Validating Empirically a Rating Approach for Quantifying the Quality of Collaboration

Georgios Kahrimanis, Irene-Angelica Chounta, and Nikolaos Avouris

Human-Computer Interaction Group, University of Patras, Rio-Patras, Greece
{kahrimanis,houren}@ece.upatras.gr, avouris@upatras.gr

Abstract. Interdisciplinarity in the Computer Supported Collaborative Learning (CSCL) research field involves the application of several methodological approaches towards analysis that range from deep-level qualitative analyses of small interaction-rich episodes of collaboration, to quantitative measures of suitably categorized events of interaction used as indicators of the success of collaboration in some of its facets. This article adopts an alternative approach to CSCL analysis that aims at taking advantage of some desired properties of each of these diverse methodological trends, involving the use of a rating scheme for the assessment of collaboration quality. After defining a set of dimensions that cover the most important aspects of collaboration, it employs appropriately trained human raters basing their assessments on substantial aspects of collaboration that are not easily formalisable. The activities studied here regard 228 collaborating dyads, working synchronously on a problem-solving task. Based on this large dataset, relations between dimensions of collaboration quality are unraveled on empirical grounds, by elaborating ratings statistically using a multidimensional scaling technique.

1 Introduction

Computer Supported Collaborative Learning (CSCL) constitutes one of the most extensively developed paradigms of research and practice in intelligent networking and collaborative systems technology. Under specific conditions, collaborative interactions can trigger collaborative knowledge building (Scardamalia and Bereiter 1996) that is beneficial for learners participating in collaborative processes.

Apart from the study of the conditions that can lead to fruitful CSCL processes, and the “learning gains” that students may obtain, analysis of collaborative interactions per se constitutes one of the core aspects of the study of CSCL (Dillenbourg et al. 1995). Interdisciplinarity in the research field involves the application of several methodological approaches towards analysis of CSCL inspired, adopted, or developed based on diverse research disciplines. Most commonly used analysis studies range from deep-level qualitative analyses of small interaction-rich episodes of collaboration, to quantitative measures of suitably categorized events of interaction that are used as indicators of the success of collaboration in some of its facets (Stahl et al. 1996; Kahrimanis et al. 2011).

Whereas the latter approach to CSCL analysis offers possibilities for practical facilities such as quick or even automated assessments of collaboration as in (Avouris

T. Daradoumis et al. (Eds.): Intelligent Adaptation & Personalization Techniques, SCI 408, pp. 295–310.
springerlink.com

et al. 2004; Strijbos et al. 2006), it is claimed that it is often based on measuring "surface" aspectzs of collaboration (Stahl et al. 1996). On the other hand, cases that belong to the former approach, such as (Roschelle 1992; Stahl 2006), may be rightly considered the most suitable for in-depth CSCL analysis, they are, however, arduous and time-consuming, since they demand much effort for analysing relatively small episodes of collaboration. Moreover, it is difficult that they scale-up to analysis of extended datasets when dealing with large-scale studies.

This article adopts an alternative approach to CSCL analysis that aims at taking advantage of some desired properties of each of these diverse methodological trends. It involves the use of a rating scheme for the assessment of collaboration quality (Kelringer and Lee, 2000). After defining a set of dimensions that cover the most important aspects of collaboration, it employs appropriately trained human agents to assign ratings of collaboration quality to each dimension, basing their assessments on substantial aspects of collaboration which are not easily formalisable. Still, the outcome of the evaluation process is provided in quantitative form, suitable for statistical manipulation. This way, the results obtained can serve as a point of reference for intelligent techniques that are based on automatable formalisations, providing new information that takes account of deeper aspects of collaboration than most top-down approaches for evaluation, personalisation and adaptive feedback.

The activities studied here regard 228 collaborating dyads, working synchronously on a computer science problem-solving task with the use of the Synergo tool (Avouris et al. 2004). Based on this large dataset, relations between dimensions of collaboration quality are determined on empirical grounds, based on the ratings of collaboration quality applied for each dimension in each collaborative session. The technique selected, which allows the systematic view of these associations in statistical means, is multidimensional scaling (Kruskal and Wish 1978; De Leeuw and Heiser, 1982; Schiffman et al. 1981; Davison 1983; Young and Hamer 1994; Borg and Groenen 1997; Cox and Cox 2001), using dimensions of collaboration quality as the unit of analysis. That way, associations between core aspects of collaboration are represented in a two-dimensional space, with distances between dimensions denoting their dissimilarities. The results obtained are in accordance with the initial design of the rating scheme used, and further particularize the relations between its dimensions. General conclusions and further steps made possible by current findings are discussed in the last section of this article.

2 Collaborative Setting

Collaborative activities studied in this article involved about 350 computer science students at the department of Electrical and Computer Engineering of the University of Patras, Greece, engaged in jointly building the diagrammatic representation of an algorithm as an assignment of a two-hour laboratory session that was part of the first-year of studies course "Introduction to Computers". These activities took place in a single laboratory room, equipped with one computer per student. Students interacted through Synergo (Avouris et al. 2004), communicating via an integrated chat tool, and jointly designing a flow-chart representation of an algorithm in Synergo's shared workspace. A capture of a user's screen while collaborating with Synergo is shown in Fig. 1. Synergo provides libraries of objects supporting the notation of several diagrammatic models. Collaborative sessions lasted from 45 to 75 minutes and students worked in dyads, which were selected randomly. They were free to use their own resources such as textbooks or the web and were permitted to ask questions to a teacher, who restricted her feedback to technical or other minor aspects. In order to motivate students to work on the exercises collaboratively, they were informed that the grade they would get for the particular lab session would be determined by both the quality of their collaboration and the completeness and correctness of their joint solution. Dyads were arranged in space in a way that it was impossible for the students to use any other means of communication apart from these provided by Synergo.

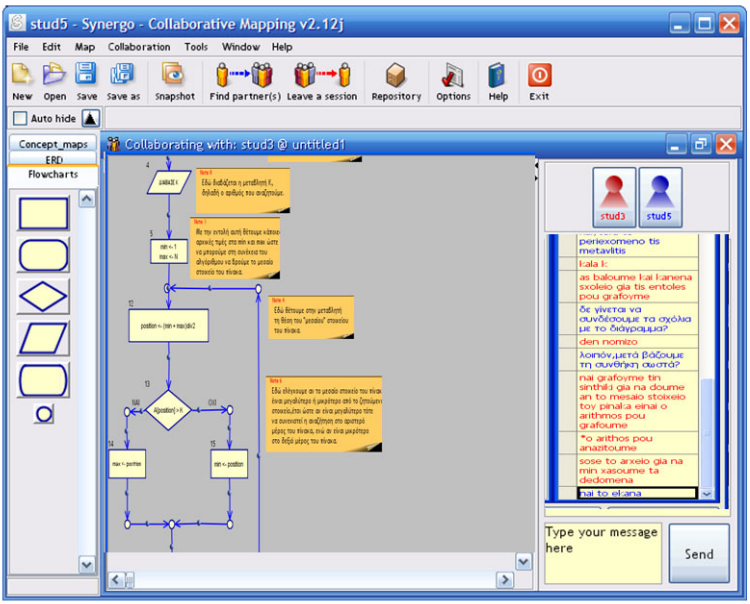

Fig. 1 Synergo in use: a capture of a user's screen collaborating on an algorithm problem

The problem domain of the task was basic algorithms in computer science. Students were asked to solve elementary algorithm exercises that are widely used for training basic algorithmic skills. For example, a basic piece of knowledge in algorithms is the concept of the variable (Samurcay 1989), which should be discriminated from the understanding of the variable that students may have from mathematics. Another major learning object was also the proper handling of algorithmic structures, such as the loop structure (Soloway et al. 1989). Participants were asked to solve specific algorithm problems requiring the use of these concepts by developing flowchart diagrams, a widely used modelling practice that provides a semiotic space for the design of algorithms (Bohl 1971). The task given to students can be considered an “intellective task” with a "demonstrably correct solution" (Laughlin 1980). The correctness of the solution is concretely defined, based on the notation of algorithms used. There are, of course, alternative ways to develop parts of the solution that are equally acceptable and correct, however, in each case, arguments on correctness and the pros and cons of each alternative can be based on solid criteria. All students were taught the knowledge demanded in order to handle the task sufficiently in university lectures before the lab sessions took place, although some of them may have been already familiar with the task domain from secondary education curricula.

3 Determining Relations between Dimensions of Collaboration Quality

The first step of analysis for the current study dealt with the definition of a conceptual framework of collaboration that is multidimensional, i.e. it defines several dimensions that cover core aspects of collaboration, and, on a second level, can be operationalised into a tool suitable for making assessments of specific instances of collaborative processes. An extended literature search led to the adoption of the work by Meier et al. (2007), who proposed a multidimensional conceptual framework suitable for the assessment of collaboration quality in synchronous interdisciplinary problem-solving through videocoenfencing systems, in a study that significantly influenced the work presented here.

The conceptual framework defined several dimensions of collaboration quality, further categorized into broader aspects of collaboration. The first broader aspect defined regards communication. On a first level, one dimension of collaboration deals with the need for the establishment of common ground of mutually shared concepts, assumptions and expectations (Clark 1996) between the participants. Common ground can be achieved and sustained if both partners (in the case of a dyad) work towards grounding their conversations on a moment-to-moment basis (Clark and Brennan 1991). Good practices in this dimension regard extra effort from the participants in order e.g. for a sender of an utterance to try to make her contribution understandable to their peer, or, for a receiver, to try to indicate understanding of what has been uttered. On a more elementary level, the framework covers, apart from the *content* of communication, the *process* of communicating as well. Practices of participants such as ensuring mutual attention (Clark 1996), and the proper management of the turn-taking mechanism are considered appropriate for the success of a collaborative process.

The second broader aspect of the framework is generally described as information processing. It covers collaborative activity that is tightly related to the task. On a first level, what is of major importance for successful collaboration in this aspect is that participants exchange and process information based on their complementary knowledge, so that they can build a shared knowledge base. In social psychology, the terms of information pooling and transactive memory (Wegner 1987) play a crucial role in describing such processes. From a collaborative learning research standpoint, similar processes can be studied under the term of knowledge acquisition. Knowledge acquisition can be achieved either by externalization of a participant's personal knowledge, or by its elicitation by their peer by asking for explanations (Fischer and Mandl 2003). On another level, after the pooling of information, collaborators have to reach a common decision on the best solution to the problem. In order to achieve that, collaborators have to evaluate the information exchanged, by stating arguments for and against the options at hand, and critically discussing different perspectives (Tindale et al. 2003).

Another important aspect of collaboration regards the coordination of participants on a broader level than the one mentioned before: on the task rather than the communicational level. It deals with practices of efficient structuring of the problem-solving process that involve issues such as the intelligent division into subtasks between participants (e.g. the optimal handling of interdependencies that may occur when subtasks build up on each other, or conflicts when group members need to access the same shared resources (Meier et al 2007)), the efficient management of the time resources available, and the proper handling of coordination demands imposed by the mediating tool's technicalities.

In addition to communication, information processing, and task coordination aspects, social aspects of collaboration are also given due attention by the conceptual framework. Under the frame of the managing interpersonal relationships aspect, the framework covers issues such as interpersonal support, helpfulness and friendliness that can be constructive for successful collaboration. Such desired practices can be reflected in the symmetry of the relationship, the extent of supportive communication and the way the conflicts are handled (Meier 2005).

Finally, the last category covered by Meier et al.'s conceptual framework (2007) deals with motivational aspects of collaboration. Orientation and dedication to the task on behalf of both participants rather than on task-irrelevant issues is considered to be a prerequisite of successful collaboration.

The conceptual framework described above was motivated and led to the definition of a rating scheme suitable for applying assessments of collaboration quality in all of its core dimensions. The rationale of this evaluation approach is described in more detail in the subsequent section.

Table 1 Meier et al.'s general aspects of collaboration

General aspect of collaboration
Communication
Joint information processing
Coordination
Interpersonal relationship
Motivation

4 The Rating Approach

4.1 Rating Scheme

The framework described in the previous section was operationalised through a rating scheme that was used as for assessing collaboration quality in its core dimensions. As an analysis tool, the rating scheme combines desirable properties of qualitative and quantitative techniques. Observed behavior can be compared to a predefined standard of exemplary collaboration that has been formed based on established CSCL theory and thorough empirical analyses of typical collaborative sessions. This can then lead to quantitative judgements of the quality of collaboration. The main advantage of the rating approach compared to common quantitative analysis is that it offers quantitative results that measure subtle aspects of collaboration (mainly being the object of study of in-depth qualitative analyses), rather than gross metrics that are usually based on quantities of events of users' interactions with the mediating tools. Simultaneously, in contrast to the approach adopted in this work, common qualitative approaches that demand in-depth analysis of collaboration usually can not be extended to more than a few rich episodes of collaboration, whereas issues of reliability and generalisability are more difficult to overcome in such cases.

Concerning the practical efficiency of analysis, rating processes are relatively time-effective, since assessment using the rating scheme is rather quick, provided that raters have been sufficiently trained. Therefore, rating schemes can be efficient when dealing with large datasets. Another important characteristic of the rating scheme analysis tool, is that it is multi-dimensional, i.e. it is used for assessing several dimensions of collaboration distinctly. In that way, the results of rating can be used for indicating which dimensions of collaboration are the most problematic and provide e.g. the opportunity for giving adaptive feedback and specialized instructions according to the demands of each separate group. This practice can be followed between different phases of an educational process that take place in sequence in e.g. an academic semester including several related lab sessions as in Meier et al. (2008). Moreover, the approach can still be useful as a first point of analysis for further, more detailed evaluation studies that necessitate more thorough research work.

Due to significant differences between the setting that lead to the definition of Meier, Spada, and Rummel's rating scheme and the current setting, a laborious process of generalising and adapting the initial conceptual framework to the current setting was followed (reported in detail in (Kahrimanis et al. 2009)). The adaptation of the rating scheme was done in two main phases: the first resulted in an adapted definition of the rating scheme's dimensions, and the second served to fine-tune the rating instructions. In the first phase of adaptation a bottom-up approach, which involved identification of "best practice" examples in the sample data, was combined with a top-down process, during which the definitions of all original dimensions were reformulated taking into account constraints and affordances characterizing the specific collaboration setting. In the second phase of adaptation, the dimensions' definitions were fine-tuned and illustrated with more detail, grounding each dimension's theoretical concepts in specific examples of collaborative practice from the data pool of the first round of adaptation (Kahrimanis et al. 2009). The resultant rating scheme specifies seven core dimensions of collaboration quality presented in Table 2.

The structure of the new scheme is in accordance with the rationale of the initial one, while definitions of dimensions do not only aim at fitting the current setting, but at being more generalisable as well.

Table 2 The adapted rating scheme

General aspect of collaboration	Dim. Num.	Dimension of collaboration
Communication	D1	Collaboration flow
	D2	Sustaining mutual understanding
Joint information processing	D3	Knowledge exchange
	D4	Argumentation
Coordination	D5	Structuring the Problem Solving Process
Interpersonal relationship	D6	Cooperative orientation
Motivation	D7	Individual task orientation

4.2 Rating Process

All dimensions were rated on the level of a collaborating dyad with the exception of individual task orientation (D7) which was rated for each participant separately. As in the initial approach by Meier et al. (2007), the ratings were applied in the scale from -2 to 2 with a step of 1 unit. One rating was assigned for each dimension of collaboration quality per collaborative session. A handbook was also developed in order to assist raters providing a rich source of detailed definitions of all dimensions, along with rating instructions and illustrative examples of episodes from the dataset. The rating process was based on video-like reproductions of the activities facilitated by the Synergo's playback tool (Kahrimanis et al. 2009).

The rating procedure was carried out in two main phases. The first one, reported in detail in (Kahrimanis et al. 2009), consisted of 101 dyads which were rated for each dimension by two raters with prior experience with the current setting, after an extended pilot phase of training. Inter-rater reliability scores were very good. The second phase, which was deemed necessary in order to extend the population of collaborative sessions, consisted of additional 149 dyads, for which the design and setting of the labs was identical with the one used in the first phase, varying only in minor aspects of task details (e.g initial values of variables were changed in order to avoid totally repeating the tasks that were given to students in the previous year's academic semester), and was thus appropriate for integrated analysis. This way, the dataset was significantly augmented and large-scale statistical elaborations from several points of view, such as the one presented later in this article, became possible. In the second phase, the ratings were applied by the same persons as in the first one and inter-rater reliability was examined for this phase as well.

4.3 Reliability of Ratings

Results of inter-rater reliability of the second rating procedure for each dimension of the scheme are illustrated in Table 3. For D7, the reliability scores for the average rating between the two students (D7a) and their absolute difference (D7b) are provided. The table contains also reliability scores for the average of the six first dimensions of the scheme (CQ).

Inter-rater reliability scores are good in reference to all empirical rules found in the literature (Fleiss 1981; Cicchetti and Sparrow 1981; Wirtz and Caspar 2002; George and Mallery 2003), although lying at somewhat lowers level than the excellent scores achieved in the first phase. Therefore, the ratings of any one of the raters could be reliably used for further elaborations.

Table 3 Inter-rater reliability scores for each dimension

General aspect of collaboration	Dimension of collaboration		ICC	ICC (adj. = r)	Cronb. α	Spear. ρ
Communication	Coll. flow	D1	.76	.77	.87	.74
	Sust. mut. underst.	D2	.79	.82	.89	.80
Joint information processing	Knowl. exchange	D3	.81	.81	.89	.74
	Argumentation	D4	.77	.77	.87	.75
Coordination	Struct. the Probl. Solv.Proc.	D5	.70	.69	.82	.71
Interpersonal relationship	Coop. orient.	D6	.82	.83	.90	.79
Motivation	Ind. task orient. (mean)	D7a	.71	.75	.84	.59
	Ind. task orient. (diff.)	D7b	.87	.87	.93	.61

5 Relations between Dimensions of Collaboration Quality

To some extent, relations between dimensions of collaboration quality can be roughly conceived from the definition of the conceptual framework and the rating scheme. Nevertheless, the extended dataset gathered offers the opportunity to validate such top-down assumptions empirically, by applying suitable statistical manipulations on the data. Moreover, the exact relations between dimensions can be detected based on the ratings applying in 228 cases.[1]

5.1 Multidimensional Scaling of Dimensions of Collaboration Quality

A systematic way to obtain an overall elaborate view on the associations between collaborative dimensions and empirically evaluate the use of the scheme regards the conduction of a MultiDimensional Scaling (MDS) analysis based on the bivariate correlations between them (Kruskal and Wish 1978; De Leeuw and Heiser, 1982; Schiffman et al. 1981; Davison 1983; Young and Hamer 1994; Borg and Groenen 1997; Cox and Cox 2001). MDS analysis is based on measures of similarities between variable pairs and no assumptions are presupposed on the distribution of the values which the variables take, the types of their similarity relations, or the way the similarity measures are obtained.

In the specific case of this study, the unit of analysis of the technique is the collaborative dimension as it is defined by the rating scheme. The algorithm takes as input the rating assigned to each dimension for 228 instances of collaborative sessions. The technique provides insightful two-dimensional diagrams representing the position of collaborative dimensions in such a way that dimensions correlated tightly are placed closer to each other in space than dimensions that do not relate that much. For the current application of the technique, disparities between correlations are represented with a spatial Euclidian distance.

The MDS algorithm used was SMACOF (Scaling by MAjorizing a COnvex Function) (De Leeuw) as it is implemented by XLSTAT (2009). This iterative algorithm aims at minimizing the normalized differences between a similarity matrix given as input (converted to a dissimilarity matrix) and the corresponding distance matrix that is represented as the outcome of the process.

5.2 Results and Internal Validation of the MDS Algorithm

The results of the application of the technique are depicted in Fig. 2 (using Kendall's τ scores for the calculation of correlations) and Fig. 3 (using Spearman's ρ for the calculation of correlations). The two diagrams are very similar and lead to the same interpretations.

[1] The cases rated in the two phases of training minus 32 cases that were left out of the final dataset due to technical problems in the logfile captures related to them (other kinds of analysis followed later than the work presented here used the logfiles and it was decided that these 32 dyads should be omitted in order to maintain a consistent dataset throughout related works).

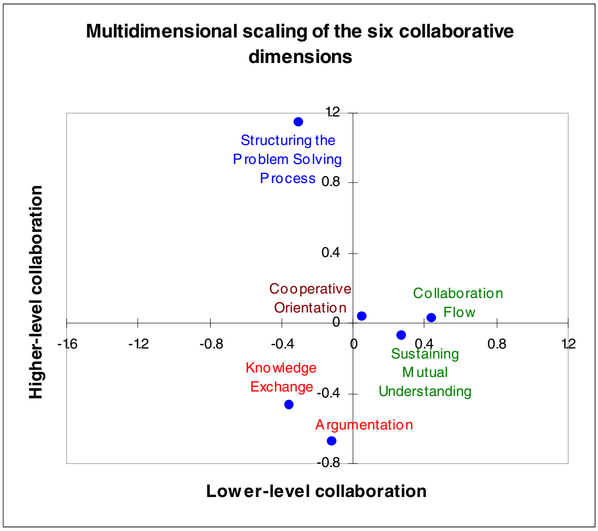

Fig. 2 Multidimensional scaling with the 6 collaborative dimensions using a similarity matrix based on Kendall's correlations

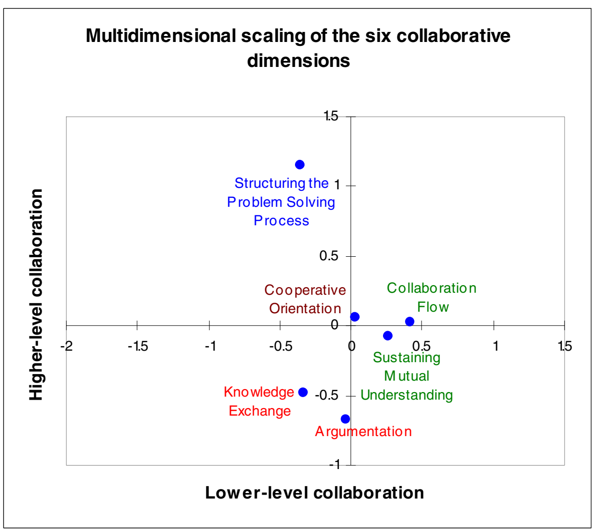

Fig. 3 Multidimensional scaling with the 6 collaborative dimensions using a similarity matrix based on Spearman's ρ correlations

The following Fig. 4 and Fig. 5 illustrate the Shepard diagrams of the application of the technique using Kendall's τ and Spearman's ρ values respectively.

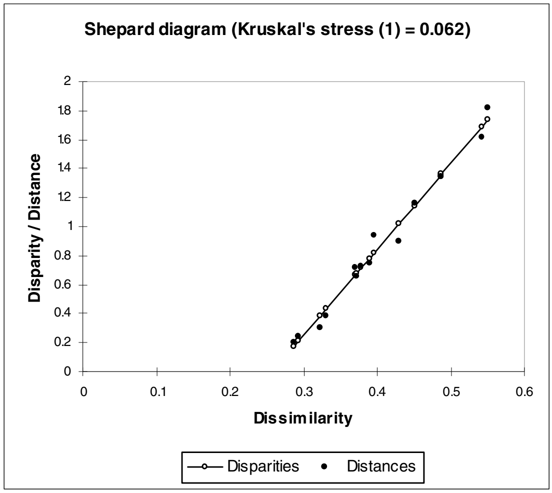

Fig. 4 Shepard diagrams of MDS algorithm using Kendall's τ correlations

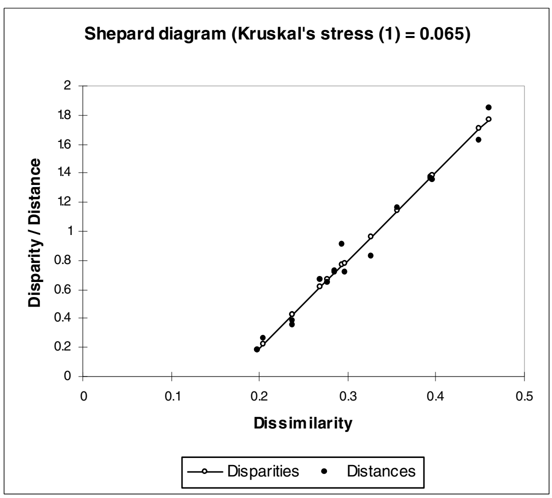

Fig. 5 Shepard diagrams of MDS algorithm using Spearman's ρ correlations

The Shepard diagram (Shepard 1962) is a scatter-plot that depicts the configured distances for the two-dimensional model in relation to the observed distances used as input (Steyvers 2002). The filled circles in the diagram represent the Euclidean distances presented by the MDS algorithm, whereas the empty circles represent the distances calculated by the monotonic regression function of the algorithm (De Leeuw 1977). The latter's slope is represented by the lines of Figures X.4 and X.5. The square root of the normalized sum of squared residuals between the filled circles and the straight line is measured by Kruskal's stress

(Kruskal 1964), which provides an estimation of the goodness-of-fit of the results. Kruskal's stress for the first application of the algorithm was measured at the acceptable level (Kruskal 1964; Borg and Groenen 1997) of 0.062 for the concluding 28th iteration of the algorithm, whereas for the second application it took a similar value (0.065) at the 30th iteration. In both cases, the convergence criterion for the final version of the model stated that the stress value of the final iteration should be improved from the previous iteration by less than 0.00001.

5.3 Interpretation of the MDS Results

As is evident from Figures 2 and 3, dimensions covering different aspects of collaboration quality cover four different parts of the two-dimensional space (the dimension belonging to the motivational aspect is not contained in the diagram since it is rated differently than other dimensions). Dimensions covering the same aspect of collaboration (denoted by the same color in the diagram) stand close to each other. Regarding the interpretation of Figure 2, the coordinates of each dimension do not denote the quality of collaboration in a quantitative manner; they are used for the representation of its distance from other dimensions. Therefore, the range of each axis should be thought of as representing aspects of collaboration that differentiate dimensions on the way they reflect different facets of this specific axis. The rationale followed in order to reach meaningful interpretations is described below.

Higher-order dimensions of collaboration are reported with higher absolute values on the vertical axis, while lower-level ones have higher absolute values on the horizontal axis (cooperative orientation, D6, which is placed near the zero-point does not straightforwardly relate to any of these axes). Thereby, the vertical axis can be considered to stand for high-level collaboration aspects and the horizontal axis to stand for lower-level collaboration aspects.

Concerning the horizontal axis, the two communicational dimensions (D1 and D2) are placed on the right of the diagram, taking positive values, whereas the two information processing dimensions are placed on the left, taking negative values. Thus, from left to right, the horizontal axis can be considered to designate the range from task-related low-level facets of collaborative activity to task-unrelated facets of collaborative activity (task-related in this case refers to these aspects of collaboration that are significantly shaped by the specific task to be solved). In the case of lower level collaborative activity, task-unrelated facets mostly refer to communicational aspects. Collaboration flow (D1) takes the largest positive (and absolute) value on the axis, since it constitutes the lowest-level dimension of the scheme. Sustaining mutual understanding (D2), on the other hand, is placed closer to the zero point and has a more noticeable Y coordinate. Among the information processing dimensions, knowledge exchange (D3) has the biggest negative value because argumentation (D4) is related more to high-level collaborative activity. Structuring the problem solving process (D5) is also placed left from the Y axis. According to the interpretation of the axes developed above, this reflects the fact that structuring the problem solving process is shaped by task-related issues in the lower level of collaboration. For example, an algorithm development problem favors practices of task coordination such as the development of different small parts of the algorithm by participants in parallel (which are arranged according to the task's demands), the proper placement of flow-chart objects by each student so that the two parts can be then combined, or the development of a part of the algorithm by one student while their partner is checking for its correctness by assigning values to variables. Such practices are highly task-dependent and would not be reproduced in the case of a task of a different kind.

A similar rationale applies to the vertical axis concerning higher-level collaboration: structuring the problem solving process (D5) takes the highest absolute value on the upper part of the diagram, while the two information processing dimensions lie on the negative part of the axis. So, in a similar way with the vertical axis, the horizontal one can be considered to denote from up to down the range from task-unrelated high-level facets of collaborative activity to task-related high-level facets collaborative activity. Among the two joint information processing dimensions, argumentation (D4) takes a significantly higher value than information processing (D3), due to the fact that the former reflects higher-level facets of this aspect of collaborative activity. Structuring the problem solving process (D5), on this axis, takes a significantly positive value. Contrary to lower level aspects that the same dimension covers, higher level aspects of structuring the problem solving process are not tight to the specific task. They mainly refer to general strategies of collaborative problem solving, such as the division of labour between participants, the evaluation of one student of the other students work, and time management concerns so that a complete solution can be delivered on time. Communicational dimensions take approximately zero values on the vertical axis, since they are related to lower level aspects of collaboration. Among the two dimensions, sustaining mutual understanding (D2) appears to have a small load on the vertical axis, due to the fact that it is, even limitedly, related more to higher level issues than collaboration flow (D1) is.

Concerning both axes, cooperative orientation (D6) is located very close to the zero point. Social aspects of collaboration that relate to cooperative orientation do not have a straightforward mapping with higher or lower task-related or non task-related facets of collaboration, even though the dimension is correlated highly with all other dimensions of the scheme, something denoted in the diagram with its central position.

In general, results obtained from the MDS algorithm are in accordance with the definition of the dimensions of the rating scheme. Distances represented by the algorithm are reasonable: a diagram of a similar rationale, applied by the researchers in a top-down manner, would probably resemble the one found empirically. Moreover, the approach offers subtler information on the exact associations between dimensions.

Concluding, it should be noted that some properties of the definition of the axes are to some extent arbitrary and their descriptions related to higher and lower-level aspects of collaboration constitute an interpretation rather than an "objective" result of the technique. The examined instance of the MDS technique could lead to the same information with the axes rotated, or their sings inversed. What would remain the same is the relative position of the dimensions (not regarding minimal differences attributed to the goodness-of-fit of the algorithm). Therefore, for the output of the algorithm presented in the figures above, the algorithm was initialized in such a way that the axes would be more interpretable, something that constitutes a common practice when applying MDS or other techniques of similar purpose in several research domains. (Gutmann 1968; Borg and Lingoes 1987).

6 Conclusions and Further Research

The work reported in the current article implemented a rating scheme based approach for the evaluation of synchronous problem-solving collaborative activities in order to gain insight into the relations between distinct dimensions of collaboration quality. The multidimensional scaling approach, which was applied in a large dataset of collaborative sessions, largely confirmed on empirical grounds the rationale of the conceptual framework of the rating tool used, as regards the relations between core dimensions of collaboration quality. Furthermore, it provided additional insight on the exact placement of each dimension of collaboration in reference to two general axes of collaborative activity.

This kind of validation adds evidence that the rating approach can be a reliable tool for evaluating aspects of CSCL activities that are not easily grasped when using strict formalizations, and allow researchers to take full advantage of the practical opportunities that it can offer: the more feasible analysis of large datasets; the provision of a research aid for the conduction of further, more focused research; and the provision of adaptive feedback to students based on their collaborative performance. A pilot study that investigates the tool's application for the latter case is reported in (Meier et al. 2008).

Future research directions related to this work can follow several paths: statistical analysis reported here can be supplemented by in-depth qualitative investigations of collaborative activities, which can shed more light into the way different dimensions of collaboration are interlinked with each other in collaborative practice. Common trends that determine the placement of dimensions of collaboration close or far from each other in the MDS representation can serve as the initial point for further qualitative analysis based on interesting instances of collaboration that may reveal subtler associations between aspects of different dimensions, or recurring patterns of the simultaneous occurrence of good or bad practices in specific dimensions.

Furthermore, the current approach can be replicated using different versions of the rating scheme or applied in different settings of collaboration. Such efforts would help to indicate the extent to which current findings are indicatory of the way dimensions of collaboration are associated with each other in general, or if they mostly pertain to the specific CSCL setting under study.

References

1. Avouris, N., Margaritis, M., Komis, V.: Modelling interaction during small-group synchronous problem-solving activities: The Synergo approach. In: 2nd Int. Workshop on Designing Computational Models of Collaborative Learning Interaction, ITS 2004, Maceio, Brasil (September 2004)
2. Bohl, M.: Flowcharting Techniques. Science Research Associates, Chicago (1971)
3. Borg, I., Lingoes, J.: Multidimensional Similarity Structure Analysis. Springer, Beverley Hills (1987)
4. Borg, I., Groenen, P.: Modern Multidimensional Scaling. Springer, Berlin (1997)
5. Cicchetti, D.V., Sparrow, S.S.: Developing criteria for establishing the interrater reliability of specific items in a given inventory. American Journal of Mental Deficiency 86, 127–137 (1981)
6. Clark, H., Brennan, S.: Grounding in communication. In: Resnick, L.B., Levine, J., Teasley, S. (eds.) Perspectives on Socially Shared Cognition, pp. 127–149. APA Press, Washington, DC (1991)
7. Clark, H.: Using language. Cambrigde University Press, Cambridge (1996)
8. Cox, T.F., Cox, M.A.A.: Multidimensional Scaling. Chapman and Hall, London (2001)
9. Davison, M.L.: Multidimensional Scaling. John Wiley and Sons, New York (1983)
10. De Leeuw, J.: Applications of convex analysis to multidimensional scaling. In: Barra, J., Brodeau, F., Romier, G., van Cutsem, B. (eds.) Recent Developments in Statistics, pp. 133–145. North Holland Publishing Company, Amsterdam (1977)
11. De Leeuw, J., Heiser, W.J.: Theory of multidimensional scaling. In: Krishnaiah, P.R., Kanal, L.N. (eds.) Handbook of Statistics, vol. 2, pp. 285–316. North-Holland, Amsterdam (1982)
12. Dillenbourg, P., Baker, M., Blaye, A., O'Malley, C.: The evolution of research on collaborative learning. In: Reimann, P., Spada, H. (eds.) Learning in Humans and Machines, pp. 189–211. Springer, Berlin (1995)
13. Fischer, F., Mandl, H.: Being there or being where? Videoconferencing and cooperative learning. In: van Oostendorp, H. (ed.) Cognition in a Digital World, pp. 205–223. Lawrence Erlbaum Associates, Mahwah (2003)
14. Fleiss, J.L.: Statistical Methods for Rates and Proportions, 2nd edn. Wiley, New York (1981)
15. George, D., Mallery, P.: SPSS for Windows Step by Step: A Simple Guide and Reference. 11.0 Update. Allyn & Bacon, Boston (2003)
16. Guttman, L.A.: A general non-metric technique for finding the smallest coordinate space for a configuration of points. Psychometrika 33, 495–506 (1968)
17. Kahrimanis, G., Meier, A., Chounta, I.-A., Voyiatzaki, E., Spada, H., Rummel, N., Avouris, N.: Assessing Collaboration Quality in Synchronous CSCL Problem-Solving Activities: Adaptation and Empirical Evaluation of a Rating Scheme. In: Cress, U., Dimitrova, V., Specht, M. (eds.) EC-TEL 2009. LNCS, vol. 5794, pp. 267–272. Springer, Heidelberg (2009)
18. Kahrimanis G., Avouris, A., Komis, V.: Interaction analysis as a tool for supporting collaboration. An overview. In: Daradoumis, T., Caballe, S., Juan, A.A., Xhafa, F. (eds.) Technology-Enhanced Systems and Tools for Collaborative Learning Scaffolding (in press)
19. Kerlinger, F.N., Lee, H.B.: Foundations of behavioral research. Harcourt College Publishers, New York (2000)

20. Kruskal, J.B.: Multidimensional scaling by optimizing goodness-of-fit to a non-metric hypothesis. Psychometrica 29, 1–27 (1964)
21. Kruskal, J.B., Wish, M.: Multidimensional Scaling. Sage Publications, London (1978)
22. Laughlin, P.R.: Social combination processes of cooperative, problem-solving groups on verbal intellective tasks. In: Fishbein, M. (ed.) Progress in Social Psychology, vol. 1, pp. 127–155. Lawrence Erlbaum, Hillsdale (1980)
23. Meier, A., Spada, H., Rummel, N.: A rating scheme for assessing the quality of computer-supported collaboration processes. International Journal of Computer-Supported Collaborative Learning 2, 63–86 (2007)
24. Meier, A., Voyiatzaki, E., Kahrimanis, G., Rummel, N., Spada, H., Avouris, N.: Teaching students how to improve their collaboration: Assessing collaboration quality and providing adaptive feedback in a CSCL setting. In: Rummel, N., Weinberger, A. (eds.) New Challenges in CSCL: Towards Adaptive Script Support, Worshop in Proceedings of the Eighth International Conference of the Learning Sciences (ICLS 2008), Utrecht, vol. 3, pp. 338–345. International Society of the Learning Sciences (June 2008)
25. Roschelle, J.: Learning by collaboration: Convergent conceptual change. Journal of the Learning Sciences 2, 235–276 (1992)
26. Scardamalia, M., Bereiter, C.: Computer support for knowledge-building communities. In: Koschmann, T. (ed.) CSCL: Theory and Practice of an Emerging Paradigm, pp. 249–268. Lawrence Erlbaum Associates, Hillsdale (1996)
27. Samurcay, R.: The concept of variable in programming: Its meaning and use in problem solving by novice programmers. In: Soloway, E., Spohrer, J.C. (eds.) Studying the Novice Programmer, pp. 161–178. Lawrence Erlbaum, Hillsdale (1989)
28. Schiffman, S.S., Reynolds, M.L., Young, F.W.: Introduction to Multidimensional Scaling - Theory, Methods, and Applications. Academic Press, New York (1981)
29. Shepard, R.N.: Analysis of proximities: Multidimensional scaling with an unknown distance function I & II. Psychometrika 27, 125–140 & 219–246 (1962)
30. Soloway, E., Bonar, J., Ehrlich, K.: Cognitive strategies and looping constructs. In: Soloway, E., Spohrer, J.C. (eds.) Studying the Novice Programmer, pp. 191–207. Lawrence Erlbaum, Hillsdale (1989)
31. Stahl, G.: Sustaining group cognition in a math chat environment. Research and Practice in Technology Enhanced Learning (RPTEL) 1(2), 85–113 (2006)
32. Stahl, G., Koschmann, T., Suthers, D.: Computer-supported collaborative learning: An historical perspective. In: Sawyer, R.K. (ed.) Cambridge Handbook of the Learning Sciences, pp. 409–426. Cambridge University Press, Cambridge (2006)
33. Steyvers, M.: Multidimensional Scaling. In: Encyclopedia of Cognitive Science. Macmillan, London (2002)
34. Strijbos, J.W., Martens, R.L., Prins, F.J., Jochems, W.M.G.: Content analysis: What are they talking about? Computers and Education 46(1), 29–48 (2006)
35. Tindale, R.S., Kameda, T., Hinsz, V.B.: Group decision making. In: Hogg, M.A., Cooper, J. (eds.) Sage Handbook of Social Psychology, pp. 381–403. Sage, London (2003)
36. Wegner, D.M.: Transactive memory: A contemporary analysis of the group mind. In: Mullen, B., Goethals, G.R. (eds.) Theories of Group Behaviour, pp. 185–208. Springer, New York (1987)
37. Wirtz, M., Caspar, F.: Beurteilerübereinstimmung und Beurteilerreliabilität. Verlag für Psychologie, Göttingen (2002)
38. XLSTAT, Addinsoft (2009), `http://www.xlstat.com`
39. Young, F.W., Hamer, R.M.: Theory and applications of multidimensional scaling. Erlbaum, Hillsdale (1994)