

A New Method to Assess the Quality of Collaborative Process in CSCL

Hans Spada, Anne Meier, Nikol Rummel, Sabine Hauser

Department of Psychology

University of Freiburg, Germany

<first name.last name>@psychologie.uni-freiburg.de

Abstract: In CSCL research, the collaborative process – the way people collaborate while working on tasks and learning – is of central importance. Instructional measures are being developed to improve the quality of the collaboration which itself determines to a great extent the results of working and learning in groups. However, assessing collaborative process is not easy. We have developed a new assessment method by quantitatively rating nine qualitatively defined characteristic dimensions of collaboration. In this paper, we first describe how these dimensions were extracted from video-recordings of dyads collaborating to solve interdisciplinary tasks. Then we explain how the resulting rating system was applied to and tested on another sample. Based on positive findings from this application, we argue that the new method can be recommended for different areas of CSCL research.

Keywords: Collaborative Process, Assessment Method, Rating System, Videoconferencing, Cognitive Dimensions, Affective Dimensions

In CSCL research, the collaborative process – e.g. the way co-learners exchange information, discuss different perspectives, take on diverse roles, coordinate their efforts in solving a joint task, or make use of technological tools – is of central importance. The quality of the collaborative process determines to a great extent the results of working and learning in groups. Instructional measures are successful if they are developed based on insights about what features of the collaborative process are relevant for successful learning and problem-solving. But analyzing and assessing collaborative process is not easy, and usually very time-consuming. In this paper, a rating system is presented that can be used to evaluate the quality of the collaborative process while reviewing it on videotape, without the need of time-consuming transcription.

In the following, a short overview of methods already used in assessing collaborative process is given. We also briefly describe the instructional experiment the data of which were used in developing the new assessment method. Next, we describe the three steps that were taken in developing and evaluating the new assessment method. First, a combination of a data-driven and a theory-based approach was used to extract nine characteristic dimensions of collaborative process that were afterwards implemented in a rating system. Second, the rating system was applied to another sample and evaluated with regard to inter-rater reliability and process-outcome validity. A further approach to testing the relevance of the rating system's dimensions involved implementing them in instructional support measures and comparing the results of instructed and non-instructed dyads of collaborators.

METHODS FOR ASSESSING COLLABORATIVE PROCESS

Throughout the learning sciences, assessing and analyzing collaborative process has become a central research topic. At the International Conference of the Learning Sciences (ICLS) 2004 at Santa Monica, for example, a symposium was devoted to discussing adequate ways to record, analyze and interpret what happens during collaborative process, with the long-term goal of assembling a “methodological toolbox” (Rummel & Spada, 2004). Many researchers in CSCL agree that the process of collaboration, in addition to traditional outcome measures, should be paid closer attention (e.g. Nurmela, Palonen, Lehtinen, & Hakkarainen, 2003). Some typical methods already in use include content analysis, discourse analysis, analysis of computer-generated quantitative log files, and social network analysis (Häkkinen, Järvelä, & Mäkitalo, 2003).

Log file data, which can be automatically generated and stored by the learning environment, can serve as an easily accessible data base for analyzing collaborative process. These log file data can be

used to identify activity patterns and participation structures in networked learning groups, which can also be graphically displayed (Nurmela et al., 2003). However, Nurmela et al. (2003) warn researchers not to rely primarily on the information provided by log file data (for example because one can never be sure whether an opened document is actually read), but to combine these structural analyses of the collaborative process with an analysis of its contents, especially the content of collaborative dialog.

Different coding schemes have been developed in order to label and to quantify what happens during collaborative process. One coding scheme that has been successfully employed in studies analyzing dialog from collaborative learning sessions (e.g. Kneser & Ploetzner, 2001; Pilkington & Parker-Jones, 1996) is the DISCOUNT scheme developed by Pilkington (1999). Aims of DISCOUNT include identifying dialog roles, tracking initiative and describing an episode's content structure. The system is applied in a hierarchical fashion: conversational episodes concerning a particular topic are broken down into exchanges, exchanges into turns, and turns into moves or even further into rhetorical predicates. The coding scheme provides the researcher with a large set of codes concerning the structure and function of these components. Researchers implementing the DISCOUNT scheme also use it to identify roles that learners take on. For example, Kneser and Ploetzner (2001) distinguished between the roles of information seeker, explainer, task performer and reflector. Bruhn, Gräsel, Fischer, and Mandl (1997) presented a different system of categories specifically designed to assess processes of knowledge co-construction in learners' discourse. These researchers suggest three central mechanisms of knowledge co-construction: externalization of knowledge, elicitation of knowledge, and different kinds of consensus building.

While most of the methods for analyzing collaborative process allow the researcher to quantify aspects of collaborative dialog and to identify particular interaction patterns and roles (e.g. the number of elicitations, the frequency of taking the role of a reflector, or the amount of time spent on coordination), it has been criticized that little is being said about the quality of the collaborative process (Häkkinen et al., 2003). One approach to assessing the quality more directly has been taken by Häkkinen et al. (2003) who developed a theory-based analysis method for rating the level of perspective taking in text-based online discussions, taking into account five distinct stages. Collazos, Guerrero, Pino, and Ochoa (2004) developed a set of five indicators in order to describe the interaction within groups that differed in the quality of their cooperative process and outcome. Other approaches have been completely data-driven and qualitative in nature, often following the ethnographic research tradition. These researchers (e.g. Guribye, Andreassen, & Wasson, 2003) placed their emphasis "on identifying concepts and patterns as they emerge from the data" (p. 388), for example when trying to understand which interactional processes are necessary in organizing distributed collaborative learning. In the focus of attention of Koschmann, Zemel, Conlee-Stevens, Young, Robbs and Barnhart (in press) have been sequences or patterns of actions through which group members achieve effective cooperation. For example, these authors were able to demonstrate "problematizing", i.e. a move by which participants call into doubt assumptions previously held by a group of learners. Ethnographic approaches are very helpful tools in identifying relevant aspects of the collaborative process, but usually do not provide quantitative results.

Our goal in developing a new assessment method has been to combine the benefits of data-driven as well as theory-driven approaches, and qualitative as well as quantitative methods. First, relevant dimensions of the collaborative process were extracted from the data in a qualitative procedure. Then these dimensions were implemented in a rating system that enables the user to evaluate the quality of collaborative process in a quantitative way, such that the resulting ratings can be subjected to statistical analyses.

OUR RESEARCH CONTEXT

The development of our new method for assessing the quality of collaborative process was embedded in a study on instructional support for computer-supported collaborative problem-solving given complementary expertise of the collaborating partners (Rummel & Spada, 2005). Dyads, each consisting of a medical student and a student of psychology, collaborated via a desktop video-conferencing system. Their task was to develop a diagnosis and a therapy plan for a given psychiatric case, which was carefully designed to require the combined application of both medical and psychological expertise in order to be solved correctly. The videoconferencing system allowed participants to see and hear each other while discussing the case. It included a shared workspace the students could use to prepare a written solution. The dyads were given two hours to solve the case, and their collaboration was videotaped. Prior to this testing phase, half of the dyads had undergone a learning phase in which they had been instructed on how to collaborate. The main goal of the study was

to compare different methods of instructional support. As data from this study were used for developing and evaluating the new assessment instrument, a short overview of the different experimental conditions is given in Table 1.

As part of this research project we have already developed, applied and evaluated several approaches for analyzing collaborative process (Rummel & Spada, in press). A first approach was based on log-file data. We counted, for example, the minutes of individual versus joint work during problem-solving; this resulted in the finding that successful dyads showed significantly longer individual work phases. To enable a more fine-grained analysis, a number of video recordings were transcribed and the dialogs coded with regard to criteria of coordination, communication, and the topics discussed. Then the instances of particular types of coordination (e.g. minutes of talk on division of labor), of communication (e.g. turns explaining new content to the partner) and of turns with specific topics were counted. Only the analyses of the coordination revealed systematic differences between successful and unsuccessful dyads. A general problem of quantifying qualitative data by coding and counting is that the number of utterances of a particular type does not provide enough information for evaluating the quality of the collaborative process. For example, more coordinative utterances do not necessarily indicate better collaboration, because too much coordinative dialog reduces the time available for the task itself. Too many coordinative utterances might even be an indicator of failed attempts to coordinate collaboration efficiently. Therefore, we decided to develop a new method that would allow us not only to describe the collaborative process in quantitative terms, but also to assess its quality. In the remainder of this paper we will present the three steps that we have taken in the development and evaluation of this new assessment method. Table 2 gives a short overview of the data used, the methods applied, and the results obtained.

Table 1: Experimental conditions in the study by Rummel and Spada (2005) on the effects of two instructional measures on collaborative work and learning

	Learning phase	Testing phase
Model condition (9 dyads)	observational learning	uninstructed collaboration
Script condition (9 dyads)	scripted collaboration	
Unscripted condition (9 dyads)	uninstructed collaboration	
Control condition (9 dyads)	no learning phase	

Table 2: Data sources, methods and results in the development and evaluation of a new method to assess characteristic dimensions of collaborative process

<p>Extracting characteristic dimensions of collaborative process and developing a rating system</p> <p><i>Data source:</i> for extracting dimensions: video-recordings of the collaboration in the testing phase and transcribed dialog of 4 dyads (2 unscripted condition and 2 control condition); for developing the rating system: transcribed dialog of these 4 dyads plus 3 additional dyads (2 model condition, 1 script condition)</p> <p><i>Method:</i> a thorough data-driven, qualitative analysis of the collaborative process of these dyads, combined with theoretical considerations based on the relevant literature; development of a rating system</p> <p><i>Results:</i> nine dimensions of collaborative process and a rating system allowing to assess them quantitatively</p>
<p>Evaluating the developed rating system with regard to inter-rater reliability and validity</p> <p><i>Data source:</i> video-recordings of collaborative work in the testing phase, and measures assessing the quality of the solution to the psychiatric case for 9 dyads (control condition)</p> <p><i>Method:</i> applying the rating system to the collaborative process of these dyads and assessing inter-rater reliability and measures of validity by calculating process-outcome correlations</p> <p><i>Results:</i> inter-rater reliability sufficient to high; high validity</p>
<p>Testing the relevance of the nine dimensions by implementing them in instructional support measures</p> <p><i>Data source:</i> data on the quality of the solution of the case from the already reported comparison of 18 dyads with and 18 dyads without instruction (Rummel & Spada, 2005)</p> <p><i>Method:</i> instructing 18 dyads on how to collaborate and comparing their outcome with that of non-instructed 18 dyads</p> <p><i>Results:</i> instructed dyads produced better outcomes → the dimensions concern relevant aspects of collaborative process</p>

EXTRACTING CHARACTERISTIC DIMENSIONS OF COLLABORATIVE PROCESS AND DEVELOPING A RATING SYSTEM

Method

In identifying relevant aspects of the collaborative process, we combined a bottom-up, data-driven and a top-down, theory-driven approach. First, in the data-driven approach, a multi-step analytical procedure built on the qualitative content analysis developed by Mayring (2003) was followed to identify process dimensions relevant for a successful collaboration (Sosa y Fink, 2003). Mayring's qualitative content analysis involves the data-driven, inductive development of categories through a stepwise reduction of transcripts, until the desired level of abstraction has been reached. In order to be able to analyze "naturally" occurring collaboration, we selected four dyads that had not received any prior instruction on how to collaborate. The collaborative dialog was transcribed. Utterances were paraphrased, generalized, and bundled into concepts according to Mayring's rules of qualitative content analysis. Higher-level concepts were formulated, and lower-level concepts subsumed. At a relatively high level of abstraction, seven categories resulted, representing characteristic features of the collaborative process. However, this set of inductively derived categories posed the problem of being not precisely enough defined and partly overlapping in content. Therefore, a complementary theory-driven approach was undertaken in order to separate them more clearly from each other, and ground them in theoretical concepts from the literature. We reviewed the literature on computer supported collaboration in order to identify aspects characteristic for successful collaboration. The focus was on dimensions of collaboration that could be directly observed from the videotaped interaction process. We neither wanted to analyze single speech acts, like in many fine-grained discourse coding schemes, nor were we looking for universal features of collaboration. Instead, we were interested in actions and interaction patterns that could be judged to be appropriate or inappropriate within the context of the given cooperative scenario. Integration of the result of the data-driven analysis with our theoretical considerations led to nine dimensions for assessing collaborative process. Finally, a rating system was developed containing a description of each of these nine dimensions, along with illustrating examples of interaction patterns and instructions on how to rate the dimensions quantitatively.

Results: Nine Dimensions of Collaborative Process and the Resulting Rating System

Successful collaboration is not possible without effective communication. In accordance with the communication theory put forward by Clark (e.g. Clark & Brennan, 1991), two important features of the communicative process are included in the rating system: *sustaining mutual understanding* (Dimension 1) and *coordinating communication* (Dimension 2). Further, collaborative problem-solving and learning can in large parts be seen as a question of information processing at the group level (Hinsz, Tindale, & Vollrath, 1997). The third and fourth dimension therefore concern processes of constructing a shared knowledge base. Two kinds of processes are distinguished, though these cannot be seen as independent: *pooling information* (Dimension 3) and *reaching consensus* (Dimension 4). Finally, collaboration can also be seen as a matter of coordination (e.g. Malone & Crowston, 1994, Barron, 2000). The focus in our rating system is on three content-unspecific aspects of coordination: *task division* (Dimension 5), *time management* (Dimension 6), and *technical coordination* (Dimension 7). In addition to these seven more cognitive oriented dimensions, two dimensions concerning motivational aspects were formulated: *shared task alignment* (Dimension 8) and *sustaining commitment* (Dimension 9). In the rest of this section, these nine dimensions are presented together with a brief glance at their theoretical background and some examples of the operationalization put forward in the rating scheme we developed. The nine categories were defined in a way to be task unspecific, i.e. they should be suitable to evaluate the quality of collaborative process for any similar task under the conditions of complementary expertise and a desktop videoconferencing setting. Table 3 gives a short overview over the resulting nine dimensions, which will subsequently be described in more detail.

Dimension 1: Sustaining mutual understanding

Sustaining mutual understanding is also known as the problem of "grounding" in communication (Clark & Brennan, 1991). Similar concepts are "convergence on central concepts", or "joint problem space" (Roschelle & Teasley, 1995). Clark and Brennan (1991) list a couple of "positive evidences" for ascertaining mutual understanding, which can be analyzed in videotaped collaboration: acknowledgements, "relevant next turns" demonstrating that the speaker has understood and is referring to what was said before, and continued attention. In a similar way the communication framework put forward by Whittaker and O'Conaill (1997) distinguishes between reference, feedback

and interpersonal cues used to coordinate the content of communication. The description in our rating scheme says that for this dimension the rater should assess, among other things, whether speakers try to make their contributions understandable (e.g. by explaining technical terms), give their partners the opportunity to ask questions and elicit feedback from their partner. Both partner should listen to each other carefully, signal their continued attention and give feedback of their understanding. As a result, the collaborators' utterances should be relating to each other.

Dimension 2: Coordinating communication

Coordinating communication refers not to the content but the process of communication. This category, which is based on the "process coordination" dimension in the framework of Whittaker and O'Conaill (1997), includes processes of turn-taking and of managing the beginning and ending of conversational episodes. In videoconferencing, collaborators can facilitate turn-taking by explicitly handing over a turn, for example by naming the next speaker or posing a question (O'Conaill & Whittaker, 1997). Conversational episodes, for example between two phases of parallel individual work, should further have a clear beginning and ending. This dimension is rated depending on how smoothly the conversation is "flowing", how well the turn-taking is being managed, and whether participants try to secure their partners attention before starting a new conversational episode.

Dimension 3: Information pooling

Information pooling, especially the pooling of unshared information, is a crucial aspect of successful collaborative problem-solving (e.g. Stasser & Titus, 1985) and knowledge construction, and even more so under the condition of complementary expertise. Information pooling is mainly a matter of externalizing knowledge (Bruhn et al., 1997), but also of asking each other questions and giving explanations. Asking for as well as giving information will be more effective if both partners keep their complementary expertise in mind (as a form of metaknowledge which helps to ensure that relevant unshared information is brought into the discussion), using their partner as a resource (Dillenbourg, Baker, Blaye, & O'Malley, 1995) and also taking over the responsibility for their own domain. Finally, explanations must be given at an appropriate level of elaboration in order to be helpful (Webb, 1989). The rater should pay attention to the following aspects: Both partners should try to contribute as much information as possible, especially the distributed information. New information should be given in an elaborated way, for example illustrated through concrete examples, and be put into the context of the task at hand.

Dimension 4: Reaching consensus

Ideally, reaching consensus, e.g. concerning a decision, should be preceded by a process of critically evaluating the given information, collecting arguments for and against the options at hand and critically discussing different perspectives. This should result in socio-cognitive conflict, which is seen as very important for learning from collaboration by many authors (see for example Dillenbourg et al., 1995), and a rather "conflict-oriented" style of negotiation (Fischer & Mandl, 2002). However, as these authors point out, in computer-mediated as well as in face-to-face collaboration, participants tend to avoid conflict, trying instead to integrate their individual perspectives without really discussing them, often resulting in a "superficial conflict-avoiding cooperation style" and "an illusion of consensus". The dimension should be rated reflecting to what extent the "ideal" way of reaching consensus was followed, especially whether proposals were critically reflected by both partners, thus avoiding a superficial consensus. The point at which a final decision is made should be clearly identifiable.

Dimension 5: Task division

Task division in general involves decomposing an overall goal into subgoals and delegating the resulting subtasks to different persons (Malone & Crowston, 1994). Further, it has been shown that particularly in the case of partners with complementary expertise, there should be both joint and individual work in a well-balanced proportion (Hermann, Rummel, & Spada, 2001). On the one hand, individual phases are important so the experts can bring their individual domain knowledge to bear; on the other hand joint phases are necessary to ensure a shared understanding of the problem to be solved, and to integrate the individual work into a coherent joint solution. The rater will observe in how far the task is split into subtasks and in how far individual as well as joint phases of work and learning are distinguishable. Drafting a plan of how to divide the task and delegate the work in the beginning together with several coordinative episodes throughout the collaboration is considered ideal. Tasks should be defined and delegated according to the partners' expertise.

Dimension 6: Time management

Time management is necessary, if (as in our scenario and probably in most CSCL settings) the time available is limited. In addition to dividing the tasks at hand into several subtasks, a suitable amount of time needs to be allotted to each working phase. In our scenario we consider it to be ideal if participants

take some time at the beginning of their collaboration in order to draft a schedule identifying the planned working phases. In rating this category one should pay attention to the following aspects: Each subtask should be allotted a certain amount of time which must both be short enough so the whole task can be finished in time and long enough so the work can realistically be done. Adherence to the schedule should be monitored throughout the collaborative process, for example by reminding each other of time limits.

Dimension 7: Technical coordination

In computer-mediated collaboration the aspect of technical coordination needs to be added to task division and time management. With Malone and Crowston (1994), coordination can be defined as managing interdependencies between activities. What distinguishes “good” technical coordination will always depend on the dependencies and the resources available within each specific computer-mediated collaboration setting. In our scenario, the dependency consists of the shared resources the desktop videoconference system provides. Collaborators have to coordinate their activities in a way that they do not impair each others work. For example, they have to clarify at any given time who may write into the shared text editor, which does not allow for simultaneous typing, or when to switch on and off the speakers for phases of individual work. Ideally, collaborators should make use of all the technical possibilities they have in order to facilitate their working process. All these aspects should be taken into account when rating this category.

Dimension 8: Shared task alignment

The term shared task alignment was borrowed from Barron (2000), who uses it to describe a collaborative orientation toward problem solving. Shared task alignment, as defined by Barron, refers to a certain way of coordinating the collaboration, e.g. by co-orienting actions around the task and taking up and expanding each others’ contributions. Our category also comprises accepting the shared task and taking on responsibility for its solution (i.e. striving to reach a good outcome), and supporting each other during collaboration. The rater judges how readily the partners take over responsibility for their joint task, how much interest and effort they put into their work, and in how far they seem willing to support each other in this process. Showing joy and/or pride during collaboration or as result of the joint accomplishment is also seen as a positive indicator for shared task alignment.

Dimension 9: Sustaining commitment

While shared task alignment describes the basic orientation participants show toward their collaborative task, sustaining commitment aims at those processes necessary to keep up a high level of task involvement and expended effort. Above all, the collaborators’ attention needs to be focused on the problem to be solved, so the problem-solving process is not impaired by competing action tendencies. There are a couple of strategies useful for the purpose of keeping up one’s motivation. Collaborators can set goals they want to reach and reward themselves (and each other) for progress toward solving the problem. If the collaborators experience failures, they should focus their attention back on the task, and if they feel their own or their partner’s motivation is decreasing, they should remind each other of the positive consequences solving the problem will have or formulate positive expectancies (e.g. that their combined abilities will suffice to solve the problem in a satisfying way). The occurrence of strategies like these is the basis for rating this category.

The Rating System

The resulting rating system contains a detailed description of each of the nine dimensions, along with questions intended to guide the rater’s attention toward certain aspects of the collaborative process. In order to further illustrate the dimensions, the transcripts of seven dyads (among them those four used during the data-driven analysis) were searched for fitting discourse episodes. For example, the following episode was selected to illustrate how dyads can sustain mutual understanding:

Dyad 14, Minute 04: Psychology student: “....Did you understand what I just said?” Medical student: “Uh-uh. That is, you mean, whether now there is a psychotic component in addition to the depression and the multiple sclerosis?” Psychology student: “Exactly!”

Instructions are given on how to rate each of the nine dimensions on a seven-point-scale from “very bad” to “very good”. The rating is done by reviewing the video-recording of the collaborative process for each dyad. The rating sheet leaves room under each dimension so raters can take notes concerning their impression of the dyad’s performance in order to aid their memory.

EVALUATING THE NEW RATING SYSTEM: INTER-RATER RELIABILITY AND PROCESS-OUTCOME VALIDITY

This paragraph describes how the instrument was applied to a sample of nine dyads in order to evaluate inter-rater reliability and the dimensions' correlations with an outcome criterion.

Method

The instrument was applied to a sample of nine dyads which collaborated freely, i.e. without prior instruction, in order to see whether the rating system was suitable to assess “natural” collaboration as it occurs in a computer-mediated setting. The sample was made up of the nine dyads in the control condition of the already mentioned experiment (Rummel & Spada, 2005; see Table 1). Transcripts of two of these dyads had already been used for the data-driven category development by Sosa y Fink (2003). All dyads were rated by two raters (A. Meier and S. Hauser); two dyads were rated jointly for training, the other seven dyads independently. To assess inter-rater reliability only the data of these seven dyads were used. Then, for all nine dyads, the ratings of the nine dimensions were correlated with an outcome criterion measuring the quality of the joint solution produced by the dyads.

Results

While working with the newly developed instrument, the raters gained the impression that the nine dimensions did indeed allow to differentiate between good and bad collaboration. All results of the statistical analyses are given in Table 4.

Table 4: Interrater-reliability of the 9 dimensions, their intercorrelations and the correlations with an outcome measure

Dimension	Interrater-reliability (intraclass correlation) (n = 7)	Correlations (Pearson, n= 9)								
		Sustaining Mutual Understanding	Coordinating Communication	Information Pooling	Reaching Consensus	Task Divison	Time Management	Technical Coordination	Shared Task Alignment	Sustaining Commitment
(1) Sustaining Mutual Understanding	.74*	--	.43	.82*	.77*	.43	.16	.34	.78*	.55
(2) Coordinating Communication	.88*		--	.08	.33	.84*	-.19	.77*	.20	.29
(3) Information Pooling	.63*			--	.83*	.00	.12	.08	.69*	.39
(4) Reaching Consensus	.87*				--	.35	-.14	.29	.64	.57
(5) Task Divison	.84*					--	.02	.90*	.43	.64
(6) Time Management	.87*						--	.11	.26	.36
(7) Technical Coordination	.45							--	.52	.68*
(8) Shared Task Alignment	.70*								--	.79*
(9) Sustaining Commitment	.56									--

* significant on the 0.05-level

Inter-rater agreement proved to be not perfect, but acceptable: From the seven independent ratings, intraclass coefficients (two way mixed effects model) were calculated as a measure of inter-rater reliability for each of the nine categories (see Table 4). The intraclass correlation was found to exceed .70 for all but three categories. It was highest for “coordinating communication”, “reaching consensus” and “time management” and lowest for “sustaining commitment” and “technical coordination. The rating instructions for the three dimensions with an inter-rater reliability below .70 are currently being revised. For the further analyses, the mean value of the two independent ratings was calculated for each dimension. For all dyads (n = 9), correlations of the nine dimensions with each other and with an external criterion - the quality of the joint solution (i.e. the outcome of the collaboration process) - were calculated. All results are given in Table 4. Based on this small sample of

nine dyads, statistical significance is only given in the case of very high correlations ($r > .67$). In the moment, the rating system is applied to a further and larger sample of a new experiment.

Not surprisingly, related categories inter-correlate moderately to highly. For example, high correlations were found between the two categories assessing the process of building a shared understanding of the problem, “information pooling” and “reaching consensus”, and the two categories assessing motivational aspects, “shared task alignment” and “sustaining commitment”. Summarizing these results, it can be concluded that the nine dimensions draw a rather coherent picture: Good dyads collaborate well concerning most of the dimensions.

For the quality of the joint solution (combined scores for the diagnosis and therapy parts), high correlations were found for “sustaining mutual understanding”, “coordinating communication”, “task division”, “technical coordination”, “shared task alignment”, and “sustaining commitment”. The lowest correlations were obtained for “information pooling” and “reaching consensus”. However, the processes assessed by these two dimensions were relevant for the first part of the joint solution, the diagnosis. Accordingly, they yielded higher correlations with the diagnosis score alone ($r = .67^*$ for “information pooling” and $r = .52$ n.s. for “reaching consensus”). Thus, the predictive validity for the outcome is moderate to high for all dimensions.

TESTING THE RELEVANCE OF THE NINE DIMENSIONS BY IMPLEMENTING THEM IN INSTRUCTIONAL SUPPORT MEASURES

Do the process characteristics that we consider to be relevant for successful cooperation actually lead to good collaborative outcomes? As we can see from the correlations between the ratings of the nine dimensions and the scores uninstructed dyads gained for their joint solution, this seems indeed to be the case. Another way of answering this question has already been taken in the experimental instructional study (Rummel & Spada, 2005). Dyads were taught to collaborate in a way which was characterized by many features resembling the dimensions of our rating system.

Method

One of the two instructed conditions in the experiment (Rummel & Spada, 2005; see Table 1) involved learning from a worked-out collaboration example (model condition). During the learning phase of the experiment, participants in this condition watched a multimedia-presentation on their computer screen. They listened to recorded scenes of the collaborative problem-solving between a psychology student and a medical student on a first psychiatric case. Animated slide-clips allowed participants to observe the development of the joint solution in the text editors of the model collaborators. An exemplary collaboration was shown in the model presentation, with many features corresponding to the characteristic dimensions of a good collaboration outlined above. Instructional explanations (such as “In the following scene you will hear how the two collaborators ask each other questions about the case. They make use of each others knowledge to clarify information given to them about the patient in the case description before they turn to the diagnosis”) as well as prompts for self-explanations were included in order to support a deeper processing of the worked-out collaboration example. The second instructional condition involved learning from scripted collaboration (script condition). Here, dyads were provided with a detailed script prescribing specific phases for their interaction. The script followed the same exemplary collaboration as presented in the model condition. The two non-instructed conditions served as controls.

Results

Results showed that both instructed conditions, model and script, outperformed the non-instructed conditions (Rummel & Spada, 2005). This implies that the dimensions represent relevant aspects of good, successful collaboration.

DISCUSSION

In this paper, we presented a new method for assessing the quality of collaborative process in computer-supported problem-solving and learning settings. Nine dimensions central to collaboration were extracted combining a data-driven analysis of collaborative process with theoretical considerations. The first two dimensions, *sustaining mutual understanding* and *coordinating communication* refer to basic communication processes which form a prerequisite for successful collaboration. The third and fourth dimension, *information pooling* and *reaching consensus*, are

relevant for the construction and maintenance of a shared understanding. *Task division*, *time management* and *technical coordination* are three dimensions reflecting the coordination of collaborative activities. Finally, the motivational aspect is covered by the two dimensions *shared task alignment* and *sustaining commitment*. The rating system we developed implementing these nine dimensions enables the user to assess the quality of the collaborative process on a relatively global level, resulting in quantitative ratings that can be subjected to statistical analyses. We have shown that the inter-rater reliability of the nine dimensions is satisfactory. Rating instructions of the less satisfying dimensions are currently under revision. Some rather high inter-correlations between the dimensions indicate that maybe a leaner instrument with fewer dimensions would be sufficient.

Correlations with the quality of the joint solution are moderate to high. These process-outcome correlations, however, are not only contingent on the reliability of our process ratings but also on the reliability with which the joint outcome was assessed. Since the participants of our study had to solve complex tasks, assessing the quality of the solution was not trivial. Process and outcome measures might show an even stronger relation when applied to problems whose solution quality is easier to evaluate. Taking together all of the results, these are promising findings. Yet, a larger sample is needed to further improve the method and replicate the results.

In our approach to assessing the quality of collaborative process, we wanted to combine the benefits of qualitative, data-driven and quantitative, concept-driven approaches. We did this by first qualitatively identifying relevant dimensions of collaborative process and then implementing them in a rating system that yields quantitative ratings. Our rating system differs from quantitative methods of coding and counting (e.g. Bruhn et al., 1997, Rummel & Spada, in press), in that it affords a more holistic assessment of the quality of collaborative processes. Compared to very fine-grained discourse coding schemes, like the DISCOUNT scheme (Pilkington, 1999), which can only be applied to transcribed dialog data, the time expenditure necessary for applying our method is considerably lower. Videotaped collaboration can be reviewed without transcribing dialog. For one hour of videotaped collaboration, about two hours of time should be calculated for reviewing and rating. Raters should be trained in advance in order to be sensitive to relevant characteristics of collaborative process.

We propose that the rating system should be applicable in most areas of CSCL research that involve collaborative problem-solving and learning on the basis of complementary expertise. Of course, the rating instructions for the “technical coordination” will have to be adjusted to the specific technical setting one wishes to analyze. Although the rating system was developed and evaluated for collaboration in dyads, we think it might also be applicable to groups of three or four collaborators.

REFERENCES

- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences*, 9 (4), 403-436.
- Bruhn, J., Gräsel, C., Fischer, F., & Mandl, H. (1997). Kategoriensystem zur Erfassung der Kokonstruktion von Wissen im Diskurs [A category system for assessing the co-construction of knowledge during discourse]. München, DFG-Projekt Ma 978/6-2: Kooperatives problemorientiertes Lernen.
- Clark, H. H. & Brennan, S. E. (1991). Grounding in Communication. In L. B. Resnik, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 17-149). Washington, DC: American Psychological Association.
- Collazos, C. A., Guerrero, L. A., Pino, J. A., & Ochoa, S. F. (2004). A method for evaluating computer-supported collaborative learning processes. *International Journal of Computer Applications in Technology*, 19 (3/4), 151-161.
- Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1995). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards a interdisciplinary learning science* (pp. 189-211). Oxford: Elsevier/Pergamon
- Fischer, F. & Mandl, H. (2003). Being there or being where? Videoconferencing and cooperative learning. In H. v. Oostendorp (Ed.), *Cognition in a digital world* (pp. 205-223). Hillsdale, NJ: Erlbaum.
- Guribye, F., Andreassen, E.F., & Wasson, B. (2003). The organisation of interaction in distributed collaborative learning. In B. Wasson, S. Ludvigsen & U. Hoppe (Eds.), *Designing for change* (pp. 333-342). Dordrecht: Kluwer.
- Häkkinen, P., Järvelä, S., & Mäkitalo, K. (2003). Sharing perspectives in virtual interaction: Review of methods of analysis. In B. Wasson, S. Ludvigsen & U. Hoppe (Eds.), *Designing for change* (pp. 333-342). Dordrecht: Kluwer.

- Hermann, F., Rummel, N., & Spada, H. (2001). Solving the case together The challenge of net-based interdisciplinary collaboration. In P. Dillenbourg, A. Eurelings, & K. Hakkarinen (Eds.), *Proceedings of the First European Conference on Computer-Supported Collaborative Learning (E-CSCL)* (pp. 293-300). Maastricht: McLuhan Institute.
- Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, 121(1), 43-64
- Kneser, C. & Ploetzner, R. (2001). Collaboration on the basis of complementary domain knowledge: observed dialogue structures and their relation to learning success. *Learning and Instruction*, 11, 53-83.
- Koschmann, T., Zemel, A., Conlee-Stevens, M., Young, N., Robbs, J., & Barnhart, A. (in press). How do people learn?. In R. Bromme, F. W. Hesse & H. Spada (Eds.), *Barriers and biases in computer-mediated knowledge communication and how they may be overcome*. Dordrecht, NL: Kluwer Academic Publishers.
- Malone, T. & Crowston, K. (1994). The interdisciplinary study of coordination. *ACM Computing Surveys*, 26 (1), 87-119.
- Mayring, P. (2003). Qualitative Inhaltsanalyse. Grundlagen und Techniken [Qualitative content analysis. Foundations and techniques]. Weinheim: Beltz.
- Nurmela, K., Palonen, T., Lehtinen, E., & Hakkarinen, K. (2003). Developing tools for analyzing CSCL process. In B. Wasson, S. Ludvigsen & U. Hoppe (Eds.), *Designing for change* (pp. 333-342). Dordrecht: Kluwer.
- O'Connaill, B. & Whittaker, S. (1997). Characterizing, predicting, and measuring video-mediated communication: A conversational approach. In K. E. Finn, A. J. Sellen & S. B. Wilbur (Eds.), *Video-mediated communication* (pp. 107-132). Mahwah, NJ: Erlbaum.
- Pilkington, R. M. (1999). Analyzing educational discourse: The DISCOUNT scheme. Technical Report No. 99/2. CBLU, University of Leeds, Leeds, UK. Retrieved November 14, 2004, from <http://www.education.bham.ac.uk/aboutus/profiles/curped/pilkington/docs/DISCoun99.htm>
- Pilkington, R. & Parker-Jones, C. (1996). Interacting with computer-based simulation: The role of dialogue. *Computers and Education*, 27 (1), 1-14.
- Roschelle, J. & Teasley, D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69-97). Berlin: Springer.
- Rummel, N. & Spada, H (2004). Cracking the nut – but which nutcracker to use? Diversity in approaches to analyzing collaborative processes in technology-supported settings. *Proceedings of the sixth international conference of the learning sciences, UCLA, Santa Monica* (pp. 23-26).
- Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem-solving in computer-mediated settings. *Journal of the Learning Sciences*, 14(2), 201-241.
- Rummel, N. & Spada, H. (in press). Sustainable support for computer-mediated collaboration: How to achieve it and how to assess it. In R. Bromme, F. W. Hesse & H. Spada (Eds.), *Barriers and biases in computer-mediated knowledge communication and how they may be overcome*. Dordrecht, NL: Kluwer Academic Publishers.
- Sosa y Fink, S. (2003). Merkmale gelungener Kooperation. Eine qualitative Analyse netzgestützter Zusammenarbeit [Characteristics of successful cooperation. A qualitative analysis of net-based collaboration]. Unpublished diploma thesis, Albert-Ludwigs-Universität, Freiburg.
- Stasser, G. & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, 48, 1467-1478.
- Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Education Research*, 13, 21-39
- Whittaker, S. & O'Conaill, B. (1997). The role of vision in face-to-face and mediated communication. In K. E. Finn, A. J. Sellen & S. B. Wilbur (Eds.), *Video-mediated communication* (pp. 23-50). Mahwah, NJ: Erlbaum.