DE INVLOED VAN SEKSE EN GROEPSSAMENSTELLING TIJDENS COMPUTER-ONDERSTEUND SAMENWERKEND LEREN

Jeroen Janssen
Gijsbert Erkens
Annemarie Schep

Research Centre Learning in Interaction
Universiteit Utrecht

Paper gepresenteerd tijdens de Onderwijs Research Dagen (ORD) 2006

Amsterdam, 10 – 12 mei 2006

Correspondentie richten aan:
Jeroen Janssen
Research Centre Learning in Interaction
Universiteit Utrecht
Postbus 80140
Telefoon: +31 (0)30 2534798
Fax: +31 (0)30 2532352
E-mail: j.j.h.m.janssen@fss.uu.nl

Nederlandse Organisatie voor Wetenschappelijk Onderzoek

SAMENVATTING

In dit onderzoek staan de effecten van sekse en groepssamenstelling tijdens computer-ondersteund samenwerkend leren (computer-supported collaborative learning, CSCL) centraal. Uit onderzoek naar samenwerking in face-to-face situaties blijkt dat vrouwen vaker affiliatieve taal gebruiken (bijvoorbeeld prijzen), terwijl mannen vaker assertieve taalgebruiken (bijvoorbeeld instructies geven). Doel van dit onderzoek was om de invloed van sekse en groepssamenstelling op synchrone online samenwerking gedurende een langere periode te onderzoeken. Gedurende acht lessen werken 25 mannelijke en 40 vrouwelijke leerlingen aan een praktische opdracht voor het vak geschiedenis. Uit de resultaten blijkt ten eerste dat vrouwen meer participeerden tijdens de online samenwerking dan mannen. Daarnaast gebruikten vrouwen vaker affiliatieve taal, terwijl mannen vaker assertieve taal gebruikten. Vrouwen gebruikten daarentegen vaker argumenten tijdens de online discussies. Ten derde blijkt uit een analyse van de samenwerkingsactiviteiten dat vrouwen meer tijd besteden aan het bewaren en aangeven van wederzijds begrip, terwijl mannen meer uitspraken deden die een negatieve invloed hebben op het groepsklimaat. Ten vierde bleken mannen en vrouwen niet te verschillen in hun oordeel over de samenwerking; die was niet positiever of negatiever. Er werden slechts enkele marginale effecten van groepssamenstelling gevonden. Dit onderzoek laat zien dat zich ook tijdens online samenwerking sekseverschillen voor kunnen doen.

Over the last decades, several advances have been made in information and communication technology. Nowadays, e-mail, real-time chat, file sharing, and instant messaging are being used by more and more people. These developments in ICT have also reached teacher's classrooms. In schools, teachers and students are increasingly turning to ICT to facilitate learning in various subjects (Lou, Abrami, & d'Apollonia, 2001). ICT applications, such as tutorials, simulations, and computer-mediated communication (CMC) are regarded as promising tools for education. *Computer-supported collaborative learning* (CSCL) is one educational application of ICT which has received considerable attention by educational researchers. The aim of CSCL is to provide students with an environment that supports and enhances collaboration between students, in order to facilitate students' learning processes (Kreijns, Kirschner, & Jochems, 2003). When using CSCL environments, students usually communicate with group members using CMC (e.g., discussion forums, chat facilities). A CSCL environment tries to offer tools that facilitate sharing of information and ideas, as well as the distribution of expertise among group members (Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003).

CSCL has been regarded as a potential tool for education for several reasons. Firstly, CSCL seems to fit well with current conceptions of learning as a social process during which the learning actively constructs knowledge (Schellens & Valcke, 2006). Secondly, research has demonstrated positive effects of using ICT in education (e.g., Fletcher-Flinn & Gravatt, 1995). When applied appropriately, ICT is found to have positive benefits for students' learning processes. Furthermore, research has also demonstrated positive effects of using collaborative learning. When students work together in small groups, they perform better compared to students working individually (Johnson & Johnson, 1999). Thus, CSCL is seen as a promising combination of ICT *and* collaborative learning. Indeed, a meta-analysis by Lou, Abrami, and d'Apollonia (2001) demonstrated that combining small group learning with ICT was more effective than combining individual learning with ICT. Finally, it has been argued that, due to the medium of communication used during CSCL (namely, CMC), traditional differences between men and women's communicative behavior (i.e., men dominating women during conversation) would be diminished (Postmes & Spears, 2002). However, this hypothesis has received mixed support. Some studies indicate an equalizing effect of using CMC instead of face-to-face (FTF) conversation (e.g., Fjermestad, 2004). Others demonstrate that, in some situations, men also dominate computer-mediated conversation (e.g., Adrianson, 2001). Thus, the expectation that CMC, through certain characteristics of the medium, would equalize gender differences is not as straightforward as it seems.

This paper addresses the influence of gender and gender group composition during CMC. In the next section, some of the most relevant differences between male and female group behavior will be described. In addition, two theoretical approaches used to explain differences in male and female collaborative behavior will be explained briefly. The subsequent section relates these theoretical approaches to male and female behavior during CMC and discusses relevant empirical findings.

EXPLAINING GENDER DIFFERENCES DURING PEER INTERACTIONS

Gender differences during FTF peer interactions have been studied extensively. In their meta-analysis of gender differences during FTF interactions, Leaper and Smith (2004) investigated three aspects of conversation that are expected to be influenced by gender:

talkativeness, affiliative language, and assertive language. First, *talkativeness*[1] refers to the amount of language spoken by a participant during conversation (e.g., the number of utterances spoken or the duration of talking). In general, women are expected to participate more than men during conversation. However, Leaper and Smith's meta-analysis demonstrated no significant differences between women and men with respect to talkativeness during peer interaction. Women participate as much as men do during peer interaction. Second, *affiliative language* is defined as language used by participants to establish or maintain socioemotional connections with group members. Examples of affiliative language are indicating agreement with group members and giving support or praise. As confirmed by Leaper and Smith's meta-analysis, women use more affiliative language during interactions with peers. Finally, language that is used to influence others and to attain task-oriented goals is referred to as *assertive language*. Examples of this type of language are criticism or disapproval, directive statements and indications of disagreement. This type of language is generally associated with male behavior. Indeed, Leaper and Smith's meta-analyses showed that men use assertive language more often during peer interactions than women did.

In conclusion, it seems the traditional stereotype that women focus more on socioemotional activities during interaction and men focus more on task-oriented and assertive activities is at least partially confirmed by empirical research. Several theoretical approaches explain these differences (Ridgeway, 2001). Two of the most prominent approaches will be described in short below.

S*ocial role theory* uses the gender division of labor that characterizes most western societies to explain gender differences during peer interactions (LePine, Hollenbeck, Ilgen, Colquitt, & Ellis, 2002; Ridgeway, 2001). Social role theory argues that because men and women have traditionally been assigned different roles in society, stereotypes have developed that associate affiliative behavior with women and assertive behavior with men. When men and women are in a mixed sex group or when they work on a group task that is associated with one gender, these gender stereotypes become salient and influence behavior accordingly. Furthermore, because men and women are exposed to different types of activities and roles, they develop different skills. Therefore, men and women may even behave differently in situations where stereotypes are not salient (Ridgeway, 2001).

According to *expectation states theory*, the relative status of an individual may determine how he or she behaves during collaboration (Chizhik, 1999; Dembo & McAuliffe, 1987; Kiesler & Sproull, 1992). Group members have preconceived expectations about their own and each other's task-relevant abilities (Thomas-Hunt & Phillips, 2004). Gender is one such characteristic that influences status and expectations. Generally, men are perceived to be more competent, in addition to possessing abilities which women lack. This explains why men may be more dominant and assertive during collaboration and why men's ideas may be accepted more easily compared to women's ideas (Chizhik, 1999; Ridgeway, 2001). Research in this tradition has shown that during group work, women are perceived to posses less expertise. Furthermore, being perceived an expert is found to be relatively positive for men, but not for women (Thomas-Hunt & Phillips, 2004).

[1] In studies on gender differences during online conversation and collaboration, the term participation is used more often. Therefore, the term *participation* will be used hereafter, instead of talkativeness.

GENDER DIFFERENCES DURING COMPUTER-MEDIATED COMMUNICATION

The previous section described differences in men and women's behavior during peer interaction and how these differences may be explained. This section focuses on gender differences during CMC.

Since gender potentially influences all types of interaction (Ridgeway, 2001), differences between male and female behavior during CMC have also attracted the attention of researchers. According to social role and expectation states theory, it may be hypothesized that during CMC, gender differences will continue to surface because of men and women's different roles, abilities, and status expectations.

In contrast, some researches have suggested that CMC may have an equalizing effect on collaboration. According these researchers, CMC reduces the amount of information and cues that participants can use to determine group member's expertise and status (Dennis, Kinney, & Hung, 1999; Kiesler & Sproull, 1992). In some cases, participants are even anonymous, thus their gender is unknown. In conclusion, it is expected that during CMC men and women will have similar levels of influence and will display similar behaviors.

Finally, a last group of researchers has argued that CMC may even place women at a greater disadvantage compared to FTF interaction. For example, because CMC is less media rich (Daft & Lengel, 1986), nonverbal cues (e.g., intonation of speech, gestures) are unavailable. Because women have a greater ability to interpret and transmit nonverbal communication, CMC may suit their style of conversation less well (Dennis et al., 1999). Furthermore, because status characteristics are less salient during CMC, participant may rely more on gender stereotypes and expectations during online collaboration. Thus, differences between men and women may even be enhanced during CMC (Postmes & Spears, 2002).

Previous studies investigating the effects of gender on *participation* during CMC have provided mixed results. As described above, some studies find that men and women participate equally during CMC (e.g., Dubrovsky, Kiesler, & Sethna, 1991; Fjermestad, 2004; Kiesler & Sproull, 1992; Postmes & Spears, 2002; Straus, 1996), whereas others report men (e.g., Herring, 1993; Wolfe, 1999) or even women (Adrianson, 2001) to dominate CMC discussions. Furthermore, gender group composition seems to affect participation during CMC. In some studies, groups composed of only males participated less intensively compared to female only or mixed groups (Savicki, Kelley, & Ammon, 2002; Savicki, Kelley, & Lingenfelter, 1996b), whereas in another study participation was highest in mixed groups (Savicki, Kelley, & Lingenfelter, 1996a).

Several studies have also described effects of gender on participants' use of *affiliative language* during CMC. For example, Wolfe (1999) found that women indicated agreement more often than men did, whereas Postmes and Spears (2002) observed that women used a more dependent communication style by asking more questions. In addition, gender group composition seems to affect use of affiliative language during CMC. In their study, Savicki et al. (1996b) found that groups composed of all females tend to use more self-disclosure and coalition language, compared to mixed or male only groups. Finally, Van der Meijden (2005) demonstrated that female only groups used significantly more greetings than did groups of mixed gender composition.

Finally, gender also influences participants' use of *assertive language* during CMC. Postmes and Spears (2002) examined male and female behavior during CMC and concluded that men were more dominant during online discussion. Additionally, Herring (1993) noted that men behave more assertively and are more likely to display their knowledge. Furthermore, the studies by Savicki et al. (1996a; 1996b) demonstrated that male only groups are more likely to attack arguments of group members and to use abusive language.

From the research described above, it can be concluded that gender and gender group composition both affect how participants behave during CMC. Some of the results found in CMC contexts mirror the results found in FTF contexts. The current study seeks to extend this body of research in several ways. First, this study was conducted in a secondary education context. Most studies investigating gender differences during CMC are carried out in organizational or university settings (e.g., Adrianson, 2001; Savicki et al., 1996b). Second, this study involved secondary education students collaborating in a CSCL environment over a longer time span (eight lessons over a period of four weeks). During many studies on gender differences a much shorter time span is used (e.g., Dennis et al., 1999; Van der Meijden, 2005); sometimes as short as 10 to 12 minutes (e.g., Postmes & Spears, 2002). During such a short period of time, CMC groups may not be able to develop adequate levels of interpersonal trust (Wilson, Straus, & McEvily, 2006), which may lead to negative behavior, influencing the effects of gender and gender group composition. Giving group members more time to develop a sound social climate may yield different results compared to giving students little time to develop a relationship. Finally, students used synchronous CMC (chat) to communicate and collaborate during this study. A considerable body of research on gender differences focuses on asynchronous CMC, such as e-mail or discussion boards (e.g., Savicki et al., 1996a; Savicki et al., 1996b). Synchronous CMC differs from asynchronous CMC, however. For example, feedback is more immediate when using synchronous CMC (Daft & Lengel, 1986). Thus, the impact of gender and gender group composition may be different in synchronous CMC environments.

RESEARCH QUESTIONS

As described above, it is expected that gender and gender group composition will affect students' participation in a synchronous CSCL environment. Furthermore, it is expected that female students and groups dominated by women will use more affiliative language, whereas male students and groups dominated by men will use more assertive language. In addition, it expected that male students will engage in different collaborative activities than female students. Subsequently, it is expected that gender and gender group composition will affect students' perceptions of their collaboration. Due to their more harmonious styles of online conversation and collaboration, female students are expected to perceive their collaboration more positively. Finally, it is expected that gender and gender group composition may have an impact on group performance. In sum, the following research questions are considered:

1. Do gender and gender group composition affect students' participation during online collaboration?
2. Do gender and gender group composition affect students' use of affiliative language during online collaboration?
3. Do gender and gender group composition affect students' use of assertive language during online collaboration?
4. Do gender and gender group composition affect students' collaborative activities during online collaboration?
5. Do gender and gender group composition affect students' perception of collaboration during online collaboration?
6. Does gender group composition affect groups' performance during online collaboration?

Participants

In total, 65 students (40 female, 25 male) from a Dutch school for secondary education participated in this study. Students were enrolled in the second stage of the pre-university track. Mean age of the participating students was 16.3 years (*SD* = .58). During the experiment, students collaborated in a CSCL environment in groups of two, three or four. In total, students were randomly assigned to 21 groups by the researchers (16 groups were female dominated, 5 groups were male dominated).

Task and materials

CSCL environment: VCRI

Students collaborated in a CSCL environment named *Virtual Collaborative Research Institute* (VCRI, Jaspers, Broeken, & Erkens, 2004). The VCRI program is a groupware program designed to facilitate collaborative learning. The program offers students several tools that they can use while they are collaborating on research projects or inquiry group tasks. For example, students can read the description of the group task and search for relevant historical information using the *Sources* tool. This information can be communicated and shared with group members, using the synchronous *Chat* tool. To write research reports and argumentative texts or essays, students can use the *Co-Writer*. The *Co-Writer* is a shared text-processor, which can be used by students to work simultaneously on the same text. Figure 1 shows a screenshot of the VCRI program. Other tools, not shown in Figure 1, include for example the *Logbook*, which students can use to record which activities they have carried out, and the *Diagrammer*, which can be used to construct argumentative diagrams.

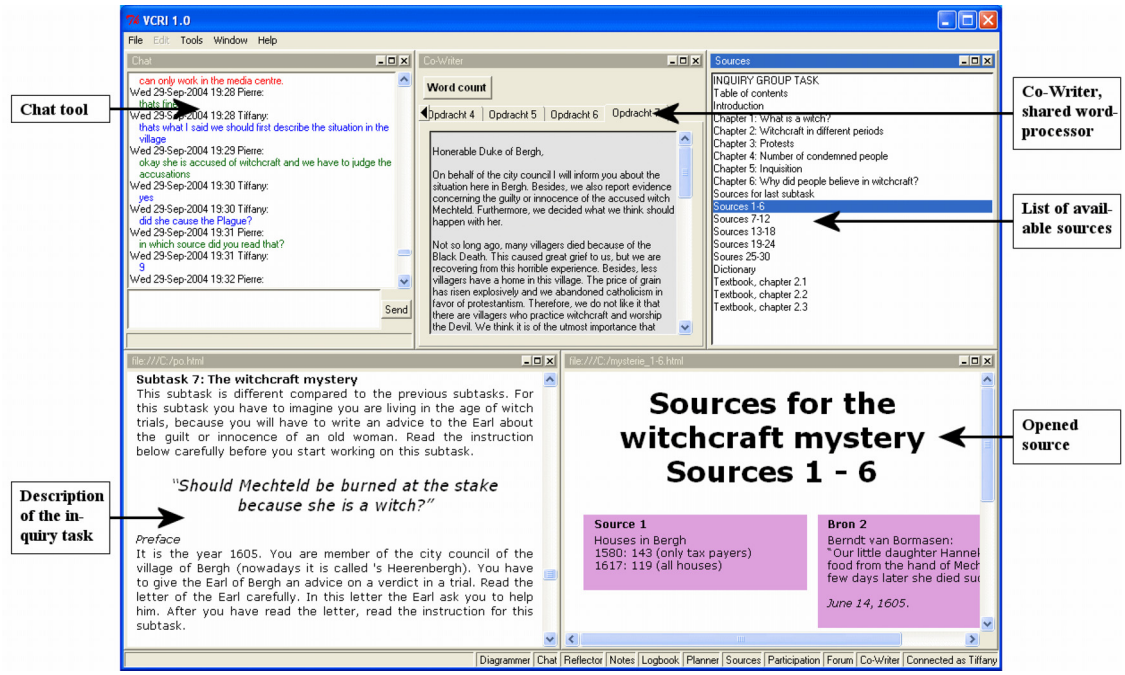

Figure 1 Screenshot of the VCRI program, detailing some of its tools.

Inquiry group task

The participating students collaborated on a historical inquiry group task. Inquiry tasks are an important part of the curriculum in the Dutch upper secondary levels. Subject of the task was witchcraft and the persecution of witches. The groups had to use different historical and

(more) contemporary sources to answer questions and co-author argumentative texts. Approximately 40 sources from textbooks and the Internet were available to the students through the *Sources* tool. Students were instructed to use the VCRI program to communicate with group members. Students were told they had eight lessons to complete the inquiry task.

The introduction of the task stressed the importance of working together as a group on the subtasks, and pointed out that group members were themselves responsible for the successful completion of the task. To successfully complete the inquiry group task, all group members had to participate during the group process. Positive interdependence and individual accountability were incorporated in the group task, thus making high levels of collaboration necessary in order to successfully complete the group task (Johnson & Johnson, 1999).

Procedure

The participating students worked in their groups on the inquiry group task during four weeks. During the first lesson, the task and the most important features of the CSCL environment were introduced to the students. After the first lesson, another seven history lessons were devoted to the inquiry group task. The teachers were standby to answer task-related questions, while the experimenters were standby to answer technical questions or to solve any technical problems. Students were allowed to work on the inquiry task during free periods. For example, students could work in the media center when they had spare time in their timetable. However, students could only access the CSCL-environment from school, not from their homes. After eight lessons the students were required to hand in their final versions of the group task. After the last lesson, a questionnaire was administered to the students in order to assess students' perceptions of their online collaboration.

Units of analysis

To answer the first four research questions, concerning the effects of gender and gender group composition on participation, affiliative language, assertive language, and collaborative activities, an appropriate unit of analysis had to be chosen. Using entire chat messages as the unit of analysis would be an intuitive choice, but during online collaboration some students only send one sentence per chat message, while others type several sentences that combine multiple clauses (Howell-Richardson & Mellar, 1996). Furthermore, even within in a single sentence, multiple concepts, ideas and statements may be expressed (Strijbos, Martens, Prins, & Jochems, 2006). For example, a compound sentence, such as "She should be burned at the stake, because she is a witch", contains two clauses: "She should be burned at the stake" and "She is a witch". Thus, it may be necessary to segment a message into even smaller parts that can be meaningful in their selves. Therefore, the chat messages sent by the participating students were segmented into *dialogue acts* (Erkens, 2004; Erkens, Jaspers, Prangsma, & Kanselaar, 2005; Janssen, Erkens, Kanselaar, & Jaspers, in press). Dialogue acts indicate the communicative function of a message (responding, informing, argumentation, commanding, or eliciting). 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.

To analyze the chat protocols, the computer program *Multiple Episode Protocol Analysis* (MEPA) was used (Erkens, 2005). A *segmentation filter* was used to segment chat messages into smaller, meaningful parts. A filter is a program, which can be specified and used in the MEPA program 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 connectives (e.g., "and", "but", "because") are used by the filter to segment messages

into dialogue acts. For example, the filter would segment abovementioned sentence "She should be burned at the stake, because she is a witch", into two dialogue acts: "She should be burned at the stake" and "because she is a witch". Using filters dramatically speeds up the segmentation process, and ensures segmentation rules are applied consistently. Dialogue acts were subsequently used as the unit of analysis for all analyses concerning participation, affiliative language, assertive language, and collaborative activities (see research questions one to four).

Measures

Participation

To determine participation, dialogue acts were used as the unit of analysis, instead of, for example, number of messages or words typed. Furthermore, a distinction was made between *long* (>5 words) and *short dialogue acts* (<=5 words). During online collaboration, short dialogue acts are used mostly for supporting, confirming and back channeling (e.g., "okay", "I agree"). Although these behaviors serve an important function during collaboration, they are considered less important for the development of online conversation and collaboration (Bonito, 2000). In contrast, group members mostly use long dialogue acts for transfer of content or information and for regulation of task and group processes. Thus, long dialogue acts are considered more useful for the development of conversation and collaboration, and are therefore considered to be more substantive contributions (Bonito, 2000).

Affiliative language

After the segmentation of chat messages into dialogue acts, it was examined whether the dialogue acts contained affiliative language. Affiliative language refers to dialogue acts that are used to establish or maintain a connection with others (Leaper & Smith, 2004). Three measures of affiliative language were distinguished: *agreement*, *acknowledgement*, and *praise*. All three are considered measures of affiliative language by Leaper and Smith (2004).

Coding of agreement, acknowledgment, and praise was done using the *Dialogue Act Coding (DAC) filter* within the aforementioned MEPA program. This filter uses over 1300 rules based on discourse markers to determine the communicative function of a dialogue act. Discourse markers are characteristic words signaling the communicative function of a phrase in conversation in natural language (Schiffrin, 1987). In total, five main categories of communicative functions are distinguished: argumentative (indicating a line of reasoning or argumentation), responsive (answers to questions and proposals), informative (transfer of information), elicitative (questions or proposals requiring a response), and imperative (commands). Each main category consists of several subcategories. For example, the elicitative category consists of verification questions, open questions, set questions, and proposals for action. In total, 29 different dialogue acts are distinguished.

Dialogue acts were coded as agreement when they contained acceptances of information and indications of understanding (e.g., "I agree"). Acknowledgments were minimal responses acknowledging group members' contributions. Finally, dialogue acts were considered praise if they contained positive evaluations of ideas and contributions (e.g., "good idea", "I like what you wrote"). Reliability of this automatic coding procedure is high (over 90% of all dialogue acts coded correctly, see Erkens et al., 2005). The automatic coding procedure ensures consistent application of the coding rules.

Assertive language

Assertive language refers to dialogue acts that are used to influence others (Leaper & Smith, 2004). Four measures of assertive language were distinguished: *directives, disagreement, argumentative language*, and *negative language*. Directives, disagreement, and negative

language were considered measures of assertive language by Leaper and Smith (2004). Argumentative language was included since argumentation is clearly used to influence and convince others (Erkens et al., 2005) and also facilitates knowledge construction (Kuhn & Udell, 2003) and collaborative problem solving (Weinberger & Fischer, 2006).

Coding of directives, disagreement and argumentative language was done using the DAC filter described above (see Affiliative language section). Dialogue acts were coded as directives if they contained commands directed at the other group members (e.g., "you have to type in the answer", "get back to work"). Disagreements contained rejections of statements and information (e.g., "I don't think so", "I don't agree"). Argumentative language indicates a line of reasoning and is used to convince group members, for example by giving counter arguments (e.g., "but that's not important here") or reasons (e.g., "because it's better to do it together"). As described above, reliability of this automated coding procedure was high.

Online collaboration is usually characterized by higher amounts of offensive behavior as opposed to FTF collaboration (Straus, 1996). This hostile expression of (negative) emotion and feelings is also known as "flaming" (Kiesler & Sproull, 1992). Furthermore, negative or offensive behavior has been found to negatively affect trust development and online collaboration (Savicki, Kelley, & Oesterreich, 1998; Wilson et al., 2006). In order to code negative language, a *negative language filter* was developed. Using this filter, dialogue acts were coded as negative language when they contained inflammatory comments or use of rude or offensive words (e.g., "shut up idiot", "this assignment sucks"). In order to establish to reliability of the automatic coding of negative language, over 5000 dialogue acts were coded by one of the researchers. The results of this coding were compared to the results of the automatic coding procedure. An overall Cohen's Kappa of .90 was found. Thus, the automatic coding of negative language was sufficiently reliable.

Collaborative activities

To answer the fourth research question, regarding the influence of gender and gender group composition on students' collaborative activities, a coding scheme was developed (Janssen et al., in press). The aim of this coding scheme was to provide insight into the task- and group-related processes taking place between students while working on the inquiry group task. This section describes the codes of the coding scheme and its interobserver reliability.

Description of the coding scheme. Different types of activities are necessary to successfully complete a group task. These four types of activities are reflected by the four different dimensions of the coding scheme. Each dimension contains two or more coding categories. In total, the scheme consists of 19 categories.

First, groups have a production function (McGrath, 1991), meaning group members have to *perform task-related activities*, such as sharing and discussing task-related information (Fischer, Bruhn, Gräsel, & Mandl, 2002; Jehn & Shah, 1997), and asking task-related questions (King, 1994). This dimension contained two categories pertaining to the discussion of relevant task-related information: *exchanging task-related information* and *asking task-related questions*.

The second dimension referred to *regulation and coordination of task-related activities*. During collaboration, group members need to coordinate and organize their activities to determine a common course of action (Erkens, 2004). Thus, metacognitive activities that regulate task performance, such as making plans and monitoring task progress, are considered important to successful group performance (De Jong, Kollöffel, Van der Meijden, Kleine Staarman, & Janssen, 2005; Van Meter & Stevens, 2000). This coding dimension contained four categories. First, *planning* involved discussion of strategies necessary to complete the task, choice of appropriate strategies, and delegation of task responsibilities. Second, *monitoring* 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 *positive* or *negative*.

Performance of social activities was the third dimension of the coding scheme. Group members have to attend to the social and emotional element of collaboration to successfully complete a group task (Kreijns, 2004; Kumpulainen & Mutanen, 1999; Rourke, Anderson, Garrison, & Archer, 1999). Behaviors such as offering positive comments and praising group members contribute to a sound social space and a positive group atmosphere (Kreijns, 2004), which may increase group members' efforts to complete the group task (Jehn & Shah, 1997; Rourke et al., 1999). On the other hand, behaviors such as swearing or displaying extreme or negative emotions may have a negative impact on group cohesion and trust development (Wilson et al., 2006). In all, this dimension contained five coding categories. First, *greetings* were included, since they contribute positively to group atmosphere and a feeling of social presence (Rourke et al., 1999). Second, *social support* remarks referred to comments that contributed positively to group atmosphere, such as exchanging positive comments, displaying positive emotions, and disclosure of personal information. Third, *social resistance* remarks referred to behaviors that had a negative impact on group atmosphere, such as insulting group members, flaming, and displaying negative emotions. Fourth, *shared understanding* referred to confirmations and indications of agreement, which serve to reach and maintain joint understanding. Similarly, *loss of shared understanding* referred to denials, disagreements, and expressions of incomprehension.

The fourth dimension referred to *regulation and coordination of social activities*. Like task-related activities, social activities require regulation as well (Erkens, 2004). During collaboration, group members are interdependent, and therefore they have to discuss collaboration strategies, monitor their collaboration process, and evaluate and reflect on the manner in which they collaborated. This is sometimes referred to as group processing (Johnson, Johnson, & Stanne, 1990; Webb & Palincsar, 1996). In total, this dimension contained four categories. First, *planning* involved discussion of collaboration strategies, such as helping each other, or proposals to work together on certain tasks. Second, *monitoring* referred to the exchange of information that could be used to monitor group processes. Last, *evaluation* involved appraisal and discussion of group processes and collaboration, which could be *positive* or *negative*.

Finally, statements that did not fit into any of the previously mentioned categories were coded as *Other*. These codes mostly referred to nonsense and off-task remarks.

Interobserver reliability. Two researchers were involved in the development and refinement of the coding scheme. In order to examine interrater agreement 601 dialogue acts were coded independently by two raters. An overall Cohen's Kappa of .86 was found, a satisfactory result. Category Kappa's (Cicchetti, Lee, Fontana, & Dowds, 1978) ranged from .69 to 1.00.

Student perception of collaboration

To investigate whether gender and gender group composition affected students' perceptions of their online collaboration, the posttest questionnaire contained two scales. The first scale consisted of six items and addressed *positive group behavior*, such as equal participation of group members, and helping group members (Webb, 1995). Sample items included: "We collaborated well during the group task", and "We helped each other during collaboration". Cronbach's alpha was found to be .89 for this scale. Thus, students' responses to the individual items were averaged into a single score for "positive group behavior". Higher scores on this measure indicate more perceived occurrences of positive group behavior.

The second scale also consisted of six items. These items addressed occurrences of *negative group behavior* such as conflicts and free riding behavior (O'Donnell & O'Kelly, 1994).

Sample items of this scale are: "I had to do most of the work during collaboration" and "There were conflicts in our group". Students' responses to the six items were averaged to obtain a score for "negative group behavior". Cronbach's alpha for this scale was found to be .75. Higher scores on this variable indicate more perceived occurrences of negative group behavior.

Group performance

The quality of the texts written by the groups for the seven subtasks of the inquiry group task was examined in order to answer the final research question regarding the influence of gender group composition on group performance. For this purpose an assessment form was developed. Using this assessment form, three quality aspects were assessed for each subtask. First, *use of sources* referred to the manner how students incorporated the historical sources that were available into their texts. This quality aspect contained two items: completeness of sources used in the written text, and copy-pasting of information from sources to the written texts. Second, c*ontent and argumentation* referred to way students formulated their answers and supported these with arguments. The inquiry group task consisted of seven subtasks and since each subtask addressed different aspects of the subject, the content and argumentation aspect was formulated differently for each subtask. The amount of items that addressed content and argumentation also differed for each subtask, since some subtasks were more extensive than others. *Text construction and language* was the third quality aspect that was examined. This aspect referred to the manner how students' written text had an adequate text construction and contained correct language. This quality aspect consisted of three items: text construction, text structure, and correctness of language.

All items of the assessment form were answered on a 3-point scale, with 0 indicating poor quality and 2 good quality. To determine whether students directly copy-pasted information from the sources to their text, the program *WCopyFind* (http://plagiarism.phys.virginia.edu/Wsoftware.html) was used. This program compares texts to the available sources and determines how many percent of the written text is copy-pasted directly from the sources. This percentage was used to determine whether the group received 0, 1, or 2 points. Groups that copy-pasted less than 34% of their text from the sources, received 2 points; groups that pasted more than 66% of their text received no points.

In total, groups could receive up to 12 points for subtasks one, five and seven, 14 points for subtasks two, four and six, and 18 points for subtask three. Thus, in total 96 points could be earned.

To check the objectivity of the scoring procedure, two researchers independently scored a number of texts. Each researcher filled out the assessment form for 8 to 10 groups. For the three above mentioned quality aspects, (a) use of sources, (b) content and argumentation, and (c) text construction and language, interrater agreement reached 88.5% (Cohen's κ = .79), 77.5% (κ = .70), and 75.0% (κ = .51) respectively. Overall, the agreement percentage was 75%, while Cohen's κ was .69.

In order to create a total score for group performance, the scores for the seven subtasks were summed. A Cronbach's alpha of .80 indicated this scoring procedure was internally consistent.

RESULTS

Participation

Table 1 shows the differences between male and female students for measures of participation. Similarly, Table 2 shows the results for differences between male dominated and female dominated groups. Multilevel analysis was used to examine the effects of gender

and gender group composition on participation. First, a multilevel model was constructed which included the total amount of time students were online. This was done to account for the fact that some students were online longer than other students (e.g., due to illness). By including time online in the multilevel analyses, the effects of gender and gender group composition could be investigated, independent of the effect of the time students were online. Second, the effect of gender was examined. For the effect of gender to be significant, its *t* value had to be significant. Furthermore, the model including gender also had to be a better model, indicated by a significant χ^2 value. Third, gender group composition was added to the model. Similarly, the effect of gender group composition was only considered to be significant when both the *t* and χ^2 value were significant.

Unsurprisingly, the multilevel analyses indicated time online was a significant predictor for all three measures of participation. Thus, when students were online longer, they participated more during online collaboration.

Table 1 Means and standard deviations for participation, affiliative language and assertive language, and multilevel analyses of the effect of gender.

	Male students (*N* =25)		Female students (*N* = 40)			
	M	SD	M	SD	*t*	χ^2
Participation						
Total dialogue acts	241.92	106.52	322.40	162.37	2.74**	7.01**
Long dialogue acts (>5 words)	82.48	46.93	122.53	72.38	3.08**	8.65**
Short dialogue acts (<=5 words)	159.44	69.29	199.88	99.88	2.09*	4.23*
Affiliative language						
Agreement	36.32	22.52	60.75	33.76	2.52**	5.85**
Acknowledgement	2.28	2.37	3.50	3.61	.72	.51
Praise	2.92	2.45	4.83	4.35	1.10	1.20
Assertive language						
Directives	9.08	8.39	7.80	5.06	-2.45**	5.69*
Disagreement	4.64	2.89	5.68	3.92	-.33	.08
Argumentative lang.	28.96	15.48	55.25	38.98	2.57**	6.29*
Negative language	3.12	3.40	1.70	2.04	-3.26**	9.27**

* $p < .05$. ** $p < .01$.

Furthermore, female students scored higher on all three measures of participation than did male students (see Table 1). The multilevel analyses indicated an effect of gender on all three measures of participation. Female students typed significantly more dialogue acts than male students did during online collaboration independent of the total amount of time they were online, $t(62) = 2.74$, $p = .00$. Furthermore, female students typed more long dialogue acts, $t(62) = 3.08$, $p = .00$, and more short dialogue acts, $t(62) = 2.09$, $p = .02$. Although there were differences between male and female dominated groups (on average, participation was higher in female dominated groups for all three participation measures), the effects of gender group composition were not significant, $t(17) = -.12$, $p = .45$; $t(17) = -.35$, $p = .37$; $t(17) = .08$, $p = .47$, respectively.

Table 2 Means and standard deviations for participation, affiliative language and assertive language, and multilevel analyses of the effect of gender group composition.

	Male dominated groups (*N* =5)		Female dominated groups (*N* = 16)			
	M	*SD*	*M*	*SD*	*t*	χ^2
Participation						
Total dialogue acts	738.80	256.33	939.06	427.66	-.12	.01
Long dialogue acts (>5 words)	290.40	130.84	344.44	186.05	-.35	.13
Short dialogue acts (<=5 words)	493.40	134.64	594.63	254.48	-.12	.01
Affiliative language						
Agreement	128.20	72.90	168.56	82.37	-.01	.00
Acknowledgement	8.40	4.62	9.69	5.06	-.35	.12
Praise	7.60	4.04	14.25	10.06	1.19	1.35
Assertive language						
Directives	25.60	15.37	25.69	12.64	.17	.03
Disagreement	11.00	4.18	18.00	9.47	1.96*	3.62
Argumentative lang.	96.60	43.63	153.19	86.40	1.00	.98
Negative language	9.00	4.74	6.31	4.78	-.46	.21

Affiliative language

Tables 1 and 2 also show the results for differences between male and female and female students and between male and female dominated groups concerning use of affiliative language during online collaboration. Again, multilevel analysis was used to examine the effects of gender and gender group composition. First, a model was constructed which included the number of dialogue acts typed by a student. This was done because female students typed significantly more dialogue acts than female students (see the previous section). By including dialogue acts, the effects of gender and gender group composition could be investigated independent of the total number of dialogue acts sent. In the second and third step, gender and gender group composition were added to the model.

As expected, multilevel analyses indicated number of dialogue acts typed was a significant predictor for all three measures of affiliative language. Students who typed more dialogue acts during online collaboration also used more affiliative language.

Furthermore, Table 1 shows that, on average, female students scored higher on all three measures of affiliative language: agreement, acknowledgement, and praise. However, only the effect of gender on indications of agreement was found to be statistically significant, $t(62) = 2.52$, $p = .01$. This indicates that, during online collaboration, female students indicated agreement significantly more than male students independent of the number of dialogue acts they typed. No effects of gender on acknowledgement and praise were found, $t(62) = .72$, $p = .27$, and $t(62) = 1.10$, $p = .14$, respectively. Table 2 also indicates that female dominated groups, on average, used more affiliative language than did male dominated groups. The effects of gender group composition were not significant, however, $t(17) = -.01$, $p = .50$; $t(17) = -.35$, $p = .37$; $t(17) = 1.19$, $p = .13$, respectively.

Assertive language

Additionally, Tables 1 and 2 show the results for differences between male and female and between male and female dominated groups concerning use of assertive language during online collaboration. The multilevel analyses were carried out in the same fashion as for affiliative language in the previous section. Again, number of dialogue acts typed was a

significant predictor for all three measures of assertive language. This indicates that students, who participated more, also used more assertive language.

Table 1 shows several differences between male and female students. The results of the multilevel analyses show female students to use significantly less directives independent of the number of dialogue acts typed, since the effect of gender was statistically significant, $t(62) = -2.45$, $p = .01$. In addition, female students used significantly less negative language independent of the number of dialogue acts typed, as indicated by the significantly negative effect of gender, $t(62) = -3.26$, $p = .00$. In contrast, a significantly positive effect of gender was found for argumentative language, $t(62) = 2.57$, $p = .01$. This indicates that female students used significantly more argumentative language than their male counterparts, independent of the number of dialogue acts typed. The multilevel analyses did not reveal an effect of gender on indications of disagreement during online collaboration, $t(62) = -.33$, $p = .37$. In addition, Table 2 shows differences between male and female dominated groups. Surprisingly, the multilevel analyses indicated an effect of group composition on disagreement, $t(17) = 1.96$, $p = .03$. This indicates that when participation is taken into account, female dominated groups disagree more than do male dominated groups. However, since the associated χ^2 was only marginally significant ($p = .06$), this result should be interpreted with caution. Gender group composition did not affect the other measures of assertive language, $t(17) = .20$, $p = .43$; $t(17) = 1.00$, $p = .17$; $t(17) = -.46$, $p = .33$, respectively.

Table 3 Means and standard deviations for collaborative activities, and multilevel analyses of the effect of gender.

	Male students (N =25)		Female students (N = 40)			
	M	*SD*	*M*	*SD*	*t*	χ^2
Performing task-related activities						
- Info exchange	24.28	21.95	37.65	49.77	-.15	.02
- Asking questions	6.56	6.77	7.63	8.70	-.85	.72
Coordinating task-related activities						
- Planning	41.88	21.82	61.55	35.10	1.16	1.33
- Monitoring	29.64	15.92	40.83	21.79	.73	.53
- Positive evaluations	2.88	2.98	5.15	3.83	1.68*	2.66
- Negative evaluations	4.28	4.83	6.05	4.48	.22	.05
Social activities						
- Greetings	10.40	9.06	10.28	5.49	-1.93*	3.39
- Social support	17.04	13.46	22.40	16.20	.09	.01
- Social resistance	4.84	4.49	3.87	3.97	-1.94*	3.70*
- Mutual understanding	47.16	26.90	75.23	41.68	2.04*	4.03*
- Loss of mutual understanding	9.04	5.42	9.90	5.12	-.90	.77
Coordinating social activities						
- Planning	4.72	4.62	6.40	4.35	.44	.19
- Monitoring	14.36	7.30	17.38	11.77	-.60	.35
- Positive evaluations	.44	.77	.60	.84	.17	.03
- Negative evaluations	.56	.82	.93	1.53	.67	.45
Other	5.08	6.44	3.15	3.79	-1.87*	3.41

* $p < .05$.

Collaborative activities

The effects of gender and gender group composition were examined in a similar fashion as for affiliative and assertive language. The analyses indicated number of dialogue acts was a significant predictor for all collaborative activities. For example, a student who typed more dialogue acts also typed more greetings compared to a student who typed very little dialogue acts.

Tables 3 and 4 show the differences between male and female students and between male and female dominated groups. For five types of collaborative activities, a significant effect of gender was found. First, the positive effect of gender on positive evaluations of task-related activities indicates, when participation was taken into account, female students sent more messages that contained positive appraisals of their group's task progress, $t(62) = 1.68$, $p = .05$. However, the corresponding χ^2 value did not reach statistical significance ($p = .10$). Second, a negative effect of gender on the number of greetings typed was found, $t(62) = -1.93$, $p = .03$. This indicates male students are more likely to type messages that contain greetings, even when participation is taken into account. Again, the χ^2 value was only marginally significant however ($p = .07$). Third, the negative effect of gender on social resistance indicates male students typed significantly more social resistance remarks, such as swearing, compared to female students, $t(62) = -1.94$, $p = .03$. Furthermore, female students devoted significantly more effort to reaching and maintaining mutual understanding, as indicated by the positive effect of gender, $t(62) = 2.04$, $p = .02$. Finally, the number of *Other* remarks was considerably higher for male students than for female students, and thus the effect of gender reached statistical significance, $t(62) = -1.87$, $p = .03$. This points to higher levels of off-task behavior for male students, even when the number of dialogue acts typed is taken into account. It should be noted however, that the corresponding χ^2 value was only marginally significant ($p = .06$).

Table 4 indicates two effects of gender group composition on students' collaborative activities. The positive effect of gender group composition for loss of mutual understanding, indicates that female dominated groups were more likely to express disagreement and incomprehension, $t(17) = 1.76$, $p = .05$. Second, the negative effect found for number of other remarks, indicates that male dominated groups were more off-task compared to female dominated groups, $t(17) = -1.95$, $p = .03$. It should be noted however, that in both cases the corresponding χ^2 values were only marginally significant ($p = .09$, $p = .06$, respectively). Thus, these results should be interpreted with caution.

Student perception of collaboration

To examine the effects of gender and gender group composition on students' perceptions of their collaboration, multilevel analyses were used. First, gender was added to the model, followed by gender group composition. On average, female students reported higher levels of positive group behavior ($M = 3.96$, $SD = .73$) and lower levels of negative group behavior ($M = 2.08$, $SD = .74$) than did male students ($M = 3.56$, $SD = .84$, and $M = 2.42$, $SD = .83$). The effects of gender were only marginally significant, however, $t(61) = 1.44$, $p = .08$, $t(61) = -1.48$, $p = .07$, respectively.

Similarly, female dominated groups reported higher levels of positive group behavior ($M = 3.83$, $SD = .57$) and lower levels of negative group behavior ($M = 2.17$, $SD = .61$) than male dominated groups ($M = 3.60$, $SD = .79$, and $M = 2.41$, $SD = .61$). No significant effects of gender group composition were found however, $t(17) = .45$, $p = .33$, $t(17) = -.41$, $p = .34$, respectively.

Table 4 Means and standard deviations for collaborative activities, and multilevel analyses of the effect of gender group composition.

	Male dominated groups (*N* =5)		Female dominated groups (*N* = 16)			
	M	*SD*	*M*	*SD*	*t*	χ^2
Performing task-related activities						
- Info exchange	65.20	45.26	111.69	118.49	.67	.44
- Asking questions	16.60	10.74	24.13	24.77	.80	.62
Coordinating task-related activities						
- Planning	136.20	58.23	176.75	77.70	.41	.17
- Monitoring	103.80	37.12	115.94	50.36	-.94	.89
- Positive evaluations	9.80	6.53	14.31	10.57	.28	.08
- Negative evaluations	16.00	10.15	16.81	9.29	-.83	.68
Social activities						
- Greetings	24.00	11.92	34.44	16.97	1.57	2.34
- Social support	55.20	12.24	65.38	38.15	-.09	.01
- Social resistance	18.20	2.86	11.56	9.06	-1.50	2.13
- Mutual understanding	164.20	77.85	210.44	103.71	-.12	.02
- Loss of mutual understanding	23.20	4.09	31.63	14.44	1.76*	2.94
Coordinating social activities						
- Planning	18.60	11.15	17.56	8.15	-1.19	1.34
- Monitoring	42.80	15.99	52.50	25.12	.31	.10
- Positive evaluations	1.60	1.52	1.69	1.74	-.33	.11
- Negative evaluations	2.00	1.41	2.56	3.83	-.12	.01
Other	19.40	9.40	9.75	8.64	-1.95*	3.54

* $p < .05$.

Group performance

To examine whether there were differences between male and female dominated groups regarding group performance, a *t* test for independent samples was used. No differences were found between male and female dominated groups for the total score on the inquiry group task, $M_{male\ dominated} = 65.40$, $SD_{male\ dominated} = 14.94$, $M_{female\ dominate} = 66.56$, $SD_{female\ dominated} = 9.00$, $t(4.94$, equal variances not assumed$) = -.17$, $p = .88$.

DISCUSSION AND CONCLUSION

In the present study, the effects of gender and gender group composition on students' online behavior during computer-supported collaborative learning were examined. Research in FTF and CSCL contexts has shown that male and female students behave differently during collaboration. This study sought to extend previous findings, by examining the effects of gender and gender group composition on secondary education students' online behavior in a synchronous CSCL environment over a longer time span.

The first research question was whether gender and gender group composition affected students' participation during online collaboration. The results demonstrate an effect of gender on participation. Female students participate more during online

collaboration in terms of their total number of dialogue acts typed, compared to male students. More specifically, female students type more short dialogue acts as well as long dialogue acts. The latter finding is important, since long dialogue acts are mainly used for information exchange and regulative processes, and are therefore considered very important for successful collaboration (Bonito, 2000). No effects of gender group composition on participation were found.

The results for the first research question are in contrast with the results generally found during FTF collaboration. For example, Leaper and Smith (2004) noted that female students are generally expected to participate more during collaboration, but found no participation differences between male and female students in their meta-analysis. Furthermore, the higher participation of female students in the current study is in contrast with some of the results found during online collaboration and conversation. In this case using CMC did not have an equalizing effect on participation (Kiesler & Sproull, 1992), nor did it place women at a disadvantage (Dennis et al., 1999). This may be due the communication medium (synchronous CMC) and the nature of the group task. Students worked on a more "open" task, for which there were neither set procedures nor clear right or wrong answers. It has been argued that under these high equivocality conditions, women may tend to compensate more for the lack of nonverbal communication during online collaboration (Dennis et al., 1999). This may explain the higher participation rates of female students in the present study.

The second research question was whether gender and gender group composition affected students' use of affiliative language during online collaboration. The results showed this to be partly the case. On average, female students typed more indications of agreement during online collaboration than male students when controlling for participation. In addition, the effect of gender on indications of agreement was significant. In contrast, the effects of gender on acknowledgements and giving praise were not significant. No effects of gender group composition on agreement, acknowledgement, or praise were found.

The results for the second research question seem to be partly in line with results found in FTF collaboration. For example, Leaper and Smith's (2004) meta-analysis showed that women scored higher on general measures of affiliative language than men did. Furthermore, these results are also in line with research in CMC collaboration. For instance, in studies by Adrianson (2001) and Wolfe (1999) women agreed more than men did. The same finding emerged in this study. This result is in line with social role theory (Ridgeway, 2001), because in the current study female students behaved more or less according to general gender stereotypes: female students used more affiliative language during online collaboration by typing more agreements.

The effect of gender and gender group composition on students' use of assertive language during online collaboration was the focus of the third research question. It was expected that male students would use more assertive language than female students. This hypothesis was partly confirmed. Indeed, male students used more directives and negative language (e.g., swearing, flaming) than female students: the effect of gender on directives and negative language was significant independent of student participation. However, female students used more argumentative language than male students did. This is in contrast with the abovementioned hypothesis, because argumentative language is used to influence group members, and can thus be considered a form of assertive language. The finding that female students use more argumentative language than male students is important, because it has been argued that using arguments is positively related to learning and knowledge construction (Weinberger & Fischer, 2006). No effects of gender group composition on assertive language were found. It is worth noting however, that the effect of gender group composition on indications of disagreement approached statistical

significance. In female dominated groups, students disagreed more compared to male dominated groups.

The results found for assertive language are partly in line with results found during FTF collaboration. Leaper and Smith (2004) demonstrated that male students scored higher on general measures of assertive language and gave more directives. Furthermore, the results of this study seem to be in line with previous CMC research. For example, the finding that male students used more negative language, seems to be in line with research by Savicki et al. (1996a) and Herring (1993). Thus, it seems that male students seemed to behave as predicted by social role theory and expectation states theory. However, the finding that female students used more argumentative language is in contrast with both theories. This could be due to the nature of the group task students had to complete. Although the task was an "open" task (i.e., no right or wrong answer), the task description contained clear descriptions about what was expected from students. This may have reduced group members' uncertainty about what they had to do in order to complete the group task successfully, thereby diminishing gender differences and thus stimulating female students to behave more assertively by giving more arguments (LePine et al., 2002).

The fourth research question was whether gender and gender group composition affected students' collaborative activities during online collaboration. Some effects of gender were found. First, male students indicated more social resistance than female students. Since social resistance remarks usually consist of displays of negative emotions and insults, this result mirrors the abovementioned finding that male students used more negative language. Second, female students devoted more effort to maintaining and signaling mutual understanding. Again, since mutual understanding remarks frequently consist of agreements, this result mirrors the abovementioned finding that female students indicate agreement more during online collaboration. Several other effects of gender approached statistical significance. Most interestingly, male students were more off-task than female students: a significant effect of gender on off-task remarks was found, although the corresponding χ^2 value was not significant. No effects of gender group composition were found, although the effect on off-task remarks approached statistical significance. In male dominated groups, students were more off-task compared to female dominated groups. It should be noted however, that off-task behavior did not occur very often during online collaboration. In conclusion, the results for the fourth research question seem to be in line with social role and expectation states theory, since male students adopted a more assertive collaborative approach, while female students used more affiliative styles of collaboration.

The fifth research question addressed the influence of gender and gender group composition on students' perception of collaboration. No significant effects of gender or gender group composition were found. Although female students reported more occurrences of positive group behavior and less occurrences of negative group behavior, the effects of gender were only marginally significant. This is surprising, for several reasons. First, this finding is contradictory to other studies which found female students to hold more positive perceptions of collaboration than male students (e.g., Adrianson, 2001). Second, because male students displayed more assertive and negative behavior during online collaboration, this could have had a negative impact on the perceptions of students. This was not the case however. Perhaps male students' negative behavior during online collaboration does not lower female students' perceptions of collaboration, nor do male-dominated groups perceive their collaboration to be more negative. It may be the case that not all negative language used by male students is intended to be negatively nor is it perceived negatively. Male students may simply have a collaborative style that consists of being more assertive and rude fitting the male stereotype, but this style does not automatically have to hamper collaboration. A final explanation may lie in the interpersonal

trust between group members. Students worked online during eight lessons over period of four weeks. This period may have been enough for students to develop interpersonal trust. Furthermore, students knew each other before the start of the study because they followed the same class. This familiarity may also have lead to higher levels of interpersonal trust. These possibly higher levels of trust may have lead to more positive perceptions of collaboration (Wilson et al., 2006).

The final research question investigated whether group composition affected group performance during online collaboration. To answer this question the quality of the group products was assessed. On average, female dominated groups only marginally outperformed male dominated groups. Thus, no effect of gender group composition was found. This may be explained by the fact that the effects of gender group composition were not significant in most cases (e.g., concerning participation and affiliative language). Thus, although female dominated groups collaborated differently compared to male dominated groups, these differences were too small to influence group performance neither negatively, nor positively. A second explanation may be that male students' collaboration styles may not necessarily be detrimental for male dominated groups. Perhaps male dominated groups do not suffer from using negative language or using less argumentative language. A third explanation may be that the collaboration in male dominated groups was not bad as suggested. Although male dominated groups used for example more negative language were more off-task, by far the largest part of their collaborative activities was on-task. In male dominated groups too, students were intensively engaged performing and coordinating task-related activities.

Several possible limitations of this study should be kept in mind when interpreting the results of this study. First, the use of a relatively small number of students and groups from a single school may limit the representativeness of the data gathered. More research involving larger number of students and groups is needed to determine whether the results of this study can be replicated. Second, as mentioned before, the group task used in this study may have influenced the results found. Future research should investigate the effects of gender and gender group composition involving different types of tasks (e.g., open or closed, much structure or little structure). Third, this study does not explain *why* female students behaved differently online than male students did. Future research could examine whether differences exist between the perceived expertise of male and female students and whether perceived expertise explains students' online behavior (Thomas-Hunt & Phillips, 2004). Finally, this study only discriminated between male and female dominated groups. Several other studies discriminated between male only, female only and mixed sex groups (e.g., Savicki et al., 2002; Van der Meijden, 2005). This study only involved one male only group and one female only group, whereas the other groups were mixed (but either dominated by male students or by female students). Examining differences between male only and female only groups was therefore not an option. This may explain why some results differ from previous research. Male students may behave more extremely in male only groups (Savicki et al., 1996a, 1996b), for example by typing even more negative messages. Future research should examine different types of group compositions (e.g., male only, male dominated, gender balanced, female dominated, and female only).

In conclusion, the present study demonstrates that gender is also "at play" (Ridgeway, 2001, p. 14185) during computer-supported collaborative learning. This study highlighted some important differences between female and male students' online collaborative behavior. The finding that female students participate more and give more arguments during online collaboration, is important because previous research has demonstrated that participation and constructing messages that contain reasons and arguments are positively related to group performance and individual learning (Chizhik, 2001; Cohen, 1994; Van der Meijden & Veenman, 2005). Thus, during the present study male

students may have been at a disadvantage compared to female students. An important question for educators and researchers is therefore how students can be stimulated to participate more and participate in ways that are more beneficial for collaboration.

ACKNOWLEDGEMENT

This project was funded by NWO, the Netherlands Organisation for Scientific Research under project number 411-02-121. The authors would also like to thank Marcel Broeken for his technical and programming assistance during the preparation and execution of the experiment.

REFERENCES

Adrianson, L. (2001). Gender and computer-mediated communication: Group processes in problem solving. *Computers in Human Behavior, 17*(1), 71-94.

Bonito, J. A. (2000). The effect of contributing substantively on perceptions of participation. *Small Group Research, 31*(5), 528-553.

Chizhik, A. W. (1999). Can students work together equitably? An analysis of task effects in collaborative group work. *Social Psychology of Education, 3*(1-2), 63-79.

Chizhik, A. W. (2001). Equity and status in group collaboration: Learning through explanations depends on task characteristics. *Social Psychology of Education, 5*(2), 179-200.

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*(3), 805-813.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*(1), 1-35.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554-571.

De Jong, F., Kollöffel, B., Van der Meijden, H., Kleine Staarman, J., & Janssen, J. (2005). Regulative processes in individual, 3D and computer supported cooperative learning contexts. *Computers in Human Behavior, 21*(4), 645-670.

Dembo, M. H., & McAuliffe, T. J. (1987). Effects of perceived ability and grade status on social interaction and influence in cooperative groups. *Journal of Educational Psychology, 79*(4), 415-423.

Dennis, A. R., Kinney, S. T., & Hung, Y. T. C. (1999). Gender differences in the effects of media richness. *Small Group Research, 30*(4), 405-437.

Dubrovsky, V. J., Kiesler, S., & Sethna, B. N. (1991). The equalization phenomenon: Status effects in computer-mediated and face-to-face decision-making groups. *Human-Computer Interaction, 6*(2), 119-146.

Erkens, G. (2004). Dynamics of coordination in collaboration. In J. van der Linden & P. Renshaw (Eds.), *Dialogic learning: Shifting perspectives to learning, instruction, and teaching* (pp. 191-216). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Erkens, G. (2005). Multiple Episode Protocol Analysis (MEPA). Version 4.10. Retrieved October 24, 2005, from http://edugate.fss.uu.nl/mepa/

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior, 21*(3), 463-486.

Fischer, F., Bruhn, J., Gräsel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction, 12*(2), 213-32.
Fjermestad, J. (2004). An analysis of communication mode in group support systems research. *Decision Support Systems, 37*(2), 239-263.
Fletcher-Flinn, C. M., & Gravatt, B. (1995). The efficacy of computer assisted instruction (CAI): A meta-analysis. *Journal of Educational Computing Research, 13*(3), 219-241.
Herring, S. C. (1993). Gender and democracy in computer-mediated communication. *Electronic Journal of Communication / Revue Electronique de Communication, 3*(2). Retrieved January 4, 2006, from http://www.cios.org/www/ejc/v3n293.htm
Howell-Richardson, C., & Mellar, H. (1996). A methodology for the analysis of patterns of participation within computer mediated communication courses. *Instructional Science, 24*(1), 47-69.
Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (in press). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education*.
Jaspers, J., Broeken, M., & Erkens, G. (2004). Virtual Collaborative Research Institute (VCRI) (Version 2.0). Utrecht: Onderwijskunde Utrecht, ICO/ISOR.
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*(4), 775-790.
Johnson, D. W., & Johnson, R. T. (1999). *Learning together and alone: Cooperative, competitive, and individualistic learning* (5th ed.). Boston: Allyn and Bacon.
Johnson, D. W., Johnson, R. T., & Stanne, M. B. (1990). Impact of group processing on achievement in cooperative groups. *Journal of Social Psychology, 130*(4), 507-516.
Kiesler, S., & Sproull, L. (1992). Group decision making and communication technology. *Organizational Behavior and Human Decision Processes, 52*(1), 96-123.
King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal, 31*(2), 338-368.
Kreijns, K. (2004). *Sociable CSCL environments: Social affordances, sociability, and social presence.* Unpublished PhD thesis, Open Universiteit, Heerlen, The Netherlands.
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*(3), 335-353.
Kuhn, D., & Udell, W. (2003). The development of agrument skills. *Child Development, 74*(5), 1245-1260.
Kumpulainen, K., & Mutanen, M. (1999). The situated dynamics of peer group interaction: An introduction to an analytic framework. *Learning and Instruction, 9*(5), 449-473.
Leaper, C., & Smith, T. E. (2004). A meta-analytic review of gender variations in children's language use: Talkativeness, affiliative speech, and assertive speech. *Developmental Psychology, 40*(6), 993-1027.
LePine, J. A., Hollenbeck, J. R., Ilgen, D. R., Colquitt, J. A., & Ellis, A. (2002). Gender composition, situational strength, and team decision-making accuracy: A criterion decomposition approach. *Organizational Behavior and Human Decision Processes, 88*(1), 445-475.
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*(5), 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*(3), 449-521.

McGrath, J. E. (1991). Time, interaction, and performance (TIP). *Small Group Research, 22*(2), 147-174.

O'Donnell, A. M., & O'Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review, 6*(4), 321-349.

Postmes, T., & Spears, R. (2002). Behavior online: Does anonymous computer communication reduce gender inequality? *Personality and Social Psychology Bulletin, 28*(8), 1074-1083.

Ridgeway, C. L. (2001). Small-group interaction and gender. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (Vol. 21, pp. 14185-141289). Amsterdam: Elsevier.

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.

Savicki, V., Kelley, M., & Ammon, B. (2002). Effects of training on computer-mediated communication in single or mixed gender small task groups. *Computers in Human Behavior, 18*(3), 257-270.

Savicki, V., Kelley, M., & Lingenfelter, D. (1996a). Gender and group composition in small task groups using computer-mediated communication. *Computers in Human Behavior, 12*(2), 209-224.

Savicki, V., Kelley, M., & Lingenfelter, D. (1996b). Gender, group composition, and task type in small task groups using computer-mediated communication. *Computers in Human Behavior, 12*(4), 549-565.

Savicki, V., Kelley, M., & Oesterreich, E. (1998). Effects of instructions on computer-mediated communication in single- or mixed-gender small task groups. *Computers in Human Behavior, 14*(1), 163-180.

Schellens, T., & Valcke, M. (2006). Fostering knowledge construction in university students through asynchronous discussion groups. *Computers & Education, 46*(4), 349-370.

Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Straus, S. G. (1996). Getting a clue: The effects of communication media and information distribution on participation and performance in computer-mediated and face-to-face groups. *Small Group Research, 27*(1), 115-142.

Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education, 46*(1), 29-48.

Thomas-Hunt, M. C., & Phillips, K. W. (2004). When what you know is not enough: Expertise and gender dynamics in task groups. *Personality and Social Psychology Bulletin, 30*(12), 1585-1598.

Van der Meijden, H. (2005). *Knowledge construction through CSCL: Student elaborations in synchronous, asynchronous, and three-dimensional learning environments.* Unpublished PhD thesis, Radboud University Nijmegen, The Netherlands.

Van der Meijden, H., & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior, 21*(5), 831-859.

Van Meter, P., & Stevens, R. J. (2000). The role of theory in the study of peer collaboration. *Journal of Experimental Education, 69*(1), 113-127.

Webb, N. M. (1995). Group collaboration in assessment: Multiple objectives, processes, and outcomes. *Educational Evaluation and Policy Analysis, 17*(2), 239-261.

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. C. Berliner (Ed.), *Handbook of educational psychology* (pp. 841-873). New York: Simon & Schuster Macmillan.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education, 46*(1), 71-95.

Wilson, J. M., Straus, S. G., & McEvily, B. (2006). All in due time: The development of trust in computer-mediated and face-to-face teams. *Organizational Behavior and Human Decision Processes, 99*(1), 16-33.

Wolfe, J. L. (1999). Why do women feel ignored? Gender differences in computer-mediated classroom interactions. *Computers and Composition, 16*(1), 153-166.