Rating Dimensions of Collaboration Quality in Synchronous Collaborating Dyads: Findings and Interpretations

Georgios Kahrimanis, Irene-Angelica Chounta, Nikolaos Avouris, Human-Computer Interaction Group, Dept. of Electrical and Computer Engineering, University of Patras, Greece
Email: kahrimanis@ece.upatras.gr, houren@ece.upatras.gr, avouris@ece.upatras.gr

Abstract: Analysis and evaluation of CSCL activities are valuable for both building new knowledge in the research field, and for informing the practice of the design and support of collaborative activities. A major objective regards the development and use of analysis tools that are simultaneously appropriate for conducting meaningful evaluations of collaborative processes and efficient for practical use. This article adopts an innovative approach to CSCL analysis, involving the use of a rating scheme for the assessment of collaboration quality in several of its core dimensions. It involves appropriately trained evaluators that assign ratings on significant aspects of collaboration. Based on statistical analyses of ratings of 228 collaborating dyads working synchronously on a computer science problem-solving task, it reports and interprets interesting findings concerning general trends of collaborative practice, associations between dimensions of collaboration quality, and the way they relate to the quality of the problem's solution.

Introduction

Analysis and evaluation of CSCL activities are valuable for both building new knowledge in the research field, and for informing the practice of the design and support of collaborative activities. Several methodological approaches of analysis of CSCL have been suggested that range from deep-level qualitative analyses of small interaction-rich episodes of collaboration, to quantitative approaches focusing on measuring significant aspects of collaborative activities by statistical means (Stahl, Koschmann, & Suthers, 2006). Moreover, the goal of analysis is often shaped by practical purposes, aiming at regulating collaborative processes in real time (Soller, Martínez, Jermann, & Muehlenbrock, 2005), scaffolding learners in order to build new knowledge (Reiser, 2002), or informing the design and structuring of the conditions of the CSCL process (Kollar, Fischer, & Hesse, 2006), including the design of the tools that mediate them (Suthers & Hundhausen, 2003).

A major objective regards the development and use of analysis tools that are simultaneously appropriate for conducting meaningful evaluations of collaborative processes, rather than being based on "surface" counts of interaction events, and efficient for practical use. A rather newly applied approach in the field regards the use of *rating schemes* for evaluation of collaborative activities that involves human agents assigning ratings to predefined aspects of collaboration. In more detail, a rating scheme or a rating scale is "a measuring instrument that requires the rater or observer to assign the rated object to categories or continua that have numerals assigned to them" (Kerlinger & Lee, 2000, p. 736, cited in Meier 2005). Rating schemes are discriminated from coding schemes in that they are used to make a judgment on a larger piece of data each time, and are based on the knowledge and the critical skill of the human agent that applies them, in contrast to coding schemes that demand from the coder to neutralize the process by following strictly defined rubrics (Kerlinger & Lee 2000). So far, the rating approach has been applied in the CSCL field for assessing the "level of perspective taking" in asynchronous online discussions (Järvelä & Häkkinen, 2003) and the quality of collaboration through videoconferencing in synchronous problem-solving (Meier, Spada, & Rummel, 2007).

As an analysis tool, the rating scheme combines desirable properties of qualitative and quantitative techniques. Observed behavior is compared to a predefined standard that has been formed based on established CSCL theory and empirical analyses of typical collaborative sessions. This can then lead to *quantitative* judgments of the *quality* of collaboration. Therefore, the rating approach offers quantitative results that reflect subtle aspects of collaboration accessible to the human intellect but not easily detectable using any strict formalizations, or content analysis rubrics. Moreover, it is more time-efficient than deeper-level qualitative approaches, suitable for supporting CSCL practices, such as the provision of feedback to participants, targeted at the aspects of collaboration for which problems are reported. In addition, the rating approach allows assignment of grades by tutors for collaborative performance and not just for the correctness of the task's outcome. Ratings can also be useful for unraveling trends in collaborative practices in large populations, where qualitative approaches are not feasible, with the advantage that the ratings that constitute the object of analysis cover deeper-level aspects of collaboration than most other quantitative approaches. Finally, results of the rating process can be valuable as research aids, used as "quick indicators where more detailed analyses are merited, thereby focusing the detail work" (Stahl et al., 2006, p. 13).

This article builds on the rating tool developed by Meier et al. (2007), and adapted by Kahrimanis et al. (2009), and applies the approach to a large dataset of 228 synchronous collaborative problem-solving dyads. It investigates common trends in this population, related to the several dimensions of collaboration quality defined, and concludes to some interesting interpretations on factors influencing collaborative performance. In addition, the associations between dimensions of collaboration quality are investigated statistically, validating empirically the design of the rating tool. A third investigation made possible by the scale of the dataset relates collaboration quality with the quality of the outcome of the process, in this case the problem solution diagram. The intuitive assumption that good collaboration leads to good task performance is tested on statistical terms. The article concludes with discussion and proposals for further research built on this work.

Collaborative Setting

The study involved about 350 first year students of 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 that was part of an introductory to computing course. Randomly formed dyads of students interacted through Synergo (Avouris, Margaritis, & Komis, 2004), communicating via an integrated chat tool, and jointly designing a flow-chart representation of an algorithm on a shared workspace. Collaborative sessions took place in a university laboratory room and lasted from 45 to 75 minutes. Students were free to use their own resources such as textbooks, however the feedback they received to their questions was restricted to technical support. In order to motivate students to work collaboratively, they were informed that the grade would be formed 50% by the quality of their collaboration and 50% by the completeness and correctness of their joint solution. Moreover, students were given general instructions on what constitutes good collaborative practice according to the core dimensions of collaboration. Dyads were spaced in a way that partners communicated exclusively through Synergo.

The domain of the task was basic computer algorithms. Students were asked to build and investigate the correctness of elementary algorithms using flowchart diagrams (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 used. All students were taught the knowledge demanded in order to handle the task sufficiently in university lectures before the lab sessions, even though some of them may have been already familiar with the task domain from secondary education.

The Rating Approach

Rating Scheme

The conceptual framework that shaped the analysis approach applied in this work was developed by Meier, et al. (2007). The framework was operationalized through a rating scheme. Due to significant differences between the setting that led to the definition of Meier’s, et al. (2007) scheme and the current setting, a laborious process of generalizing and adapting the initial framework to the current setting was followed (reported in Kahrimanis et al., 2009). The resultant scheme specifies seven core dimensions of collaboration quality, shown in Table 1.

Table 1: Dimensions of collaboration quality.

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

Rating Procedure and Reliability

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. The rating was conducted using video-like reproductions of the collaborative processes provided by the Synergo playback tool. The pilot phase involved the rating of the 1/3 of the dataset jointly by the two raters, aiming at overcoming discrepancies between them and potential misunderstandings on the rationale of the rating process, and simultaneously avoiding “over-training” that may lead to artificial consensus reached by succumbing to trivial aspects of collaboration. The second 1/3 of the dataset was rated in parallel by each rater individually and the comparison of concordance in their ratings led to very good inter-rater reliability results (Kahrimanis et al., 2009). The last 1/3 of the data was divided in two equal parts and each rater assigned ratings to each part separately.

The second phase, which was deemed necessary in order to extend the population of collaborative sessions, consisted of additional 149 dyads for which the setting of the labs was identical with the first phase, varying only in minor aspects of task details, being thus appropriate for integrated analysis. This way, large-scale statistical elaborations became possible. In the second phase, the ratings were applied by the same raters as in the first one and new tests of inter-rater reliability were made. The two raters co-rated 1/3 of the dataset (50) dyads, and split the remaining 99 dyads in two approximately equal parts, so that each rated half of them.

Overall results of inter-rater reliability of the whole dataset for each dimension of the rating scheme are shown in Table 2. For dimension D7, the reliability scores for the average rating (D7a) and the absolute difference between the ratings of the two students (D7b) are shown. The scores are good for all dimensions and indicate that the rating process was reliable according to established empirical rules (Fleiss, 1981; Cicchetti & Sparrow, 1981; Garson, 2009).

The final integrated dataset consists of the sum of dyads from the two phases (101 + 149 = 260) minus instances of collaborative sessions that were not considered due to technical problems in the logging mechanism of the Synergo tool caused by network failures during a few sessions.

Table 2: Reliability scores for each dimension of collaboration quality.

General aspect of collaboration covered	Dimension of collaboration	ICC	ICC (cons. adj.) = r	Cronbach's. α	Spearman's. ρ	Kendall's τ	Same rating
Communication	D1	.85	.86	.92	.85	.77	67%
	D2	.86	.87	.93	.84	.77	64%
Joint information processing	D3	.91	.91	.95	.89	.83	74%
	D4	.87	.87	.93	.86	.78	64%
Coordination	D5	.86	.86	.92	.83	.78	72%
Interpersonal Relationship	D6	.92	.92	.96	.89	.83	77%
Motivation	D7a (average)	.83	.84	.91	.80	.76	73%
	D7b (abs. diff.)	.88	.89	.94	.78	.74	74%
	Average D1-D6	.95	.95	.98	.93	.82	-

Measuring Dimensions of Collaboration Quality

Descriptive Findings

The first part of analysis regards descriptive statistical investigations of the dyads' collaboration quality for each dimension of the rating scheme. The current sample extends the size of a typical dataset used in CSCL research and can lead to interesting insights on distinct aspects of collaborative behavior as they emerge in real-world, large-scale activities. Therefore, descriptive measures such as the mean, median, mode, and standard deviation were calculated (Table 3), as well as the distribution of grades for each dimension (Figure 1).

Table 3: Descriptive metrics of collaborative dimensions of the rating scheme.

	D1	D2	D3	D4	D5	D6	D7a	D7b	Avg. D1-D6
mean	0.41	0.43	0.62	0.50	0.48	0.84	1.52	0.73	0.55
median	1	1	1	1	1	1	2	0	0.83
mode	1	1	2	2	2	2	2	0	2
std. dev.	1.24	1.31	1.34	1.44	1.35	1.33	0.67	1.15	1.14

The distribution of ratings for all collaboration dimensions indicates a positive bias, i.e. the dyads that were rated with a rating greater than 0 in all collaboration dimensions outnumber the dyads that were rated with a negative rating. Figure 1 illustrates the histograms of the distributions of values for dimensions D1 to D6.

The first two communicational dimensions (D1, D2) appear to have the lowest mean values compared to all others, although they are still positive. Dyads tended to score higher in information processing dimensions (D3, D4) than in communicational ones. Between these two dimensions, better ratings were obtained for D3, *knowledge exchange* (which describes lower-level practices of the information processing aspect), while D4, *argumentation* ratings followed a more diverse distribution. Ratings for D5 were more moderate, influenced by

45.6% of the dyads that where characterized by negative or neutral collaboration quality in *structuring the problem solving process*. On the contrary, D6, *cooperative orientation* was deemed successful for most of the dyads (65.8% got at least 1) and negative instances were rather rare, something reflected in the mean value of this dimension which is the highest reported among D1-D6.

Concerning dimension D7, which is used to assess *individual task orientation* at the level of each collaborating partner, its average scores (D7a) tended to be significantly higher than ratings of any other dimension. 56.1% of the dyads were assigned with the highest score in D7 for both the collaborating partners, while 71.5% had an average rating of at least 1. Cases where the average of the ratings for the two students was 0 or less were limited to 8.3% of the whole population. Concerning D7b, in most cases (63.2%) it was equal to zero, i.e. both students were assigned the same rating in D7. From this subset, 88.9% relates to cases where both students where highly committed to the task, while the remaining dyads were not assigned with any negative ratings either. Therefore, no dyads in the whole dataset were reported for which both of the partners were not motivated to solve the task. This can be explained by the circumstances of the lab, where at least one of the students took the responsibility to, at least, solve the problem on her own. On the other hand, in one out of five dyads (19.7%), significant imbalance in the commitment towards the task was noticed (difference of at least 2 grades in the scale), indicating that one student tended to take responsibility and dominate the process. Finally, the overall average rating of collaboration quality had a mean value equal to 0.55 and a median value of 0.83. 35.5% of the dyads had an average of 0 or less, whereas for the rest 64.5% the average was positive. 42.5% of the dyads had an average value of 1 or more.

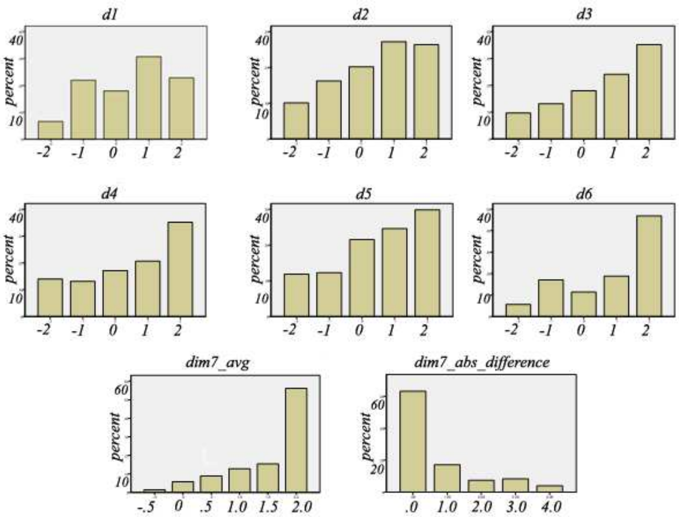

Figure 1. Distributions of values of dimensions D1-D6 of the rating scheme and the average and absolute difference in ratings of D7 (dim7_avg, dim7_abs_dif) for each participant.

Discussion on Measuring Collaboration Quality

On a general level, collaboration quality of the examined population of dyads was deemed successful in statistical terms. The relatively good performance of most dyads can be attributed to two main factors. First, a major influence was due to the motivations given to the students as 50% of their final grade was based on the assessment of the quality of collaboration; second, in the introduction of the collaborative activities by the tutors, students were encouraged to follow desired practices for each collaborative dimension.

Regarding the distinct dimensions of collaboration, students performed best at motivational and interpersonal aspects, they were mostly good at joint information processing aspects and had moderate performance on communicational and task coordination aspects. From these findings, a clear trend can be discerned that discriminates dimensions of collaboration quality between those that can be mastered by students based on their general or intuitive skills and those for which good collaboration quality can be achieved after a period of appropriation or if collaborative competences have been developed from prior experience. For the first category of dimensions, conscious intentions of collaborating partners to improve on them are usually achieved based on motivations and simple explanations of good practices in these specific dimensions, such as to communicate their task-related elementary knowledge, or to be pleasant and kind at the social level. On the other hand, general instructions given to students on how they should perform in dimensions such as *sustaining mutual understanding* and *structuring the problem solving process* are not so easily comprehensible by them. Moreover, even if instructions are communicated sufficiently, in order to achieve good performance more collaboration-specific skills are needed.

From this discussion, it can be inferred that, apart from giving sufficient motivations to students, a systematic approach towards achieving better collaboration quality in future activities should involve appropriate instructions on collaboration with more emphasis on the dimensions more difficult to learn. This can be done by arranging some targeted training sessions, by scripting the collaborative process in order to scaffold collaborative performance especially with regards to D5 (*Coordination*), or by providing feedback to participants if possible.

Associations between Dimensions of Collaboration

Associations between dimensions of collaboration quality can be conjectured by the definition of the rating scheme. Still, the extended dataset gathered offers the opportunity to validate such top-down assumptions empirically, by applying suitable statistical manipulations, as the relations between dimensions can be detected based on the ratings of collaboration quality applied in 228 cases. A suitable statistical technique is *MultiDimensional Scaling (MDS)* analysis based on the bivariate correlations between dimensions (Young, & Hamer, 1994).

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. Calculations are applied to a correlation matrix between the dimensions, as it was formed based on the ratings of each of the 228 dyads. 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 spatial Euclidian distances. The MDS algorithm used was SMACOF (Scaling by MAjorizing a COnvex Function) (De Leeuw, 1977).

Results and Internal Validation of the MDS Algorithm

The results of the application of the technique are depicted in Figure 2 (using Kendall's τ scores in the correlation matrix). A similar diagram was the result of the application of the same approach using Spearman's ρ coefficients instead (Kahrimanis, Chounta, & Avouris, 2010).

Moreover, in Figure 2, the Shepard diagram (Shepard, 1962) is included that serves for the evaluation of the algorithm (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 Leuw, 1977). Moreover, Kruskal's stress for the application of the algorithm was measured at the acceptable level (Kruskal, 1964; Borg, & Groenen, 1997) of 0.062 for the concluding 28th iteration of the algorithm.

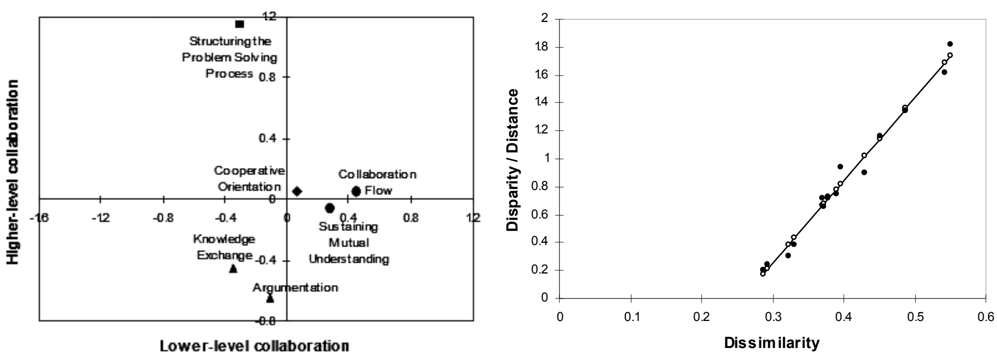

Figure 2. Multidimensional scaling diagram (left) and Shepard diagram (right) with collaboration quality dimensions D1-D6 using a similarity matrix based on Kendall's τ correlation score.

Interpretation of MDS Analysis

As it is evident from Figure 2, dimensions covering different aspects of collaboration quality cover different parts of the two-dimensional space. Dimensions covering the same aspect of collaboration (denoted by the same symbol in the diagram) stand close to each other. Regarding the interpretation of Figure 2, the coordinates of each dimension are used for the representation of its distance from other dimensions. Therefore, the range of each axis should be thought of as representing aspects that differentiate dimensions in the way they reflect different facets of collaboration. The rationale followed in diagram interpretation is described next.

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 (D6, *cooperative orientation*, which is placed near the zero point does not 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* (D3 and D4) 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 ones.

In the case of lower level collaborative activity, task-unrelated facets mostly refer to communication. *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 zero and has a more noticeable Y coordinate. Among the information processing dimensions, *knowledge exchange* (D3) has the highest negative value since *argumentation* (D4) is related more to high-level collaborative activity. Coordination (*Structuring the problem solving process* D5) is also placed to the left of the Y axis. According to the proposed axes interpretation, this reflects the fact that D5 is shaped by task-related issues in the lower level of collaboration in contrast to task-unrelated aspects in higher-level collaboration. In general, results obtained from the MDS analysis 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.

In conclusion, it should be noted that some properties of the definition of the axes are to some extent arbitrary and their relation to higher and lower-level aspects of collaboration constitute an interpretation rather than an "objective" result of the technique. For the output of the algorithm presented in Figure 2, the algorithm was initialized in such a way that the axes would be more interpretable, something that constitutes a common practice when applying MDS in other research domains (Guttman, 1968; Borg & Lingoes, 1987).

Relation between Collaboration Quality and Solution Quality

As already discussed, students were motivated for collaborating, while the outcome of the collaborative task, i.e. the algorithm diagram delivered, counted for 50% of the grade. The solution to the problem was graded separately, in a scale from 0 to 10. The assessments of the solution were quite straightforward since the task in all its variations had a demonstrably correct solution (Laughlin, 1980) with a few easily defined alternatives in parts of the diagrams, in some cases. The histogram of the grades (for N= 228 dyads) is provided in Figure 3.

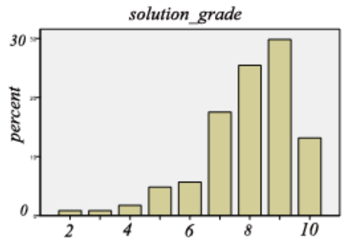

Figure 3. Frequencies of solution grades in the population in percentages.

The mean value for the sample was 7.96, with median = 8, and mode = 9 (68 times in 228 cases). 8.3% of the dyads were graded with 5 or less indicating a bad solution, 23.3% with 6 or 7 and 57% with 8 or more. Grades were generally high, something expected due to the nature of the tasks, which, although not too trivial, were of limited difficulty in order to allow researchers to focus their interest on the study of collaboration.

Table 4: Correlations between solution grades and dimensions of the rating scheme and their derivatives.

	D1	D2	D3	D4	D5	D6	D7a	D7b	Avg D1-D6
Solution Grade	.248** .301**	319** .391**	.287**. .352**	310** .380**	.402** .487**	.288** .344**	.311** .369**	-.238** -.282**	.349** .453**

** p < 0.01

The most interesting aspect related to the solution quality, regards the correlation of the solution grades with collaborative performance in general and with particular dimensions of the rating scheme in particular. Table 4 summarizes these correlations for the dataset (N=228). The values are Kendall's τ (the top cell) and Spearman's ρ (the bottom cell). For all combinations presented in Table 4, almost-medium to almost-high correlation scores between dimensions of collaboration quality and solution grade are observed. Still, the level of correlation differs significantly between different pairs.

The dimension that is correlated with the grade of the solution at the highest level is D5: *structuring the problem solving process*. This is reasonable because serious problems in the way dyads coordinate the collaborative process on a higher level, that often lead to mismanagement of their time, have often direct consequences on the completeness and correctness of the solution. The correlation is heavily influenced by cases where students do not manage to complete the task and therefore get lower grades. Next, D2: *sustaining mutual understanding* correlates at the second highest level with solution grade. It is thus apparent, that comprehensible communication between students can favor the development of a good solution, without, nevertheless, playing a determinant role. The correlation of the other communicational dimension: D1, *collaboration flow,* can be totally attributed to the variance it shares with D2. Moreover, information processing dimensions D3 and D4, are correlated at a medium level with solution grade, These are mostly related to the development and evaluation of the task per se and dyads that frequently exchange information and provide explanations to each other and get engaged in argumentation-intense collaborative sessions are more likely to develop a correct and complete diagram. D7: *individual task orientation* is also correlated with the solution grade judging by both its derivative measures. Summing up, with the exception of D2, in relation to the MDS analysis reported above, dimensions that are linked to higher-level aspects of collaboration are correlated with collaborative outcome at the highest extent.

Regarding the correlation between solution grade and average collaboration quality, which is τ=.349, ρ=.453, it indicates medium to high correlation. This result is in accordance with the experience of CSCL practitioners that good collaboration is bound to relate to a good solution. It is also in accordance with similar evidence provided in some CSCL studies in different knowledge domains that prove the relationship between collaborative outcome and different kinds of measures of interaction (Chiu, 2004; Manlove, Lazonder, & de Jong, 2006). Nevertheless, the extent of this association is not so strong as to allow judgment of the solution quality by the measure of collaboration quality alone.

Conclusions and Further Work

This article presented the application of an innovative evaluation approach in a large set of synchronous collaborative activities. Based on an approach of adopting an existent rating scheme (Meier, et al. 2007) to the needs of the CSCL setting (Kahrimanis et al., 2009), the dataset used permitted various statistical elaborations.

Descriptive statistics revealed that, in relation to collaboration quality, students performed best at motivational and interpersonal aspects, were mostly good at joint information processing aspects and their performance lied at a moderately good level in communicational and task coordination aspects. It was thus indicated that the hardest aspects of collaboration are those that demand experience with the communicative setting and collaborative practice in general. Still, a majority of dyads got a positive grade on most collaborative dimensions. Success in collaboration was possible in most cases, based on general instructions from lab supervisors plus motivations for collaborating, through the grading rules used.

A multidimensional scaling analysis of the associations between dimensions of collaboration quality served to further validate the design of the rating scheme from an empirical standpoint, reassuring its initial assumptions and shedding light on the exact placing of each dimension in relation to others.

Solution quality was also correlated with dimensions of the rating scheme at a medium-to-high level. Task coordination and time management were the aspects correlating at the highest level with the correctness and completeness of the diagrams delivered, while, in general, dimensions covering higher-level aspects of collaboration were correlated more with solution quality than lower-level aspects. Findings reported are noteworthy, not allowing however assessment of collaboration quality to be based solely on solution quality.

Potential future work involves the use of the current approach and the findings obtained as a research aid for further deeper-level qualitative analysis of collaboration sessions of similar type. Moreover, the behavior of large populations of collaborative dyads in relation to dimensions of collaboration quality can be further investigated by applying other advanced statistical techniques such as clustering of collaborating dyads. A further goal regards the unraveling of common patterns of problematic dimensions of collaboration and trying to overcome them by appropriately restructuring the activity design. Furthermore, since, as it is claimed, the approach followed in this study leads to meaningful results in quantitative form, it can be used as a basis for evaluating less subtle logfile-based automated techniques of CSCL interaction analysis, and possibly reshape them so that they are optimally targeted towards reflecting substantial aspects of collaboration quality.

Endnotes

(1) According to an established empirical rule (Cohen, 1988; Hopkins, 2000), a correlation of r = .1 is considered to be low, of r =.3 medium, and of r=.5 or more high. This rule of thumb was initially proposed for Pearson product-moment coefficient r. These are implied here for Spearman's ρ scores (that resembles Person's r for non-parametric data).

References

Avouris N., Margaritis M., & Komis V. (2004). Modelling interaction during small-group synchronous

problem-solving activities: The Synergo approach, *ITS2004, Maceio, Brasil, September 2004*.
Bohl, M. (1971). *Flowcharting Techniques*. Chicago: Science Research Associates.
Borg, I., Lingoes, J. (1987). *Multidimensional Similarity Structure Analysis*. Beverley Hills: Springer-Verlag.
Borg, I, & Groenen P. (1997). *Modern Multidimensional Scaling*. Berlin: Springer-Verlag.
Chiu, C. H. (2004). Evaluating system-based strategies for managing conflict in collaborative concept mapping. *Journal of Computer Assisted Learning*, 20, 124-132.
Cicchetti D.V., & Sparrow, S.S (1981). Developing criteria for establishing the interrater reliability of specific items in a given inventory. *American Journal of Mental Deficiency*, 86, 127-137.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.
De Leeuw, J. (1977), Applications of convex analysis to multidimensional scaling, in J. Barra, F. Brodeau, G. Romier & B. van Cutsem, (Eds), *Recent developments in statistics* (pp. 133-145). Amsterdam, The Netherlands: North Holland Publishing Company.
Fleiss, J.L. (1981) *Statistical Methods for Rates and Proportions*, 2nd. Edition. New York: Wiley.
Garson, G. D. (2009). Quantitative Research in Public Administration. Retrieved November 4, 2009, from http://faculty.chass.ncsu.edu/garson/PA765/pa765syl.htm.
Guttman, L. A. (1968). A general non-metric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, 33, 495–506.
Hopkins WG. (2000). A scale of magnitudes for effect statistics. In *A new view of statistics*. http://www.sportsci.org/resource/stats/effectmag.html#cohen. Retrieved on 21/08/2009
Järvelä, S., & Häkkinen, P. (2003). The levels of web-based discussions: Using perspective-taking theory as an analytical tool. In H. van Oostendorp (Ed.), *Cognition in a digital world* (pp. 77-95). Mahwah, NJ, USA: Lawrence Erlbaum Associates.
Kahrimanis, G., Meier, A., Chounta, I.A., Voyiatzaki, E., Spada, H., Rummel, N., & Avouris, N. (2009). Assessing collaboration quality in synchronous CSCL problem-solving activities: Adaptation and empirical evaluation of a rating scheme. *Lecture Notes in Computer Science*, 5794/2009, 267-272, Berlin: Springer-Verlag.
Kahrimanis, Chounta, I.A., Avouris, N. (2010). Determining relations between core dimensions of collaboration quality: A multidimensional scaling approach. Int. Conf. on Intelligent Networking and Collaborative Systems, INCOS, Thessaloniki, November 2010.
Kerlinger, F. N., & Lee, H. B. (2000). *Foundations of behavioral research*. New York: Harcourt College Publ.
Kollar, I., Fischer, F., & Hesse, F.W. (2006). Collaboration scripts–a conceptual analysis. *Educational Psychology Review*, 18, 159-182.
Kruskal, J. B. 1964. Multidimensional scaling by optimizing goodness-of-fit to a non-metric hypothesis. *Psychometrica*, 29, 1-27.
Laughlin, P. R. (1980). Social combination processes of cooperative, problem-solving groups on verbal intellective tasks. In M. Fishbein (Ed.), *Progress in social psychology, vol.1* (pp. 127-155). Hillsdale, NJ, USA: Lawrence Erlbaum.
Manlove, S., Lazonder, A. W., & de Jong, T. (2006). Regulative support for collaborative scientific inquiry learning. *Journal of Computer Assisted Learning*, 22, 87-98.
Meier, A. (2005). Evaluating Collaboration: A Rating Scheme for Assessing the Quality of Net-Based, Interdisciplinary, Collaborative Problem Solving. Unpublished diploma thesis, Institut für Psychologie, Albert-Ludwigs-Universität Freiburg im Breisgau, Freiburg, Germany.
Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *Int. Journal of Computer-Supported Collaborative Learning*, 2, 63–86.
Reiser, B.J. (2002). Why scaffolding should sometimes make tasks more difficult for learners. In *Proceedings of the Computer Support for Collaborative Learning (CSCL 2002), Boulder, Colorado, USA*, 255 - 264.
Shepard, R.N. (1962). Analysis of proximities: Multidimensional scaling with an unknown distance function I & II. *Psychometrika*, 27, 125-140, 219-246.
Soller, A., Martínez, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of the state of the art technology for supporting collaborative learning. *Int. J. on Artificial Intelligence in Education*, 15, 261–290.
Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences,* 409-426. Cambridge: Cambridge University Press.
Steyvers, M. (2002). Multidimensional Scaling. In *Encyclopedia of Cognitive Science*. London, Macmillan.
Suthers, D. D., & Hundhausen, C. (2003). An experimental study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences*, 12(2), 183–219.
Young, F.W., & Hamer, R.M. (1994). *Theory and applications of multidimensional scaling*. Hillsdale: Erlbaum.