Hermann, F., Rummel, N., & Spada, H. (2001). Solving the case together: The challenge of net-based interdisciplinary collaboration. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *Proceedings of the First European Conference on Computer-Supported Collaborative Learning* (pp. 293-300). Maastricht: McLuhan Institute.
Kollar, I., Fischer, F., & Slotta, J. D. (2005). Internal and external collaboration scripts in webbased science learning at schools. In T. Koschmann, D. Suthers, & T. -W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years* (pp. 331-340). Mahwah, NJ: Lawrence Erlbaum.
Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology, 14*, 343-346.
Kraut, R. E., Fussell, S. R., Brennan, S. E., & Siegel, J. (2002). Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. In P. Hinds & S. Kiesler (Eds.), *Distributed work* (pp. 137-162). Cambridge, MA: MIT Press.
Mayer, R. E. (2001). *Multimedia learning*. Cambridge: University Press.
O'Donnell, A. M., & Dansereau, D. F. (1992). Scripted cooperation in student dyads: A method for analyzing and enhancing academic learning and performance. In R. Hertz-Lazarowitz & N. Miller (Eds.), *Interaction in cooperative groups: The theoretical anatomy of group learning* (pp. 120-141). New York: Cambridge University Press.
Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*, 1-29.
Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem-solving in a desktop-videoconferencing setting. *The Journal of the Learning Sciences, 14*(2), 201-241.
Stasser, G., & Titus, W. (1985). Pooling unshared information in group decision making: Biased information sampling during group discussion. *Journal of Personality and Social Psychology, 48*, 1467-1478.
Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory & Language, 35*(3), 371-391.
VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47*, 513-539.

Acknowledgments

This work is part of the Virtual PhD Program (VGK) "Knowledge Acquisition and Knowledge Exchange with New Media", which is supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG).