Rating the quality of collaboration during networked problem solving activities

Eleni Voyiatzaki¹, Anne Meier², Georgios Kahrimanis¹, Nikol Rummel², Hans Spada², Nikolaos Avouris¹,

¹University of Patras, Greece, {evoyiatz, kahrimanis}@ece.upatras.gr, avouris@upatras.gr
²University of Freiburg, Germany, {ameier, rummel, spada}@psychologie.uni-freiburg.de

Abstract

This paper describes the adaptation and generalization of a rating scheme for judging the quality of collaboration during network-based problem solving activities with Synergo, a network-based synchronous collaborative drawing environment including a shared whiteboard and a chat tool. The goal of this study was to develop a tool for assessing the strength and weaknesses in students’ collaboration in large samples. A second goal was to show that the rating scheme’s dimensions can capture the main aspects of effective collaboration across diverse settings. Adaptation of the rating scheme involved adjusting the dimensions’ definitions to the affordances of the Synergo tool and characteristics of the given algorithm task, and testing it in a sample of six dyads selected from a broad range of studies. The final, adapted rating scheme includes seven dimensions: (i) collaboration flow, (ii) sustaining mutual understanding, (iii) exchanging knowledge and giving explanations, (iv) argumentation, (v) structuring the problem solving process and time management, (vi) cooperative orientation, and (vii) individual task orientation. For each dimension, we provide a detailed definition and case studies that illustrate how the rating scheme’s dimensions can be used to describe and evaluate differentially the quality of students’ collaboration. We outline perspectives for future research applying the newly developed rating scheme.

Keywords

Quality of collaboration, synchronous collaborative learning, algorithm teaching, rating scheme

Introduction

This paper describes the adaptation and generalization of a rating scheme for judging the quality of collaboration during network-based problem solving activities. The work described is part of an ongoing interdisciplinary collaboration between the two research teams of the University of Patras, Greece and the University of Freiburg, Germany, and aims at the exchange of analysis tools and data sets in order to broaden the scope of analysis methods and tools available for Computer-Supported Collaborative Learning (CSCL) support. In the study reported in this paper, a rating scheme, which had been developed by the Freiburg team for assessing collaboration quality on several dimensions (Meier, Spada, & Rummel, 2007), was adapted to fit data gathered by the Patras team. These data were collected during several studies that have been conducted in the frame of the regular first-year curriculum in the computer engineering department, with over 200 participating students over the last five years. These studies had served, among other purposes, to improve both the design of the networked learning tools in use, and the teaching curriculum itself. To be able to take into account the quality of the collaboration between groups of students as an additional dependent variable, in future studies as well as for re-analysis of the available data sets, a tool was needed with which the logfiles of students’ collaborative activities can be analyzed on a large scale. The rating scheme developed and applied by the Freiburg team had been designed for exactly this purpose. However, it needed to be adapted to the new CSCL context: Originally, the rating scheme had been developed for analyzing collaboration activities in the domain of interdisciplinary problem solving between medical students and students of psychology who communicated using a video-conferencing system (Rummel & Spada, 2005). This original setting differed from the one used in Patras.
in the following aspects: in the University of Patras setting, students interacted not through a videoconferencing system, but through Synergo (Avouris, Margaritis, Komis, 2004), a tool for networked learning, which includes a shared whiteboard and a chat tool. The domain, algorithm building in computer science, was also quite different from the one studied at Freiburg (medical decision making). Furthermore, while the Freiburg study used advanced students from two different knowledge backgrounds (medicine and psychology), students in the Patras case were first-year computer engineering students with similar knowledge backgrounds. The goal of the adaptation study was not only to come up with a useful tool for analyzing data, but also to show that the rating scheme’s dimensions, and thus its underlying theoretical model, are capable of capturing the main aspects of collaboration quality across different CSCL settings. In other words, we see this study as an important test of the rating scheme’s practical applicability and theoretical validity.

Adapting the rating scheme for a new setting

Synergo: A networked learning setting

The collaborative activities studied involved dyads of first year computer engineering students. Their joint task was to build a diagrammatic representation of an algorithm that was provided in textual form. The collaborative activity took place during a typical laboratory class and lasted approximately 45 minutes. The students interacted through Synergo (Avouris, et al., 2004), a network based synchronous collaborative drawing tool. Synergo includes a shared whiteboard, in which students can build the diagrammatic representation (i.e. a flowchart) of an algorithm, and a chat tool through which they can communicate during problem solving. Both tools are displayed in the same window (cf. Fig 1). In addition, Synergo generates a logfile of students’ activity, stored together with the final drawing. The “Synergo Analysis Tool” allows researchers to play back these logfiles in a video-like format. In that way, evaluators can review the whole activity, and focus in different episodes that have occurred and require more detailed inspection. A screenshot from a Synergo Analysis Tool in playback mode is shown in Figure 1.

Figure 1: Screenshot of the Synergo analysis tool in playback mode, showing the contents of the shared whiteboard and chat produced by two students

Adapting the original rating scheme to this new setting required a good understanding of the communicative affordances offered by Synergo. Most importantly, students were using two spaces of communication and interaction in combination: the shared whiteboard and the chat. They therefore used two “languages” of interaction. In the shared whiteboard, they used indirect communication, characterised
as “feedthrough” in Computer Supported Collaborative Work literature (Dix, A., J. Finlay, G. Abowd and R. Beale, 2004). This may involve actions like demonstrations, manifestations, and visual evidence of issues negotiated in chat (Gutwin and Greenberg, 2002). The conventions in the notation of flowcharts are of crucial importance to this channel of communication and the possibilities of indirect communication that it affords. Language conveyed through chat has also different characteristics and properties than language in face-to-face or even videoconferencing systems, as visual and auditory cues that offer a secondary mode of communication in face-to-face communication and teleconferencing cannot be conveyed. On the other hand, messages are persistent; i.e. they can be reviewed anytime after the posting. This allows collaborators to use verbal deixis that can refer back to earlier messages in the dialogue (Garcia and Jacobs, 1999).

The process of adaptation

Adaptation involved several rounds of adjusting the original dimensions’ definitions to the affordances of Synergo and characteristics of the task. We applied the given rating dimensions to a sample of well understood collaboration incidents of the algorithm task, identifying aspects of collaboration that were not covered by these dimensions. The sample consisted of from six student dyads who had participated in prior studies conducted by the Patras team (Voyiatzaki et al., 2008). We also combined in a top-down process, considerations on how the differences in task, sample, and CSCL setting would change the meaning of “successful” collaboration in the Patras sample, with the bottom-up approach that involved identification of examples of “best practice” in the sample data. We took into account the new communicative affordances, like persistency of dialogue messages. For example, turn-taking, as an important aspect of successful communication, had a completely different manifestation in the new context of chat communication and whiteboard activity, to that of videoconferencing and shared document. For instance, the dimension assessing how well students coordinated their communication on a moment-to-moment basis, manifested by their turn-taking in the original rating scheme, was now redefined to focus on the “flow” and coherence in the interplay of actions in the shared whiteboard and chat messages. In a similar process, we reformulated the definitions of all dimensions of the original rating scheme. The first version of the adapted rating scheme was then applied to several case studies from the Patras data set. Four dyads from four different studies conducted by the Patras team were selected in order to cover collaboration by a broad rage of participants and in a broad range of settings. Two raters from the Patras team and one rater from the Freiburg team analyzed the data with the help of the adapted rating scheme, answering two main questions: 1) Are the dimensions able to cover the main aspects of collaboration quality in the given scenario? 2) How well do the definitions fit the spectrum of collaboration observed in the sample dyads? Following this analysis, further changes were made to the rating scheme’s dimensions. For example, students in the original scenario held complementary knowledge and therefore frequently exchanged their expertise-specific information during collaboration. This activity had been part of the original “information pooling” dimension. Students in the Patras scenario, on the other hand, shared the same knowledge background. Accordingly, the four case studies showed that plain information pooling was not very frequent in these dyads. Therefore, the focus of this dimension was shifted towards the explanations students gave for their actions and as a response to questions from their partner. The adapted dimension, accordingly, was named “knowledge exchange and giving explanations”, rather than “information pooling”. Examples from the four case studies were selected to illustrate the definitions of the dimensions in the adapted rating scheme. The rating scheme was then applied to the original Greek logfiles of two more dyads by two raters from the Patras team, confirming that the dimensions were now able to capture the most important aspects of students’ collaboration.

The final rating scheme

The final, adapted rating scheme includes seven dimensions. Each dimension defines a standard for a specific aspect of good collaboration: First of all, collaborators have to communicate successfully using chat as well as actions in the shared whiteboard. To achieve successful communication, partners have to maintain 1) collaboration flow; i.e. engage in a coherent exchange of information through and maintain a joint focus. Further, they need to 2) sustain mutual understanding, i.e. work towards “common ground” (Clark & Brennan, 1991). Regarding work on the actual algorithm task, a very important dimension is that of 3) exchanging knowledge and giving explanations (i.e. self- and other-directed explanations).
addition, students have to engage in 4) **argumentation** in order to ensure a good solution and to foster their own learning progress. This dimension refers to all activities involved in double-checking the problem-solving process, avoiding mistakes, and maintaining a critical discussion. To ensure a timely and orderly solution to the given problem, students also have to coordinate their collaboration well by 5) **structuring the problem solving process and time management**. Finally, students have to maintain a 6) **cooperative orientation** (e.g., constructive handling of disagreements), and a high level of 7) **individual task orientation** throughout their collaboration. More extensive descriptions of each dimension, along with illustrative examples from several case studies are provided in the next section.

A **rating handbook** states the scope and purpose of each dimension, gives an operational definition in a short paragraph, and provides the rater with illustrative examples. The rating handbook is intended to be used in the framework of a **rater training** that involves further illustration of the dimensions, and a more precise anchoring of the scales with the help of (playback) videos taken from the actual sample that is about to be analyzed.

### The dimensions of the rating scheme of the quality of collaboration

To illustrate the way in which the rating scheme works with data from the new sample, we selected dyads from four different studies. The first three student dyads (Dyads 1, to 3) took part in three studies conducted at the University of Patras. The forth one (Dyad 4) is from a more recent study, conducted at the University of Macedonia in Thessaloniki, Greece. All studies were implemented in the regular curriculum of first-year computer science course, using Synergo as a networked learning environment.

Students in all cases worked on the same task, i.e. building a diagrammatic representation of the algorithm for binary search in a set of numbers. All students had prior knowledge of the binary search algorithm: in the first study they had been given the opportunity to try the task individually before working on it collaboratively, and in the other cases they had already studied binary search as part of their curriculum. However, students in the four studies differed strongly in the amount of instruction they had received regarding collaboration: Students in the first study did not receive any specific instruction at all. They were only informed that collaboration would be important for their joint task, and that they would receive credit if their activity logfiles showed that they had collaborated well. Students in the second study received general, written instructions how to achieve successful collaboration (e.g. by creating the algorithm cooperatively, by asking and answering questions, by giving explanation and examples, and by encouraging one’s partner). Students in the third study also received written instructions, but additionally engaged in a collaborative practice task, on a simpler algorithm, before working on the binary search task one week later. Finally, students in the forth study received written instruction and, in addition, were able to observe an exemplary case of collaboration prior to their own collaboration on the task: Two teachers built an algorithm using Synergo prior to the class, which was played back to the students and was commented by one teacher who pointed out examples of good or bad collaboration practices.

All four dyads were above the class average in their solution to the task, but not excellent. We use data from these four studies here, not to demonstrate the effects of instruction (which would require a much bigger sample), but to illustrate the application of the rating schemes dimension to collaboration data obtained under a broad range of conditions. Exact values for the ratings are not reported, though, because the specific anchoring of the rating scale should be dependent on the variability of behaviours in the sample and setting in which the tool is used, which cannot be inferred from the small sample at hand. An evaluation of the rating scheme’s reliability and validity in a bigger sample will be conducted as part of an experiment testing the effectiveness of adaptive feedback for improving the quality of students’ collaboration (Meier et al., submitted). In the following, we will provide a brief definition of each of the seven dimensions of the rating scheme, followed by illustrative examples from one or more of the case studies. We will refer to the students in each dyad as either “Red” or “Blue” (because their chat messages appeared in either red or blue in Synergo) followed by the number of the dyad they belong to (1, to 4).

**Dimension 1: Collaboration flow**

The most important criterion for this dimension is that there is a seamless “flow” of interlocking actions and chat messages while the students are collaborating on their joint task. This implies that students react to actions in the whiteboard, and to questions and proposals in the chat, by typing chat messages or by
taking actions in the whiteboard themselves. Students need to maintain mutual awareness, i.e. each student keeps track of what the partner is currently working on.

Collaboration flow was not very smooth in Dyad 1, for example, where students often worked in the whiteboard in parallel, creating and deleting objects without referring to their actions in the chat. On the other hand, actions and chat entries in Dyad 3, showed a high level of coherence most of the time (e.g. BLUE3’s actions refer directly to RED3’s orders), and it was easy for the observer to follow the “flow” of the dialog. Only in the last minutes the chat got partially disorganized, probably due to time pressure. The students in Dyad 4 also showed good collaboration flow: Their actions and chat entries show a high level of coherence, as in the following example:

RED4 creates an important portion of the algorithm in the whiteboard and Blue signals
00:11:46 BLUE4  AGREE'
(and then they work together on the same portion for a while)
00:12:25 RED4  yes go on (RED4 prompts BLUE4 )
00:12:26 BLUE4  I THINK WE PROCEED WELL (BLUE4 continuous by inserting another part )

Dimension 2: Sustaining mutual understanding

The standards for this dimension demand that partners work towards a shared basis of understanding, i.e. “common ground” (Clark & Brennan, 1991). Thus, this dimension refers to the basic processes of communication on a moment-to-moment basis. Students should take care to communicate and act in a way that is understandable for their partner. They should ask for and give feedback, for example by using backchannels (like “okay”), posing questions, or signalling agreement or disagreement in the chat. They can also demonstrate their understanding by replying with a relevant next message (Clark & Brennan, 1991) or taking a relevant next action. If misunderstandings occur, students should ask for and give clarifications (e.g. explanations for technical terms and variables).

Students In all four dyads seemed to understand the need to sustain mutual understanding, but they dealt with this need in different means. Partners of Dyads 2 and 4, for example, did not hesitate to frequently ask questions and give answers, even if they were not correct, while partners of Dyad 1 were responding with some delay and answering more elliptically: They seemed to be more patient and wait until they were able to understand their partner’s intentions either from their messages or their actions. Nevertheless, they managed to sustain mutual understanding well enough to proceed without misunderstandings. An example from Dyad 3 shows an episode where students work actively on establishing common ground while they monitor the part of the solution they have produced. At the end of this episode, mutual understanding is established:

00:22:37 BLUE3:  after the start what does that mean?
00:22:48 BLUE3:  found...false
00:22:54 RED3:  in the start-box ?
00:23:25 BLUE3:  no the box below that
00:23:37 RED3:  in the diamond-shape?
00:24:02 BLUE3:  no
00:24:22 BLUE3:  the box that contains FOUND
00:24:30 RED3:  oh yes

Dimension 3: Knowledge exchange and giving explanations

The standards for this dimension demand that collaborators make use of their knowledge resources in the process of solving their joint task. Students should demand and give elaborated explanations, using each other as a resource for obtaining the information they need in order to solve their joint task. In proceeding with their task, students should explain why they are doing what they are doing.

Students in all four dyads exchanged only little knowledge, even if this would have helped them with their task. In cases where information was exchanged, the discussion often did not include elaborated explanations. Some episodes from the logfiles, however, show positive examples of knowledge exchange and giving explanations, like the following one, from Dyad3:

BLUE3 answers a question posed earlier by RED3 by giving an elaborated explanation. This is also a good example of sustaining mutual understanding by using short “backchannels” (like “I got it”) and demonstrating understanding by relevant next messages (“L is the number…” ) in the chat.
00:25:12 RED3: we want a variable that when the object of search is found to take a value and the algorithm to end
00:25:31 BLUE3: aha m
00:25:43 RED3: in order to go to the loop it has to be false otherwise it is out of the question
00:26:06 BLUE3: a ok I got it
00:26:17 RED3: so when it is found true...indeed we do it arbitrary because it comfortable for us
00:27:01 BLUE3: L is the number we give in order to find it from the table
00:27:20 RED3: l and r are like your min and max
00:28:15 BLUE3: ok

Dimension 4: Argumentation
This dimension assesses if students double-check their solution in order to avoid mistakes, and if they maintain a critical discussion during their problem solving. For example, students should ask for and provide justifications for solution steps (e.g. parts of the algorithm) they are proposing. If students disagree, they should work to dissolve their disagreement in a fact-based discussion, searching for good arguments for and against the options under discussion, until a sustainable consensus is established. Students should also check whether their solution fulfills certain solution criteria (e.g. whether their algorithm will work as it is supposed to do), if important parts are missing, or if there are any mistakes.

Argumentation quality varied among the four dyads, with sparse and shallow arguments uttered in Dyad 2 and 3, and deeper negotiation and “common thinking” evident in the other two dyads. It is interesting to mention that deep negotiation took place in the discussion of Dyad 1 despite difficulties with communication flow and knowledge exchange/giving explanations. The following episode from Dyad 1 is indicative of the effort to proceed with the solution through critical discussion, even if the arguments are not always correct.

00:18:33 BLUE1 wait a sec
00:18:35 BLUE1 how is it going to be done
00:18:43 RED1 wait a sec (Red is adding an object and deleting it without any comments)
00:19:05 BLUE1 if we divide b we won’t know we’ll lose the position (Red creates free space for new insertions)
00:20:20 RED1 b becomes half and we search again (one of the most critical points for the solution)
00:20:59 BLUE1 we don’t want b to be lost to show the position (this complements Red’s argument)
00:20:59 RED1 it’s not necessary to put a new variable
00:21:25 RED1 yes but if we put c we have to rewrite the algorithm (c is a new variable)
00:21:29 RED1 for c
00:21:37 RED1 while now we have it ready for b (The objection of Red at 20:59 is right even if this explanation, here, is wrong)

Dimension 5: Structuring the problem solving process and time management
This dimension assesses the quality of coordination: Collaborators should make sure they finish all parts of their task in time, identify important subtasks and allocate sufficient resources for them, and they should monitor the remaining time. As a consequence, they should be able to finish all parts of their task in time. Students in all four dyads struggled with the task of structuring their problem solving process.

This was a weak point, in all dyads, and at the same time it can be a clear objective and taught relatively easily. Students in Dyad 4, for example, never monitored the time that remained for their joint task, and consequentially ran out of time (the limit was 45 minutes) in the end, with no time left to finish the last part of their assignment, i.e. to document their algorithm through “sticky” notes.

00:41:35 BLUE4: IS DOCUMENTATION NEEDED?
00:41:36 RED4: no it's good
00:41:44 RED4: Sticky notes??
00:42:29 BLUE4: WRITE IT
00:43:38 BLUE4: what else is needed
00:44:12 RED4: we have not enough time to document properly so should we deliver?
00:44:48 BLUE4: but we did not finish!
Dimension 6: Cooperative orientation

This dimension refers to the social aspect of collaboration, and to the underlying orientation collaborators have towards their task and their collaboration partner. There may be some competition regarding the task (e.g. argumentation in the case of disagreement), but not on the social plane (e.g. competing for grades or status with the teacher). Students take on symmetrical or complementary roles (e.g. tutor and tutee), and do not compete for power during collaboration. As a consequence, both have the chance to be equally active during collaboration, even though they may engage in different kinds of actions and roles. Collaborators respect each other as persons and interact in a polite and friendly way. If conflicts arise, they are treated as problems to be solved collaboratively, e.g. by fact-based discussion. The following example from Dyad 4 shows that collaborative orientation can sometimes be expressed explicitly:

00:00:41 RED4: hi
00:00:57 BLUE4: hi who are you
...(students type their names in the chat)
00:01:53 RED4: ok, let's have a good collaboration

In Dyads 1 and 4, both partners contributed actively to the joint solution, collaborating rather well. In Dyads 2 and 3, on the other hand, the interaction was asymmetric, with one student being excluded from the problem-solving process by a more dominant partner. For example, BLUE2 clearly dominated the problem-solving process over long parts, doing nearly all the work on the algorithm himself, usually without even explaining to RED2 what he was doing.

Dimension 7: Individual task orientation

This dimension refers to the effort exerted by each individual student during collaboration. It is the only dimension that is assessed at the level of the individual students rather than at the group level. Ideally, the student should be focussed on the task at hand, engage actively in the collaboration, and also motivate his or her partner to engage actively. In Dyad 3, for example, only RED3 was actually focussed on and actively engaged in the task, while BLUE3 withdrew more and more, showing only little task orientation. Rating of individual task orientation revealed only small differences among the partners in the other three dyads, as all students appeared to be very motivated and all engaged actively. Sometimes, they even tried to motivate each other: In the following example from Dyad 1, BLUE1 expresses his confidence in his own and his partner’s competence to solve the problem, and focuses his own and his partner’s attention back on the task.

00:52:30 RED1: my friend, the diagram 's a mess
00:52:42 BLUE1: wait
00:52:47 BLUE1: calm down
00:52:49 BLUE1: it is going to work
00:53:23 BLUE1: we do not make b read
00:53:36 BLUE1: what’s b?

Discussion

The adaptation study described in this paper yielded, first of all, a tool for assessing collaboration quality in networked learning activities with Synergo. The case studies discussed illustrate how the rating scheme’s dimensions can be used to describe and evaluate differentially the quality of students’ collaboration. This successful adaptation shows that the theoretical model underlying the rating scheme’s dimensions, described in more detail in Meier et al. (2007) is valid across very different learning environments, tasks and samples settings. The case studies also show that even relatively successful student dyads still had problems to manage their collaboration in a way that would lead to optimal results regarding both their joint task solution and their individual learning outcomes. For example, evaluating the case study dyads on the dimension of “knowledge exchange/giving explanations” showed that students typically did not exchange much information, let alone give elaborated explanations. Thus, an important goal in our future work will be to focus on such specific weaknesses in students’ collaboration, and develop means of training students to collaborate better. In an experimental study we explore the possibility of giving adaptive feedback to students based on an offline assessment of their collaboration.
on a joint task with the help of the adapted rating scheme (Meier et al., submitted). It is expected that feedback that is adapted to the multidimensional assessment of students’ collaboration quality with the help of the rating scheme, will help students to focus on those aspects of their collaboration that need the most attention, and thus will be more effective than giving to them general information about what constitutes successful collaboration. Tests of the rating scheme’s reliability and validity in a larger sample, which are beyond the scope of the study presented in this paper, will also be conducted as part of data analysis in this experiment. Further work is planned in order to facilitate the rating process, and make it easily and effectively applicable by trained teachers, for regular use in evaluating and facilitating networked learning in their classes. Collaboration logfiles from various studies conducted at the University of Patras can be used as a rich data basis to refine the rating process and to create training material. The intention is to support teachers in assessing specific strengths and weaknesses in their students’ collaboration, and to help them provide helpful feedback to their students, teaching them how to collaborate successfully.

References


Acknowledgment

This research project was partly funded by the Kaleidoscope Network of Excellence project Cavicola (Computer-based Analysis and Visualization of Collaborative Learning Activities).