Available online at www.sciencedirect.com
ScienceDirect
Computers in Human Behavior 23 (2007) 1105–1125
Computers in Human Behavior
www.elsevier.com/locate/comphumbeh

Visualization of agreement and discussion processes during computer-supported collaborative learning

Jeroen Janssen *, Gijsbert Erkens, Gellof Kanselaar

Research Centre Learning in Interaction, Utrecht University, P.O. Box 80140, 3508 TC, Utrecht, The Netherlands

Available online 15 November 2006

Abstract

This study examined the effects of the shared space (SS) on students' behaviors in a computer-supported collaborative learning (CSCL) environment. The SS visualizes discussion and agreement during online discussions. It was hypothesized the SS would increase the media richness of the CSCL-environment, would stimulate critical and exploratory group-norms, would lead to more positive perceptions of online collaboration, and would have an impact on students' collaborative activities. In total, 59 students working in 20 groups had access to the SS visualization, while 58 students working in 20 groups did not. The results show that students with access to the SS visualization: (a) perceived higher media richness; (b) had a more exploratory group-norm perception; (b) perceived more positive group behavior; (c) perceived their group's task strategies to be more effective; (d) engaged in different collaborative activities and (e) performed better on one part of the group task. These results demonstrate the potential benefits of visualizing agreement and discussion during CSCL.

Keywords: Computer-supported collaborative learning; Collaboration; Computer-mediated communication; Visualizations; Secondary education

Teachers and students are increasingly using ICT to facilitate learning of various subjects (Lou, Abrami, & d'Apollonia, 2001). Computer-supported collaborative learning

* Corresponding author. Tel.: +31 30 2534798; fax: +31 30 2532532.
E-mail address: j.j.h.m.janssen@fss.uu.nl (J. Janssen).

0747-5632/$ - see front matter
doi:10.1016/j.chb.2006.10.005

(CSCL) is one application of ICT that has received considerable attention by educational researchers. CSCL aims to provide students with an environment that supports and enhances collaboration, in order to facilitate their learning processes (Kreijns, Kirschner, & Jochems, 2003). During CSCL students usually communicate with group members using discussion forums or chat rooms.

Several studies demonstrated CSCL to be an effective tool for education. For example, a meta-analysis by Lou et al. (2001) found that combining small group learning with ICT was more effective than combining individual learning with ICT. Additionally, Cavanaugh (2001) demonstrated the effectiveness of interactive distance education technologies. Thus, the perceived potential of CSCL seems to be, at least partially, supported by research.

Notwithstanding the positive effects of using CSCL, many studies have also demonstrated possible pitfalls when using CSCL (Kreijns et al., 2003). For example, students communicating through computer-mediated communication (CMC) sometimes perceive their discussions as more confusing (Thompson & Coovert, 2003), demonstrate higher levels of personal conflict (Hobman, Bordia, Irmer, & Chang, 2002) or participate in unsustained, low quality discussions (Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003). In sum, positive and productive interaction is sometimes lacking during CSCL.

The following sections discuss two problems that may occur during CSCL, namely communication and discussion difficulties, and students' difficulties to conduct critical, yet constructive discussions. Possible explanations for these problems will be described in short. This is followed by a description of how these problems can be addressed.

1. Communication difficulties during CSCL

Research has demonstrated that it is difficult for group members to communicate during CSCL (Fjermestad, 2004; Fuks, Pimentel, & Lucena, 2006). Some researchers have argued that the communication problems found during CSCL may be due to the medium itself. Traditional text-based CMC systems, such as chat, are seen as media that are low in *media richness* (Daft & Lengel, 1986). Media richness is defined as a medium's ability to facilitate communication and the establishment of shared meaning. Factors such as the ability of the medium to transmit multiple cues (e.g., facial expressions, gestures), and the immediacy of feedback influence its media richness. As media richness decreases, students will have more difficulties conveying their opinions and will have more difficulties determining the meaning of group members' messages.

Furthermore, when working on group tasks students usually work on complex problems without demonstrably correct answers, which require them to resolve differing viewpoints. The type of communication usually used during CSCL, may not be suited to the types of tasks group members work on (Mennecke, Valacich, & Wheeler, 2000). The low media richness of CSCL-environments may constrain collaboration in such a way that it does not transmit the type of communication that group members need to solve their task successfully. This may lead to communication problems and decreased task performance.

2. Lack of critical but constructive discussion during CSCL

When group members collaborate, they are usually working on complex problems, which require the input of all group members. Ideally, group members engage in discus-

sions that are critical, but also constructive. This means that group members are critical of their own and the other group members' ideas, that criticism is accepted, and that they offer explanations for their opinions. These types of discussions have been called exploratory discussions and have been found to enhance learning during group work (Wegerif, Mercer, & Dawes, 1999). However, research has shown that students rarely give arguments and counter arguments (Kuhn & Udell, 2003; Munneke, Andriessen, Kanselaar, & Kirschner, 2006), nor do they offer explanations for their ideas regularly during CSCL (Van der Meijden & Veenman, 2005).

This absence of critical but constructive discussion may be explained in several ways. First, students may not know how to conduct such discussions and may not posses the necessary skills (Weinberger & Fischer, 2006). Second, as stated above, students may find it difficult to conduct constructive debates in a CSCL-environment and may have difficulties interpreting discussions. For example, they may not know whether group members agree or disagree with them. This possibly hampers argumentation and discussion (Adrianson & Hjelmquist, 1991). Finally, groups may possess group-norms that stimulate consensus among group members, instead of critical or exploratory discussion. Group-norms that stimulate consensus instead of critical discussion can contribute to the low quality of some online discussions (Erkens, Prangsma, & Jaspers, 2006; Postmes, Spears, & Cihangir, 2001). In conclusion, the relative absence of critical discussion during CSCL may have different causes. These causes need to be addressed in order to facilitate critical but constructive discussions during CSCL.

3. Addressing communication and discussion problems using visualizations

This section describes how visualizations of online dialogue may help address the previously described communication problems and the relative lack of critical but constructive discussion. Several researchers noted the lack of social cues of CSCL-environments (e.g., Donath, 2002). For users of chat rooms or discussion boards, it is often very difficult to quickly determine, who the participants of online discussions are, or what the social norms of the online group are (Lee, Girgensohn, & Zhang, 2004). This lack of awareness may constrain social interaction and lead to lower perceived quality of the social space (Kreijns, Kirschner, Jochems, & Van Buuren, 2004). To address this problem, several researchers have turned to visualization techniques that visualize important social features of the environment. For example, Donath (2002) and Janssen, Erkens, Kanselaar, and Jaspers (2006) developed techniques for visualizing participation during online discussions. It is expected that by using such visualizations, social awareness can be increased, which may in turn lead to more productive interaction.

However, visualizing participation may not be sufficient to overcome communication problems and to stimulate critical discussions, because it only visualizes quantitative aspects of collaboration, not qualitative. Therefore, a visualization called shared space (SS) was developed. The SS visualizes whether group members are agreeing or disagreeing about a topic during online discussion. This visualization has been implemented in an existing CSCL-environment, called Virtual Collaborative Research Institute (VCRI, see Jaspers, Broeken, & Erkens, 2005). More specifically, the SS is an extension of the *Chat tool* of the VCRI-program. Students use this tool to communicate synchronously. The SS analyzes all messages entered in the Chat tool. For a more detailed description of the VCRI-program, the reader is referred to the Tasks and Materials section.

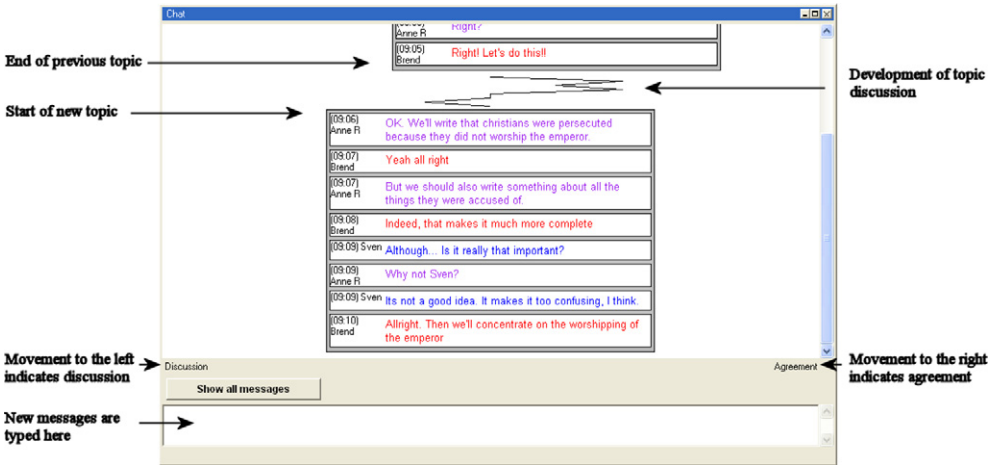

Fig. 1. Screenshot of the chat tool with shared space visualization.

First, the SS discerns discussion topics using time intervals. When students do not type messages for more than 59 s, a new topic begins. Fig. 1 shows a screenshot of VCRI's Chat tool with SS visualization. The screenshot shows the end of one topic, and the beginning of a new one.

Second, the SS analyzes the content of each chat message in order to determine whether it indicates discussion or agreement. For this purpose, the SS determines the communicative function of the message. This is done using the dialogue act coding (DAC) filter (see Erkens, Jaspers, Prangsma, & Kanselaar, 2005). This filter uses over 1300 rules based on discourse markers to determine the communicative function of a chat message. Discourse markers are characteristic words or phrases signaling the communicative function of a message. In total, five main categories of communicative functions are distinguished: argumentative, responsive, informative, elicitative, and imperative. Each category consists of several subcategories, 29 in total. Of these, confirmations, acknowledgements, and positive evaluations are considered indications of agreement, while denials, verification questions, negative evaluations, and counterarguments are considered indications of discussion or debate. In a prior study (Erkens et al., 2005), the reliability of the DAC filter was tested and found to be acceptable (over 90% of all messages coded correctly).

Finally, after establishing whether the message indicates discussion or agreement, the SS moves the current topic to the left or to the right in small steps. When a message indicates discussion, the SS moves the topic to the left; when it indicates agreement, the SS moves the topic to the right. The lines above the topics visualize the development of the online discussion. For example, in Fig. 1, at the beginning of the topic, the SS indicated agreement (the line moves to the right), whereas later on the SS indicated debate (the line goes to the left).

It is hypothesized that the SS visualization will help group members overcome the communication and discussion problems described above for several reasons. First, the SS may increase the media richness of the CSCL-environment. Because the SS visualizes discussion and agreement, it may be easier for students to determine the meaning of messages. Additionally, it may be easier to identify the different views and positions held by group

members. Moreover, the SS may help group members to determine whether there is shared understanding about a topic.

Second, the SS provides group members with feedback about the manner in which they are conducting their discussions. For example, when the SS keeps moving to the right, this tells group members they may be engaged in an uncritical discussion. Thus, the feedback provided by the SS visualization may increase students' awareness about their conversational strategies and their group-norms.

Finally, by providing them with feedback and raising their awareness, the SS may help students to engage in group processing. This occurs when group members discuss how well their group is functioning and how group processes may be improved (Yager, Johnson, Johnson, & Snider, 1986). During these discussions group members may be stimulated to adopt more critical or exploratory group-norms.

In conclusion, it is expected that SS visualization may address some of the communication problems that occur during CSCL, and may help group members to collaborate and discuss more productively.

4. Research questions

This paper investigates the effects of the SS visualization on online collaboration. The following research questions are addressed: Do students with access to SS visualization, compared to students without access, ...

1. ... perceive higher media richness when using the Chat tool?
2. ... perceive different, more critical or exploratory group-norms?
3. ... perceive their online collaboration and communication more positively?
4. ... engage in different collaborative activities?
5. ... perform better on an inquiry group task?

4.1. Method and instrumentation

4.1.1. Design

A posttest-only design with a treatment and a control group was used to answer the research questions. Treatment group students had access to the Chat tool with SS visualization, whereas control group students used the same Chat tool but without SS visualization. In order to stimulate productive interaction, students worked in small, three-person groups (Schellens & Valcke, 2006). However, due to the size of some classes and illness of students, some groups consisted of two or four students. In total, there were 33 groups of three, five groups of two, and two groups of four students. Each group of students was randomly assigned to either the treatment or the control group. The treatment group consisted of 20 groups (59 students; two two-person, 17 three-person, and one four-person group). Similarly, the control group also consisted of 20 groups (58 students; three two-person, 16 three-person, and one four-person group).

4.1.2. Participants

All 117 eleventh-grade (54 male, 63 female) participants came from five different history classes from two secondary schools in The Netherlands. These students were enrolled in

the second stage of the pre-university education track. Mean age of the students was 16.17 years ($SD = 0.60$, $Min = 15$, $Max = 18$). Students were randomly assigned to a group by the researchers. In order to prevent combinations of students who could not get along with each other, their teachers checked the group compositions. As a result, three students were re-assigned.

4.2. Task and materials

4.2.1. CSCL-environment

Group members collaborated in a CSCL-environment called VCRI. The VCRI is a groupware program designed to support collaborative learning on research projects and inquiry tasks. Students use the program to collaborate in small groups. Every group member works at one computer. Students use the *Chat* tool to communicate synchronously with group members (see Fig. 2). To read the description of their group task or to search and read relevant information, students can use the *Sources* tool. This tool lists a number of sources, which can be opened and read from the screen. Group members use the *Cowriter* as a shared word processor. Using the *Cowriter*, group members can work simultaneously on different parts of their texts. To collaboratively construct (argumentative) diagrams, students can use the *Diagrammer*. The VCRI-program contains several other tools not shown in Fig. 2. For example, the *Planner*, which can be used to develop plans and assign tasks.

An alternative version of the VCRI-program was available for teachers. Using this so-called *Coach*-program, teachers can monitor the online discussions of their students. Teachers can also send messages in order to answer students' questions, or to warn students in case of misbehavior. Furthermore, teachers have access to the texts students are writing in the *Cowriter*. This way, teachers can monitor the progress of their groups.

4.2.2. Inquiry group task

Participating students worked together on a historical inquiry group task. Subject of the task was "The first four centuries of Christianity" and consisted of three parts. The

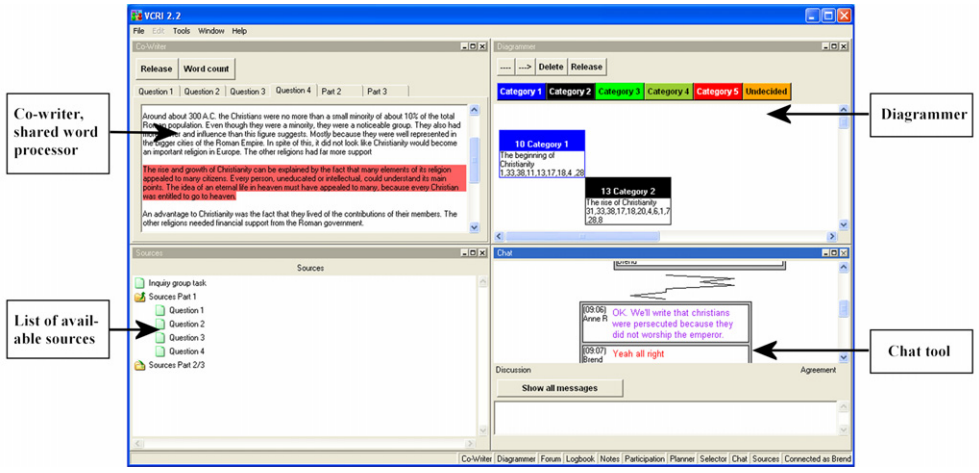

Fig. 2. Screenshot of the VCRI program.

introduction of the task stressed the importance of working together as a group to successfully complete the inquiry task. Students were told they had eight lessons to hand in their reports, and they would receive a group grade for their reports.

For the first part of the inquiry task, the groups had to answer four different questions pertaining to the first four centuries of Christianity. To answer these questions, 12 different sources were available to the students. These sources were, for example, fragments from the New Testament, and historical texts from the Roman era. Additionally, students could search the Internet or their textbooks for more information. To complete the second part of the task, the groups had to study 40 different sources about the subject. These sources needed to be categorized into up to five different categories. Furthermore, group members were instructed to construct a diagram of their categorization using the Diagrammer. Finally, students had to write a short text, explaining how and why they categorized the different sources. For the final part of the inquiry task, group members had to collaboratively write an essay of at least 1200 words. The essay had to explain why and how Christianity developed from a small 'cult' into the main religion of the Roman Empire.

Comparable to the task used by Munneke et al. (2006), the task can be characterized as an open-ended task without a standard procedure and no single right answer. Furthermore, the inquiry task was quite complex and comprehensive; therefore, no single group member was likely to solve the task on his or her own. Thus, the participation of all group members was necessary to successfully complete the task.

4.2.3. Procedure

During the lessons each student worked on a separate computer in a computer lab. Students sat as far from their teammates as possible, in order to stimulate them to use to the VCRI-program to communicate with their other group members. Before the first computer lesson, students received information about the task and the group compositions from their teachers. During the lessons, the teachers answered task-related questions, while the experimenters solved technical problems. Students were allowed to work on the inquiry group task during free periods. Thus, they could work on the task in the school's media center when they had spare time in their timetable. After eight lessons, the groups handed in their final versions of the group task.

After the last lesson, a questionnaire was administered to the students. This questionnaire contained several items pertaining to perceived media richness, group-norm perception, and perception of online collaboration and communication. In total, the questionnaire contained 48 items. Students expressed their opinions using a 5-point scale ranging from 1 (=completely disagree) to 5 (=completely agree). Due to absence or sickness, 20 students did not complete the posttest questionnaire. Thus, the total number of respondents was 97 students.

4.3. Measures

4.3.1. Perceived media richness of the Chat tool

To measure perceived media richness of the Chat tool, the questionnaire contained a 15-item scale that addressed various aspects pertaining to the media richness of the Chat tool. The items addressed whether students were aware of agreements and disagreements during online discussions, and whether they could explain things easily to group members (Dennis, Kinney, & Hung, 1999). "During discussion in the Chat it is clear whether there

is agreement among group members'' is a sample item from the scale. Students' ratings were averaged to create a score for ''*perceived media richness*'' ($\alpha = 0.92$). Higher scores indicate higher perceived media richness.

4.3.2. Group-norm perception

To measure students' perceptions of group-norms, the questionnaire completed by the students contained three scales. The first scale consisted of three items, and asked students whether they perceived their group as having *critical group-norms*. The items were based on the work of Postmes et al. (2001). A sample item of this scale was: ''Our group is critical''. The second scale investigated whether students perceived their group as having *consensual group-norms*. This scale also consisted of three items based on the work of Postmes et al. (2001). An example from this scale is: ''In this group people generally adapt to each other''. Finally, the third scale examined whether group members perceived their group as having *exploratory group-norms*. This scale consisted of seven items, modeled after the ground rules for exploratory talk formulated (Wegerif et al., 1999). ''During discussions, criticism and counterarguments were accepted'' is a sample item of this scale.

For all three scales, students' responses to the items were averaged to obtain scores for *critical* ($\alpha = 0.84$), *consensual* ($\alpha = 0.59$), and *exploratory* ($\alpha = 0.74$) *group-norm perception*. Higher scores on these variables correspond to a more critical, consensual, or exploratory *group-norm* perception.

4.3.3. Perception of online collaboration and communication

To measure how students perceived their online collaboration and communication the questionnaire contained three scales. The first scale consisted of seven items (e.g., ''We helped each other during collaboration'') and addressed *positive group behavior* (Janssen, Erkens, & Schep, 2006), such as equal participation of group members, and helping group members (Webb, 1995).

Five items formed the second scale. These items (e.g., ''There were conflicts in our group'') addressed occurrences of *negative group behavior* (Janssen et al., 2006) such as conflicts and free riding behavior (O'Donnell & O'Kelly, 1994).

The final scale addressed students' *perceived effectiveness of their group's task strategies*. This scale was based on the work of Saavedra, Earley, and Van Dyne (1993) and consisted of eight items that assessed the choices made and the strategies chosen by the group members. An example from this scale is: ''We planned our group work effectively''.

Again, for all three scales students' responses were averaged to compute a score for *positive* ($\alpha = 0.82$) and *negative group behavior* ($\alpha = 0.68$), as well as for *perceived effectiveness of group task strategies* ($\alpha = 0.81$).

4.3.4. Collaborative activities

In order to examine the influence of the SS on students' collaborative activities, a coding scheme was used (Janssen et al., 2006). The aim of this coding scheme was to provide insight into the task- and group-related processes during students' online collaboration.

4.3.4.1. Description of the coding scheme. The coding scheme consists of four different dimensions. Each dimension contains two or more coding categories. Furthermore, the scheme includes several additional categories (e.g., technical aspects) that did not belong

to any of the four dimensions. In total, the scheme consists of 19 categories. The first dimension referred to *performance of task-related activities* aimed at carrying out the task (Jehn & Shah, 1997). This dimension contained two categories pertaining to the discussion of relevant task-related information: exchanging and sharing task-related information (TaskExch) and asking task-related questions (TaskQues). The abbreviations of the codes are given between parentheses.

The second dimension referred to *regulation and coordination of task-related activities*, containing four categories. Metacognitive activities that regulate task performance (e.g., making plans, monitoring task progress), are considered important to successful performance in electronic learning environments (Narciss, Proske, & Koerndle, 2006; Van der Meijden & Veenman, 2005). First, planning (MTaskPlan) involved discussion of strategies necessary to complete the task, and delegation of task responsibilities. Second, monitoring (MTaskMoni) involved exchange of information that could be used to monitor task performance and progress, and assessing the amount of time available. Finally, evaluation involved appraisal and discussion of task performance and progress, which could be either positive (MTaskEvl+) or negative (MTaskEvl−).

Since group members also have to attend to the social and emotional element of collaboration to successfully complete a group task (Rourke, Anderson, Garrison, & Archer, 1999) *performance of social activities* constituted the third dimension. This dimension contained five categories. First, greetings (SociGree) contribute positively to group atmosphere (Rourke et al., 1999). Second, social support remarks (SociSupp) referred to comments that contributed positively to group atmosphere, such as exchanging positive comments, and disclosure of personal information. Third, social resistance remarks (SociResi) referred to behaviors that contributed negatively to group atmosphere, such as insults and displaying negative emotions. Fourth, shared understanding (SociUnd+) referred to confirmations and indications of agreement, which serve to reach and maintain joint understanding. Similarly, loss of shared understanding (SociUnd−) referred to denials, and expressions of incomprehension.

The fourth dimension referred to *regulation and coordination of social activities*. These might be called meta-social activities. For example, group members need to discuss collaboration strategies or reflect on the manner in which they collaborated. This dimension contained four categories. First, planning (MSociPlan) involved discussion of collaboration strategies, such as helping each other, or proposals to work together on certain tasks. Second, monitoring (MSociMoni) referred to the exchange of information that could be used to monitor group processes. Finally, evaluation involved appraisal and discussion of group processes and collaboration, which could be positive (MSociEvl+) or negative (MSociEvl−).

Statements that addressed neutral, negative, or positive technical aspects of the CSCL-environment were also included in the coding scheme (TechNeut, TechNega, and TechPosi). Finally, statements that did not fit into any of the previously mentioned categories were coded as Other, referring to nonsense and off-task remarks.

4.3.4.2. Segmentation and coding procedure. During online collaboration some students only send one sentence per message, while others type several sentences that combine multiple clauses. Furthermore, even within in a single sentence, multiple concepts or statements may be expressed (Strijbos, Martens, Prins, & Jochems, 2006). Thus, it may be necessary to segment a chat message into smaller parts that are meaningful in their selves.

Therefore, the chat messages were segmented into *dialogue acts* (Erkens et al., 2005). One dialogue act corresponds to a sentence or a part of a compound sentence that can be regarded meaningful in itself and has a single communicative function.

Segmentation and coding were done using the multiple episode protocol analysis (MEPA) computer program (Erkens, 2005). Messages were segmented into dialogue acts using a *segmentation filter*. A filter is a program, which can be specified and used in MEPA for automatic rule based coding or data manipulation. The segmentation filter automatically segments messages into dialogue acts, using over 150 decision rules. Punctuation marks (e.g., full stop, exclamation mark, question mark, comma) and connecting phrases (e.g., "and if", or "but if") are used to segment messages into dialogue acts. Using filters speeds up segmentation, and ensures segmentation rules are applied consistently. After the segmentation process, the dialogue acts were subsequently coded using the coding scheme.

4.3.4.3. Interobserver reliability. In an earlier study (Janssen et al., 2006), a satisfactory overall Cohen's Kappa of .86 was found. The category Kappa's (Cicchetti, Lee, Fontana, & Dowds, 1978) ranged from .67 to 1.00. For the purpose of the current study, one rater coded 796 collaborative activities from four random protocols from the previous study. The results of this coding were compared to the results of the previous study. An overall Cohen's Kappa of .94 was found (category Kappa range: .78–1.00).

4.3.5. Group performance scores

In order to measure the effect of the SS on group performance, an assessment form was developed for each part of the inquiry task. The assessment form for the first part addressed (1) the *conceptual content and the quality of the argumentation* of the answers, and (2) the *quality of the presentation*of the answers (Scarloss, 2002). *Conceptual content and quality of argumentation* were assessed using one item on a 4-point scale. For example, an answer that received one point, contained little or no relevant historical concepts and little or no argumentation, whereas a an answer that received four points contained all relevant historical concepts and adequate argumentation. *Quality of the presentation* was assessed using five items (e.g., correctness of the language used, copy-pasting of sources directly into the text, structure of the written answer) that were rated on a 3-point scale. For example, concerning correctness of the language used, groups received zero points if their answer contained seven or more language errors, one point if it contained four to six errors, and two points if it contained three or less errors.

During part two of the task, group members needed to study and categorize sources into categories, construct a diagram of their categorization in the *Diagrammer*, and explain how and why they categorized the sources. The assessment form for this part consisted of three items, which assessed the quality and completeness of the constructed diagram and the quality of the explanation. These items were rated on a 3-point scale. For instance, concerning quality of the explanation, groups received zero points if they did not formulate an explanation, one point if they wrote an explanation that did not explain completely why certain categories had been chosen, and two points if the explanation addressed why categories had been chosen and why sources had been placed in certain categories.

For the last part of the inquiry task, group members needed to collectively write an essay. Comparable to part one, *conceptual content and quality of argumentation* were assessed using three items rated on a 3-point scale. *Quality of the presentation* of the essay

Table 1
Interrater reliability of the group performance assessment

Part of the inquiry task	% Agreement (range)	Cohen's κ (range)
Part one		
Conceptual content and argumentation	89.29% (85.71–100.00%)	0.85 (0.73–1.00)
Presentation	84.52% (82.14–85.71%)	0.74 (0.71–0.76)
Part two	91.67% (87.50–100.00%)	0.87 (0.80–1.00)
Part three		
Conceptual content and argumentation	95.24% (85.71–100.00%)	0.90 (0.59–1.00)
Presentation	85.71% (85.71–85.71%)	0.72 (0.61–0.73)

was assessed using five items (e.g., structure of the essay, correctness of language used, correct use of historical sources) on a 3-point scale. This was done in a similar fashion as for part one of the inquiry task.

To check the objectivity of the assessment procedure, two researchers scored seven inquiry tasks. The results of reliability analysis are presented in Table 1. The percentages agreement and Cohen's kappa's indicate the assessment procedure was reliable.

4.3.6. Data analysis

To investigate the effects of the SS on students' collaborative activities during CSCL, an independent samples *t*-test with condition (SS or no SS) as an independent variable could be used. However, students' collaborative activities are *nonindependent* (Kashy & Kenny, 2000), which makes a *t*-test inappropriate. This is caused by *mutual influence* (Kenny, Mannetti, Pierro, Livi, & Kashy, 2002). That is, what one group member says, is influenced by, and influences the contributions of other group members. Therefore, students who are in the same group behave in more or less similar ways. Thus, it is expected that students who are, for example, in a group with group members who are focused on task-related activities, will also be stimulated to focus on task-related activities. To address the problem of nonindependence, *multilevel analysis* was used to examine the effects of the SS.

The multilevel analyses involved estimating two models: an empty model and a model including one or more predictor variables. By comparing the deviance of the latter model to the empty model, a decrease in deviance can be calculated. When this decrease in deviance is significant (tested with a χ^2-test), the latter model is considered a better model. Additionally, the estimated parameters of the predictor variables can be tested for significance by dividing the regression coefficient, β, by its standard error, yielding a *t*-value. A significant *t*-value indicates a significant effect of the predictor.

The line of reasoning concerning nonindependence of students' collaborative activities can be extended to the other individual measures used in this study (research questions one, two, and three). Thus, the effects of the SS on students' perceived media richness, *group-norm* perception, and perception of online collaboration and communication, will be examined using multilevel analysis as well.

4.4. Results

4.4.1. Perceived media richness of the chat tool

On average, treatment group students perceived higher media richness ($M = 3.26$, $SD = 0.80$) compared to control group students ($M = 3.01$, $SD = 0.76$). Furthermore,

the effect of the SS was significant, $\beta = .124$, $p = .03$. However, the associated χ^2-value was only marginally significant, $\chi^2 = 2.41$, $p = .06$.

4.4.2. Group-norm perception

It was expected that treatment group students would perceive their *group-norms* as more critical and less consensual. The results presented in Table 2 show that this expectation was only partially confirmed. Regarding critical group-norm perception, no effects of the SS were found, $\beta = .000$, $p = 1.00$. Similarly, no differences were found regarding consensual group-norm perception, $\beta = .022$, $p = .18$. However, the multilevel analyses revealed a significant effect of the SS on exploratory group-norm perception, $\beta = .108$, $p = .01$. This indicates that treatment group students perceived that their groups were engaged more in critical but constructive online discussions, compared to control group students.

4.4.3. Perception of online collaboration and communication

The results for research question three are presented in Table 3. Treatment group students reported more occurrences of positive group behavior compared to control group students, $\beta = .155$, $p = .01$. Moreover, treatment group students perceived their group's task strategies to be more effective, compared to control group students, $\beta = .165$, $p = .00$. However, treatment group students reported similar levels of negative group behavior compared to control group students, $\beta = -.100$, $p = .05$.

4.4.4. Collaborative activities

In Table 4, the mean frequencies for collaborative activities are presented. The numbers in parentheses indicate the percentages of the total number of collaborative activities that were devoted to a specific activity. To examine the effect of the SS on students' collaborative activities, multilevel analyses were used as well. In this case, two predictors were added to the multilevel models. Besides condition (SS or no SS), number of dialogue acts typed was included in the model. This was done to account for the fact that some groups typed more dialogue acts and were generally more active than others. By including this predictor, the effect of the SS could be investigated independent of number of dialogue acts typed by students.

Table 4 also lists the results of the multilevel analyses. Number of dialogue acts typed was a significant predictor for all collaborative activities, except negative evaluations of social activities (*MSociEvl−*). Thus, in most cases participation was related to collaborative activities. For example, the more a student participated during online discussions, the more questions he or she asked (TaskQues).

Additionally, condition was a significant predictor for five collaborative activities. First, the SS had a negative effect on the number of task-related remarks (TaskExch) made by students, $\beta = 3.45$, $p = .04$. This effect should be interpreted with caution however, as the associated χ^2-value was only marginally significant ($p = .08$). Second, having access to the SS was negatively related to the number of task-related questions (TaskQues) asked by a student, $\beta = -2.18$, $p = .00$. Third, the β for SociUnd+, indicates a negative effect of the SS on the number of messages which were aimed at reaching and maintaining mutual understanding, $\beta = -6.95$, $p = .02$. Fourth, the SS had a positive effect on students' use of positive evaluations of group activities, $\beta = .20$, $p = .04$. Again, this effect should be interpreted with caution, as the χ^2-value was only marginally significant ($p = .08$). Finally, the

Table 2
Means for group-norm perception and results of multilevel analyses

	Treatment group students ($N = 48$)		Control group students ($N = 49$)		β	SE	χ^2
	M	SD	*M*	SD			
Critical group-norm perception	3.24	0.86	3.25	0.65	0.000	0.085	0.000
Consensual group-norm perception	3.50	0.89	3.46	0.66	0.022	0.062	0.123
Exploratory group-norm perception	3.82	0.53	3.60	0.53	0.108*	0.053	3.933*

Note: Mean scores along scales ranging from 1 (=completely disagree) to 5 (=completely agree).
* $p < 0.05$.

Table 3
Means for perception of online collaboration and communication and results of multilevel analyses

	Treatment group students ($N = 48$)		Control group students ($N = 49$)		β	SE	χ^2
	M	SD	M	SD			
Positive group behavior	3.93	0.54	3.62	0.58	0.155**	0.067	4.909*
Negative group behavior	2.34	0.72	2.54	0.68	−0.100	0.079	1.532
Effectiveness of group task strategy	3.73	0.56	3.42	0.62	0.165**	0.065	6.066**

Note: Mean scores along scales ranging from 1 (=completely disagree) to 5 (=completely agree).
* $p < 0.05$.
** $p < 0.01$.

Table 4
Mean frequencies and standard deviations of collaboration acts and results of multilevel analyses of the effects of condition

	Treatment group students ($N = 59$)			Control group students ($N = 58$)			Total ($N = 117$)			Effect of condition		
	M			*M*			*M*					
	Frequency	(%)	SD	Frequency	(%)	SD	Frequency	(%)	SD	β	SE	χ^2
Performing task-related activities												
Info exchange (TaskExch)	19.75	(7.43)	18.89	24.84	(6.63)	27.89	22.27	(7.02)	23.81	−3.45*	2.41	2.00
Asking questions (TaskQues)	8.17	(4.11)	6.99	12.33	(3.01)	10.42	10.23	(3.56)	9.06	−2.18**	0.81	6.62**
Coordinating/regulating task-related activities												
Planning (MTaskPlan)	61.32	(21.61)	41.84	63.71	(21.47)	45.35	62.50	(21.54)	43.44	−2.62	4.29	0.37
Monitoring (MTaskMoni)	36.31	(12.93)	22.25	36.48	(12.64)	22.62	36.39	(12.78)	22.34	0.12	−0.85	0.23
Positive evaluations (MTaskEvl+)	6.12	(1.87)	5.51	5.72	(1.98)	6.14	5.92	(1.92)	5.80	0.12	0.63	0.03
Negative evaluations (MTaskEvl−)	5.92	(2.32)	5.66	6.31	(1.77)	5.07	6.11	(2.04)	5.36	−0.32	0.35	0.85
Performing social activities												
Greetings (SociGree)	8.71	(3.31)	7.10	10.31	(3.06)	10.31	9.50	(3.18)	8.68	−0.98	1.12	0.76
Social support (SociSupp)	31.93	(8.64)	24.05	30.03	(0.51)	30.03	30.99	(9.58)	30.49	0.28	2.88	0.01
Social resistance (SociResi)	10.69	(2.91)	9.87	8.67	(3.66)	8.67	9.69	(3.29)	10.12	0.90	1.08	0.69
Mutual understanding (SociUnd+)	54.15	(21.69)	31.74	65.26	(19.09)	65.26	59.66	(20.38)	41.59	−6.95*	3.68	3.44*
Loss of mutual understanding (SociUnd−)	11.68	(4.12)	8.66	11.29	(4.07)	11.29	11.49	(4.09)	8.32	−0.04	0.73	0.00
Coordinating/Regulating social activities												
Planning (MSociPlan)	4.98	(1.29)	4.98	3.98	(1.56)	3.59	4.49	(1.42)	4.36	0.46	0.46	0.98

(*continued on next page*)

Table 4 (*continued*)

	Treatment group students ($N = 59$)			Control group students ($N = 58$)			Total ($N = 117$)			Effect of condition		
	M			*M*			*M*					
	Frequency	(%)	SD	Frequency	(%)	SD	Frequency	(%)	SD	β	SE	χ^2
Monitoring (MSociMoni)	14.25	(3.83)	10.68	12.36	(5.07)	11.89	13.32	(4.45)	11.29	0.76	1.07	0.50
Positive evaluations (MSociEvl+)	0.76	(0.07)	1.81	0.34	(0.32)	1.21	0.56	(0.20)	1.55	0.20*	0.14	2.06
Negative evaluations (MSociEvl−)	0.63	(0.15)	0.96	0.48	(0.25)	1.14	0.56	(0.20)	1.05	0.06	0.09	0.47
Technical												
Neutral technical (TechNeut)	4.02	(1.41)	3.85	4.45	(1.38)	4.92	4.23	(1.39)	4.40	−0.33	0.46	0.54
Negative technical (TechNega)	2.14	(1.09)	2.65	3.24	(0.64)	4.50	2.68	(0.86)	3.71	−0.61*	0.32	3.39*
Positive technical (TechPosi)	0.49	(0.10)	0.88	0.34	(0.13)	0.78	0.42	(0.12)	0.83	0.06	0.06	0.81
Other	6.20	(1.14)	10.36	3.84	(2.77)	6.98	5.03	(1.96)	8.89	1.12	0.91	1.50

* $p < 0.05$.
** $p < 0.01$.

SS had a negative effect on the number of negative technical remarks (TechNega) made, $\beta = -.61$, $p = .02$. Treatment group students typed less negative comments about the program.

4.4.5. Group performance scores

The performance scores for the different parts of the inquiry task are given in Table 5. Note that, since the performance scores were calculated for groups, not individual students, the number of observations is different from, for example, Table 4. As group performance is a *group level* variable, one-tailed *t*-tests for independent samples were used to examine differences between treatment and control groups. From Table 5 it becomes clear that treatment groups received significantly higher scores for conceptual content and quality of argumentation for part one of the task, $t(38) = 1.88$, $p = .03$, $d = 0.59$. Furthermore, the quality of the presentation of part one of the inquiry task was significantly higher for treatment groups, $t(38) = 2.52$, $p = .01$, $d = 0.80$. No significant differences were found for the other parts of the task.

To determine the effect of the SS on group members' contribution to the written group products, additional analyses were performed. For each student, the number of characters typed in the Cowriter was calculated. Subsequently, for each group a Gini coefficient was calculated (Kiesler & Sproull, 1992). This coefficient indicated the equality of contributions to the Cowriter (0 = perfect equality, 1 = perfect inequality). On average, students contributed rather equally to the group products ($M = 0.22$, SD = 0.11). Furthermore, no significant differences were found between treatment and control groups, $t(38) = -1.06$, $p = .30$.

4.5. Conclusion and discussion

During online conversations in CSCL environments, group members often experience communication difficulties (Fjermestad, 2004). Furthermore, the discussions conducted by group members are often shallow and uncritical. The present study addressed these problems by investigating the effect of visualizing agreement and discussion between

Table 5
Means and standard deviations for treatment and control groups for group performance scores

Group performance scores	Treatment groups ($N = 20$)		Control groups ($N = 20$)		*t*	*p*	*d*
	M	SD	*M*	SD			
Part one							
Conceptual content and argumentation	0.58	0.14	0.50	0.11	1.88	0.03*	0.59
Presentation	0.63	0.11	0.54	0.09	2.52	0.01*	0.80
Part two	0.65	0.24	0.64	0.24	0.11	0.46	0.03
Part three							
Conceptual content and argumentation	0.68	0.27	0.68	0.33	0.00	1.00	0.00
Presentation	0.53	0.25	0.45	0.23	1.05	0.15	0.33

Note: To increase comparability, scores for the different parts were standardized, with 0 being the minimum amount of points and 1 the maximum.

* $p < 0.05$.

group members during CSCL. A visualization called shared space (SS) was developed and implemented in an existing CSCL-environment. Based on an in-depth, automated analysis of chat messages typed by group members, the SS visualizes whether there is agreement or debate amongst them. It was hypothesized that giving students access to the SS visualization would be beneficial to online collaboration. In order to examine this hypothesis, a posttest-only design with a treatment and a control group was used. Treatment group students had access to the SS visualization, whereas control group students did not.

The first research question examined whether treatment group students perceived online communication to be easier and more efficient than control group students. This hypothesis was confirmed, because questionnaire data confirmed treatment group students perceived the medium as more media rich.

The second research question considered whether the SS had an impact on students' group-norm perceptions. No differences were found regarding critical and consensual group perception. However, treatment group students reported their group held a more exploratory group-norm perception. Therefore, it may be concluded the SS helped group members to value critical but constructive online discussions.

The third research question investigated the effect of the SS on students' perceptions of online collaboration. Treatment group students reported more occurrences of positive group behavior and perceived their group's task strategies to be more effective, compared to control group students. However, reported levels of negative group behavior were similar for both groups. In conclusion, the SS seemed to have a positive effect on students' perceptions of their online collaboration.

The fourth research question addressed whether treatment group students were engaged in different collaborative activities than control group students. For some activities this hypothesis was confirmed. For example, treatment group students exchanged relatively less task-related information or questions than control group students. Furthermore, they typed fewer messages aimed at reaching and maintaining mutual understanding. This may indicate the SS helped students to correctly interpret whether there was agreement or discussion during their online discussions, thereby decreasing the need to maintain mutual understanding. Finally, the SS influenced treatment group students to send more positive evaluations of their collaboration. This indicates that SS stimulated students to engage in group processing, which is important for effective collaboration (Yager et al., 1986).

The last research question concerned whether treatment groups performed better on the inquiry group task. Treatment groups performed significantly better on the first part of the task, but not on the second and third part. This partial effect of the SS may be explained by the limited effect of the SS on students' collaborative activities. Students were mostly busy regulating their task performance by making plans and monitoring task progress. Consequently, they devoted less effort to content-related activities, such as exchanging task-related information. In fact, treatment group students typed less task-related informative remarks and questions. This may explain why the SS only had a partial impact on group performance.

Finally, some possible limitations of this study should be borne in mind. First, this study employed a complex, open-ended group task. More research is needed to determine whether the results of this study can be replicated using different types of group tasks. For other types of tasks, CMC may be better suited (Mennecke et al., 2000), thus decreasing the need for visualization of agreement and discussion. For example, CMC has been found

to a suitable medium for idea generation (Fjermestad, 2004). Thus, when working on idea generation tasks, communication difficulties may have a less detrimental effect, decreasing the need for visualization of agreement and discussion. On the other hand, CMC seems to be less suitable for groups that need to negotiate conflicts of interest (Mennecke et al., 2000). As Munneke et al. (2006) argued, in such situations the task of working in a CSCL environment, studying information sources, co-writing texts, and conducting online discussions, may a very difficult one, thereby increasing the need for a tool such as the SS. Future research should explore the effects of the SS for different types of task (e.g., idea generation versus negotiation).

Secondly, this study did not take into account the influence of individual or group-level factors on the effects of the SS. For example, in groups composed of students who were familiar with each other, the effects of the SS may have been smaller since group member familiarity facilitates online communication (Adams, Roch, & Ayman, 2005). Thus, future research should investigate the possible differential effects of visualizing agreement and discussion for different types of groups (e.g., familiar versus non-familiar).

In sum, the present study found several positive effects of visualizing agreement and discussion during online collaboration. At first glance, the effects of the SS seem to concentrate on students' perceptions (e.g., treatment group students perceived their collaboration to be more positive). However, the present study also found some effects of the SS on students' behavior: treatment group students collaborated somewhat differently and performed better on some parts of the inquiry group task. Thus, the effects of the SS seem to extend beyond merely influencing students' perceptions.

Acknowledgements

This study is part of the Computerized Representation of Coordination in Collaborative Learning (CRoCiCL) project. This project is funded by NWO, the Netherlands Organisation for Scientific Research under project number 411-02-121. Furthermore, the authors thank Marcel Broeken and Jos Jaspers for their technical assistance.

References

Adams, S. J., Roch, S. G., & Ayman, R. (2005). Communication medium and member familiarity: The effects on decision time, accuracy, and satisfaction. *Small Group Research, 36*, 321–353.

Adrianson, L., & Hjelmquist, E. (1991). Group processes in face-to-face and computer mediated communication. *Behaviour and Information Technology, 10*, 281–296.

Cavanaugh, C. S. (2001). The effectiveness of interactive distance education technologies in K-12 learning: A meta-analysis. *International Journal of Educational Telecommunications, 7*, 73–88.

Cicchetti, D. V., Lee, C., Fontana, A. F., & Dowds, B. N. (1978). A computer program for assessing specific category rater agreement for qualitative data. *Educational and Psychological Measurement, 38*, 805–813.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*, 554–571.

Dennis, A. R., Kinney, S. T., & Hung, Y. T. C. (1999). Gender differences in the effects of media richness. *Small Group Research, 30*, 405–437.

Donath, J. (2002). Visualizing online conversations. *Communications of the ACM, 45*(4), 45–49.

Erkens, G. (2005). Multiple Episode Protocol Analysis (MEPA). Version 4.10. http://edugate.fss.uu.nl/mepa/ accessed 24.10.05.

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*, 463–486.

Erkens, G., Prangsma, M., & Jaspers, J. (2006). Planning and coordinating activities in collaborative learning. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.). *Collaborative learning reasoning, and technology* (pp. 233–263). Mahwah, NJ: Lawrence Erlbaum Associates.

Fjermestad, J. (2004). An analysis of communication mode in group support systems research. *Decision Support Systems, 37*, 239–263.

Fuks, H., Pimentel, M., & Lucena, C. J. P. (2006). R-U-Typing-2-Me? Evolving a chat tool to increase understanding in learning activities. *International Journal of Computer-Supported Collaborative Learning, 1*, 117–142.

Hobman, E. V., Bordia, P., Irmer, B., & Chang, A. (2002). The expression of conflict in computer-mediated and face-to-face groups. *Small Group Research, 33*, 439–465.

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2006). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers and Education,* doi:10.1016/j.Compedu.2006.10.004.

Janssen, J., Erkens, G., Schep, A. (2006). *De invloed van sekse en groepssamenstelling tijdens computer-ondersteund samenwerkend leren [The influence of gender and group composition during computer-supported collaborative learning]*. Paper presented at the Educational Research Days 2006 (ORD 2006), Amsterdam, The Netherlands.

Jaspers, J., Broeken, M., & Erkens, G. (2005). Virtual Collaborative Research Institute (VCRI). Version 2.2. http://edugate.fss.uu.nl/~crocicl/ Accessed 00.10.05.

Jehn, K. A., & Shah, P. P. (1997). Interpersonal relationships and task performance: An examination of mediation processes in friendship and acquaintance groups. *Journal of Personality and Social Psychology, 72*, 775–790.

Kashy, D. A., & Kenny, D. A. (2000). The analysis of data from dyads and groups. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 451–477). Cambridge: Cambridge University Press.

Kenny, D. A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology, 83*, 126–137.

Kiesler, S., & Sproull, L. (1992). Group decision making and communication technology. *Organizational Behavior and Human Decision Processes, 52*, 96–123.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior, 19*, 335–353.

Kreijns, K., Kirschner, P. A., Jochems, W., & Van Buuren, H. (2004). Measuring perceived quality of social space in distributed learning groups. *Computers in Human Behavior, 20*, 607–632.

Kuhn, D., & Udell, W. (2003). The development of agrument skills. *Child Development, 74*, 1245–1260.

Lee, A., Girgensohn, A., & Zhang, J. (2004). Bowsers to support awareness and social interaction. *IEEE Computer Graphics and Applications, 24*(5), 66–75.

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction, 13*, 487–509.

Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research, 71*, 449–521.

Mennecke, B. E., Valacich, J. S., & Wheeler, B. C. (2000). The effects of media and task on user performance: A test of the task-media fit hypothesis. *Group Decision and Negotiation, 9*, 507–529.

Munneke, L., Andriessen, J., Kanselaar, G., & Kirschner, P. (2006). Supporting interactive argumentation: Influence of representational tools on discussing a wicked problem. *Computers in Human Behavior,* doi:10.1016/j.chb.2006.10.003.

Narciss, S., Proske, A., & Koerndle, H. (2006). Promoting self-regulated learning in web-based learning environments. *Computers in Human Behavior,* doi:10.1016/j.chb.2006.10.006.

O'Donnell, A. M., & O'Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review, 6*, 321–349.

Postmes, T., Spears, R., & Cihangir, S. (2001). Quality of decision making and group norms. *Journal of Personality and Social Psychology, 80*, 918–930.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education, 14*(2), 50–71.

Saavedra, R., Earley, P., & Van Dyne, L. (1993). Complex interdependence in task-performing groups. *Journal of Applied Psychology, 78*, 61–72.

Scarloss, B. A. (2002, April). Assessing complex academic performance at the group level. Paper presented at the American Education Research Association (AERA) annual meeting, New Orleans, LO.

Schellens, T., & Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers and Education, 46*, 349–370.

Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers and Education, 46*, 29–48.

Thompson, L. F., & Coovert, M. D. (2003). Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction and postdiscussion accuracy. *Group Dynamics, 7*, 135–151.

Van der Meijden, H., & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior, 21*, 831–859.

Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis, 17*, 239–261.

Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction, 9*, 493–516.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers and Education, 46*, 71–95.

Yager, S., Johnson, R. T., Johnson, D. W., & Snider, B. (1986). The impact of gro up processing on achievement in cooperative learning groups. *Journal of Social Psychology, 126*, 389–397.