Title: Awareness of Group Performance in a CSCL Environment: Effects of Peer Feedback and Reflection

Abstract: This study investigated the effects of a peer feedback tool and a reflection tool on social and cognitive performance during computer supported collaborative learning (CSCL). A CSCL-environment was augmented with a peer feedback tool (Radar) and a reflection tool (Reflector) in order to make group members aware of both their individual and their group behavior. Radar visualizes how group members perceive their own social and cognitive performance and that of their peers during collaboration along five dimensions. Reflector stimulates group members to reflect upon their own performance and the performance of the group. A 2x2 factorial between-subjects design was used to examine whether Radar and Reflector would lead to better team development, more group satisfaction, lower levels of group conflict, more positive attitudes toward problem-based collaboration, and a better group product. Results show that groups with Radar perceived their team as being better developed, experienced lower conflict levels, and had a more positive attitude towards collaborative problem solving than groups without Radar. The quality of group products, however, did not differ. The results demonstrate that peer feedback on the social performance of individual group members can enhance the performance and attitudes of a CSCL-group.
AWARENESS OF GROUP PERFORMANCE IN A CSCL ENVIRONMENT: EFFECTS OF PEER FEEDBACK AND REFLECTION

Abstract

This study investigated the effects of a peer feedback tool and a reflection tool on social and cognitive performance during computer supported collaborative learning (CSCL). A CSCL-environment was augmented with a peer feedback tool (Radar) and a reflection tool (Reflector) in order to make group members aware of both their individual and their group behavior. Radar visualizes how group members perceive their own social and cognitive performance and that of their peers during collaboration along five dimensions. Reflector stimulates group members to reflect upon their own performance and the performance of the group. A 2x2 factorial between-subjects design was used to examine whether Radar and Reflector would lead to better team development, more group satisfaction, lower levels of group conflict, more positive attitudes toward problem-based collaboration, and a better group product. Results show that groups with Radar perceived their team as being better developed, experienced lower conflict levels, and had a more positive attitude towards collaborative problem solving than groups without Radar. The quality of group products, however, did not differ. The results demonstrate that peer feedback on the social performance of individual group members can enhance the performance and attitudes of a CSCL-group.

KEYWORDS: groupware; collaborative learning; peer feedback; reflection; group performance.
Introduction

Well performing teams go through several stages of group development (e.g., Gersick, 1988; Tuckman & Jensen, 1977). While the research showing this was carried out in face-to-face teams, Tuckman and Jensen’s (1977) concept of group development stages also seems to be relevant to virtual learning groups (Johnson, Suriya, Yoon, Berret, & La Fleur, 2002). Tuckman and Jensen observed and distinguished five stages, namely: (1) forming (i.e., getting to know each other and the task at hand), (2) storming (i.e., establishing roles and positions within the group), (3) norming (i.e., reaching consensus about behavior, goals en strategies, (4) performing (i.e., reaching conclusions and delivering results), and (5) adjourning (i.e., dismantling of the group when the task is completed). Each of these five stages involves two aspects: interpersonal relationships (i.e., social and socio-emotional aspects) and behavior to accomplish the task (i.e., cognitive aspects). It is especially the social or socio-emotional aspects of group development processes such as developing positive affective relationships, group cohesiveness, feelings of trust, and a sense of community, that are very important for a group to reach their full potential (Kreijns & Kirschner, 2004).

However, most computer supported collaborative learning (CSCL) environments focus primarily on the support of cognitive processes in collaboration, and limit the possibility for social processes to take place (Kreijns & Kirschner, 2004). Moreover, group members are often not fully aware that their behavior is not in the best interest of the groups’ development or product (Karau & Williams, 1993). Therefore, in this study, a CSCL-environment was augmented with a peer feedback tool (Radar) and a reflection tool (Reflector) in order to support the social processes during collaboration, and to make group members aware of their individual and group behavior. The aim of this study is to examine the effects of these two tools on team development, group satisfaction, level of group conflict, attitude towards collaborative problem solving, and the quality of the groups’ product.
Awareness of Group Performance in CSCL

Social and Cognitive Processes in CSCL

Collaborative learning can be defined as the “mutual engagement of participants in a coordinated effort to solve the problem together” (Rochelle & Teasley, 1995, p. 70). From a cognitive or task-related perspective, collaborative learning can lead to deeper level learning, critical thinking, shared understanding, long term retention of the learned material (e.g. Garrison, Anderson, & Archer, 2001; Johnson & Johnson, 1999). From a social or non-task related perspective, collaborative learning can stimulate students to develop social and communication skills (teamwork skills), more positive attitudes towards group members, better social relationships and higher levels of group cohesion (Boud, Cohen, & Sampson, 1999; Johnson, Johnson, & Smith, 2007). These effects can be reinforced when collaborative learning is embedded in an authentic context and applied to ill-structured and complex learning tasks (Jonassen, 1991, 1994).

A key element in successful collaborative learning is social interaction (Kreijns, Kirschner, & Jochems, 2003; Liaw & Huang, 2000; Northrup, 2001). Social interaction is not only important for the cognitive (task-related) processes in collaboration, such as discussion, reasoning, reflection, critical thinking and creating a shared understanding of the problem (Kreijns, Kirschner, & Jochems, 2003), it is equally important for the social (non-task related) processes in collaboration, such as developing positive affective relationships, group cohesiveness, feelings of trust, and a sense of community (Kreijns & Kirschner, 2004). These social processes allow group members to get to know and understand each other, and are necessary to become a ‘healthy’ community of learning (Gunawardena, 1995). Both cognitive and social processes are necessary to collaboratively complete a task, solve a problem or construct knowledge (Kreijns, Kirschner, & Jochems, 2003). These processes, however, do not automatically happen by simply bringing learners together (Fischer, Bruhn, Gräsel, & Mandl, 2002). Therefore, Kreijns, Kirschner, and Jochems (2003) stated that the primary aim of computer supported collaborative learning (CSCL) is to provide an environment that supports
and enhances collaboration between students, in order to enhance students’ cognitive processes as well as their social processes.

**Social and Cognitive Benefits of Groups in CSCL environments**

The rapid development of information and communication technologies (ICT), has led to computer applications (e.g., e-mail, chat, discussion forums, video conferencing, simulations, 3-D models, visualizations and external representations) which have proven to be useful tools or widgets to support collaborative learning (Janssen, Erkens, Kanselaar, & Jaspers, 2007).

Several researchers report cognitive and social benefits for groups in CSCL environments as compared to contiguous (i.e., face-to-face) groups. First, concerning the *cognitive aspects* of collaboration, researchers have found that students working in CSCL-environments report higher levels of learning (Hertz-Lazarowitz & Bar-Natan, 2002), make higher quality decisions, deliver more complete reports, participate more equally (Fjermestad, 2004; Janssen, Erkens, Kanselaar, & Jaspers, 2007), and engage in more complex, broader, and challenging discussions (Benbunan-Fich, Hiltz, & Turoff, 2003) than students working face-to-face. With respect to the *social aspects*, students working in CSCL-environments report higher levels of satisfaction compared to students in contiguous groups (Fjermestad, 2004).

There are, however, also contradictory results. First, concerning the *cognitive aspects* of collaboration, students working in CSCL-environments sometimes perceive their discussions as more confusing (Thompson & Coover, 2003), as being less productive (Straus, 1997; Straus & McGrath, 1994) and need more time to reach consensus and to make decisions (Fjermestad, 2004) than students working face-to-face. Second students working in CSCL-environments have been found to show lower levels of participation (Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003), to experience higher levels of conflict (Hobman, Bordia, Irmer, & Chang, 2002), to experience lower levels group cohesiveness (Straus, 1997; Straus & McGrath, 1994) and to experience lower levels of satisfaction (Baltes, Dickson, Sherman, Bauer, & LaGanke, 2002) as
compared to students working in contiguous groups. In other words, students working in CSCL-environments do not always reach their full potential.

Finally, there are also studies which show that there is little difference between face-to-face and CSCL-groups, especially with respect to characteristic problems and difficulties (O'Donnell & O'Kelly, 1994; Salomon & Globerson, 1989), such as social loafing (i.e., where group members invest less effort in a group, compared to working individually), or the free rider effect (i.e., where students let other group members do the work for them).

**Effects of CSCL-design on Social and Cognitive Behavior**

Two important reasons for the disparity between the potential of groups working in CSCL-environments and their performance lies in (1) the design of the CSCL-environment, and (2) the social and cognitive behavior of the group members. With respect to the former, most CSCL-environments focus primarily on supporting cognitive or task-related processes in collaboration and limit the possibility for social or non-task related processes to take place (Kreijns & Kirschner, 2004). For instance, despite technological advances, most CSCL environments use still text-based computer mediated communication (CMC) systems making use of email, chat and/or discussion boards, which cannot easily convey visual nonverbal cues (Kreijns, Kirschner, & Jochems, 2003). The absence of these cues can cause specific problems for effective communication and interaction between group members since this removes possibilities for the exchange of socio-emotional and affective information, and decreases information available about group members’ presence, self-image, attitudes, moods, actions and reactions (Short, Williams, & Christie, 1976). According to Short et al. the functions of these nonverbal cues are in some way related to forming, building or maintaining social relationships. Therefore, CMC can have negative effects on impression formation and group members’ social behavior (e.g. Garton & Wellman, 1995; Walther, Anderson, & Park, 1994).

With respect to the latter, group members form interpersonal perceptions during
interaction (Kenny, 1994). Based on what they see and experience, they form impressions (e.g., norms, values, beliefs) about themselves, the group, other group members, and what the other group members think of them. These self-, other- and meta-perceptions are based on the perceived cognitive behaviors (e.g., productivity) and social behaviors (e.g., dominance and friendliness) that occur during interaction. Based upon these perceptions, group members determine their own social and cognitive behavior, and develop social relationships with each other. However, research has shown that group members’ perceptions of their own performance (i.e., self-perception) and their perceptions of group performance are generally unrealistically positive, resulting in an illusion of group productivity (Stroebe, Diehl, & Abakoumkin, 1992). This tendency to believe that their group is performing effectively, while it often is not, can result in a reduce of effort by group members, a phenomenon also known as social loafing (Williams, Harkins, & Latané, 1981), which further undermines the groups’ social and cognitive performance. However, group members are often not aware that they are loafing, or are unwilling to admit it (Karau & Williams, 1993).

To overcome this obstacle to social and cognitive performance, CSCL environments can be augmented with computer tools or widgets that act as social contextual facilitators relevant for the learner’s social interaction (Kirschner, Strijbos, Kreijns, & Beers, 2004). These tools, also known as ‘social affordance devices’, can positively affect social and cognitive performances in a CSCL environment (Kirschner, et al., 2004). Two operationalizations of such tools are used in this research, namely a peer feedback tool to make group members aware of the social and cognitive behavior of themselves, their peers, and how this is perceived by others, and a reflection tool to stimulate group members to reflect upon their individual behavior, why their peers see them the way they do, and to also reflect collaboratively (i.e., co-reflect) on the performance of the group as a whole. The aim of these tools is to make group members aware of their social and cognitive behavior and to enhance their social and cognitive performance and that of the group. The next sections deal with aspects central to these tools, namely peer feedback
and reflection.

**Peer Feedback**

Feedback can be described as information provided to an individual to increase performance (Kluger and DeNisi, 1996). In a learning situation, this information can come from many sources such as teachers, computers, fellow students, and so forth. This study centers on peer feedback; information, provided by peers (i.e., fellow learners), which is intended to increase performance. This information can be provided on the outcome performance (i.e., outcome feedback), or on how one is performing (i.e., process feedback). Feedback can be given by individuals or groups, and can also be received by individuals or groups. In this study, peers provide process feedback at the individual and the group level, in order to enhance interpersonal behavior. It is expected that enhancement of interpersonal behavior will have a positive effect on a group’s social performance (Geister, Konradt & Hertel, 2006; McLeod & Liker, 1992), as well as an indirect positive effect on a group’s cognitive performance (Kreijns, Kirschner, & Jochems, 2003).

Process feedback can include cognitive or task-related information (e.g., task behaviors, actions and strategies), or social or non-task related information, such as information about interpersonal behavior (e.g., dominance and friendliness) or teamwork (Geister, Konradt & Hertel, 2006). McLeod and Liker (1992) found that process feedback at the group level on the interpersonal behavior of student group members, such as dominance and group orientedness, changed the dominance behavior of individual group members. Two other studies investigating individualized peer feedback on interpersonal behavior of group members (e.g., communication and collaboration), found that such feedback led to increased cooperation, communication, satisfaction and motivation in group members (Dominick, Reilly, & McGourty, 1997; Druskat & Wolff, 1999).

The premise behind one of the tools developed for and used in this research is that an
individualized peer feedback tool that provides group members information about the social and cognitive behavior of themselves, their peers, and the group as a whole, will positively alter the social and cognitive behavior of individuals and group.

**Reflection**

Simply providing group members’ with information on their cognitive and social behavior is not enough to positively alter their behavior (Prins, Sluijsmans, & Kirschner, 2006). Group members also need to process this information and ask themselves whether they understand, accept, and agree with the feedback. In other words, they must reflect upon the feedback. Reflection is the intellectual and affective activities individuals engage in to explore their experiences (e.g., behaviors, ideas, feelings) in order to reach new understandings and appreciations (Boud, Keogh, & Walker, 1985). The feedback receiver needs to be challenged to reflect on his/her own performance, and determine whether the feedback provides clues for behavioral change (Prins, Sluijsmans, & Kirschner, 2006). Therefore, it is expected that peer feedback in combination with reflection will even be more effective than feedback alone (e.g., Schön, 1987).

According to Boud, Keogh, and Walker (1985), reflection can lead to new perspectives on experience, changes in behavior, readiness for application, and commitment to action. Therefore, reflection on peer feedback should make group members more aware of their own individual behavior, how their behavior affects others, and whether they should alter their behavior. Awareness can be defined as the “understanding of the activities of others, which provides a context for your own activity” (Dourish & Bellotti, 1992, p. 107).

Members of CSCL groups often have a limited view on the behavior, ideas, and feelings of their peers, because – as stated - most CSCL environments use text-based computer mediated communication systems, such as email or discussion boards that are not capable of transferring visual non-verbal cues (Kreijns & Kirschner, 2004). The absence of these visual non-verbal cues
can cause specific problems for effective communication and impression formation between group members, since it removes possibilities for the expression of socio-emotional state of mind and decreases the information available about the other’s presence, self-image, attitudes, moods, actions, and reactions (Jacobson, 1999; Short, Williams, & Christie, 1976).

Based upon these ideas, a shared reflection tool in which group members individually reflect and provide information on (1) their own individual perspective on group performance, (2) their own personal contribution to group performance, (3) how their own behavior is perceived by others, (4) their own personal perspective on their own behavior was developed and studied. Because group performance is determined by the individual effort of all group members, this tool also stimulates group members to collaboratively reflect (i.e., co-reflect) on the group performance and reach a shared conclusion about this. Co-reflection is defined as “a collaborative critical thinking process involving cognitive and affective interactions between two or more individuals who explore their experiences in order to reach new intersubjective understandings and appreciations” (Yukawa, 2006; p. 206).

Research Questions

This study investigated the effect of a peer feedback tool and a reflection tool for both the social and cognitive behavior of individual group members working in a CSCL-environment, and the social and cognitive performance of the group as a whole. To this end, an existing CSCL-environment was augmented with two independent, but complementary, tools. The first was an individualized peer feedback tool - Radar - which was meant to stimulate and provide group members with information about the social and cognitive behavior of themselves, their peers, and the group as a whole. This information was presented from both the perspectives of the group members themselves (i.e., self perceptions), their peers (i.e., peer perceptions) and the group as a whole. The second tool was a shared reflection tool - Reflector - which was meant to stimulate group members to reflect on and provide information about their personal perspectives
on the group’s performance, their own contributions, their own behavior and how this behavior was perceived by their peers, as well as to co-reflect on the group performance and reach shared understanding on this.

The following research questions will be addressed:

1. Do groups with Radar and Reflector show higher differences between self assessments and peer assessments between three successive measurement moments, than groups with only Radar?

   Expected is that the peer feedback provided by Radar at the first assessment should make group members aware of their unrealistic self perceptions and peer perceptions, resulting in a decrease of self assessment and peer assessment scores at a subsequent assessment. Also, a combination of Radar and Reflector should lead to even lower self assessment and peer assessment scores than groups with only Radar.

2. Do groups with Radar and Reflector show more congruency between self assessments and peer assessments at T3 than groups without Radar and/or Reflector alone?

   Expected is that Radar and Reflector will cause group members to adjust their unrealistic positive self perceptions towards more realistic perceptions of their peers between a second and third measurement. Therefore, groups with Radar and Reflector should show the highest positive correlations between self assessments and peer assessments.

3. Do members of groups with Radar and Reflector perceive themselves and others to exhibit better social and cognitive behavior than those in groups without Radar and/or Reflector?

   Expected is that both Radar and Reflector should positively affect perceived social and cognitive behavior, with a combination of Radar and Reflector being
Awareness of Group Performance in CSCL

most effective.

4. Do groups with Radar and Reflector perform better socially than groups without Radar and/or Reflector? In other words, do groups using Radar and Reflector develop better, have higher group satisfaction, have lower levels of group conflict, and have more positive attitude towards collaborative problem solving than groups without Radar and/or Reflector?

Expected is that both Radar as Reflector will positively affect the social behavior in the group, and that this should lead to an increase in the social performance of the group. A combination of both tools should be most effective.

5. Do groups with Radar and Reflector perform better cognitively than groups without Radar and/or Reflector? In other words, do groups with Radar and Reflector produce a group product of higher quality than groups without Radar and/or Reflector?

Expected is that both Radar as Reflector will positively affect the social behavior in the group and that this should indirectly lead to an increase in the cognitive performance of the group. A combination of both tools should be most effective.

Method

Participants

Participants were 39 fourth-year students (19 male, 20 female), with an average age of 16 (\(M = 15.54, \ SD = .60, \ Min =14, \ Max = 17\)), from an academic high school in The Netherlands. Students came from two classes and were enrolled in the second stage of the pre-university education track which encompasses the final three years of high school. The participants were randomly assigned by the researchers to groups of three or four, and to one of the four conditions (see Design). Group compositions were heterogeneous in ability and gender.
**Design**

A 2x2 between-subjects factorial design was used with the factors Radar unavailable (~Ra) – available (+Ra), and Reflector unavailable (~Rf) – available (+Rf). This leads to four conditions (~Ra~Rf, +Ra~Rf, ~Ra+Rf, +Ra+Rf). The condition with Radar and Reflector (+Ra+Rf) consisted of 11 students (2 groups of 4, and 1 group of 3), without Radar but with Reflector (~Ra+Rf) of 12 students (3 groups of 4), and with Radar but without Reflector (+Ra~Rf) and without both tools (~Ra~Rf) of 8 students (2 groups of 4).

--- INSERT TABLE 1 ---

**Measures (see Table 1)**

*Social behavior.* The perceived social behavior in the group is measured by the self assessments and peer assessments in Radar on four variables, namely ‘influence’, ‘friendliness’, ‘cooperativeness’ and ‘reliability’. These variables are rated on a continuous scale ranging from 0 to 4 (0 = none, 4 = very high).

*Cognitive behavior.* The perceived cognitive behavior in the group is measured by the self assessments and peer assessments in Radar on the variable ‘productivity’, that was rated on a continuous scale ranging from 0 to 4 (0 = none, 4 = very high).

*Cognitive performance.* The grade given to the groups’ collaborative writing task (i.e., the essay) was used as a measure of cognitive performance. The essays were graded by two researchers, both experienced in grading essays. The inter-rater reliability was high (n = 10, Cronbach’s α = .86).

*Social performance.* To measure social performance, previously validated instruments (Strijbos, Martens, Jochems, & Broers, 2007) were translated into Dutch and transformed into 5-point Likert scales (1 = totally disagree, 5 = totally agree; see Table 2). The Team Development Satisfaction scale provides information on the perceived level of group cohesion. The Group-process Satisfaction scale provides information on the perceived satisfaction with general group
functioning. The Intra-group Conflicts scale provides information on the perceived level of conflict between group members. The Attitude towards Collaborative Problem Solving scale provides information on the perceived level of group effectiveness and how group members felt about working and solving problems in a group.

--- INSERT TABLE 2 ---

Task and procedure

The students collaborated in groups of three or four on a collaborative writing task in sociology. Every student worked at a computer. Each group had to write one essay about *Fitna* - a very contentious film - which argues that Islam encourages, among other things, terrorism, anti-Semitism, sexism, violence against women, and Islamic universalism. This task was considered highly civically relevant by the school. The collaborative writing task consisted of two 90-minute sessions separated by one week. The groups collaborated in a CSCL environment called Virtual Collaborative Research Institute (VCRI; Jaspers, Broeken, & Erkens, 2002) which is a groupware program designed to support collaborative learning on research projects and inquiry tasks. VCRI will be further described in the Instruments section. Students were instructed to use VCRI to communicate with the other group members and to make complete use of the tools for peer feedback and reflection when the experimental condition allowed this. Students received content information and definitions regarding to the five variables on which they had to assess themselves and their peers. Students were told that they had four lessons to complete the task, that it would be graded by their teacher, and that it would affect their grade for the course. The introduction to the task stressed the importance of working together as a group and pointed out that each individual group member was responsible for the successful completion of the group task. To successfully complete the task, all group members had to participate.

During collaboration, groups with a peer feedback tool (i.e., +RA~RF, +RA+RF) used the tool at the beginning of the experiment (T1), halfway through the experiment (i.e., at the end
of the first session; T2), and at the end of the second and final session (T3). The groups with a reflection tool (i.e., ~RA+RF, +RA+RF), used the tool twice, namely halfway through the experiment (T2) and at the end of the final session (T3). While groups with Radar and/or Reflector used the tools, groups without Radar and/or Reflector continued working on their collaborative writing task. Groups with Radar and/or Reflector received extra time for their collaborative writing task so that time-on-task was equal for all conditions. At the end of the final session (T3), the peer assessment- and reflection tools became available for all conditions so that all participants could assess their peers and reflect on their behaviors. Finally, all participants completed a 30-item questionnaire measuring the social performance of the group.

**Tools**

*Virtual Collaborative Research Institute (VCRI).* The Virtual Collaborative Research Institute (VCRI) is a groupware program that supports collaborative working and learning on research projects and inquiry tasks (Jaspers, Broeken, & Erkens, 2004). VCRI contains more than 10 different tools, but only 6 were used for this experiment (see Figure 1).

--- INSERT FIGURE 1 ---

The Chat tool (top left) is used for synchronous communication between group members. The chat history is automatically stored and can be re-read by participants at any time. Users can search for relevant historical information using the Sources tool (top centre). The Co-Writer (top right) is a shared word-processor, which can be used to write a group text. Using the Co-Writer, students can simultaneously work on different parts of their texts. Notes (bottom left) is a note pad which allows the user to make notes and to copy and paste selected information. Radar for peer feedback (bottom centre) and Reflector for reflection (bottom right) will be described in the following sections. Windows of the available tools are automatically arranged on the screen, when students log on to the VCRI.
Peer feedback tool (Radar). The VCRI was augmented with a peer feedback tool for stimulating and facilitating information of group members’ social and cognitive behavior. This information is visualized in a radar diagram; therefore the peer feedback tool is named ‘Group Radar’ (in this article Radar; see Figure 2). Radar provides users with anonymous information on how their cognitive and social behavior is perceived by themselves, their peers, and the group as a whole. The information gathered is based on specific traits that have been found to tacitly affect how one ‘rates’ other people (Brok, Brekelmans, & Wubbels, 2006). Radar provides information on five traits that are important for assessing behavior in groups. Four are related to social or interpersonal behavior, namely (1) influence; (2) friendliness; (3) cooperation; (4) reliability; and one to cognitive behavior, namely (5) productivity. These traits are derived from studies on interpersonal perceptions, interaction, group functioning, and group effectiveness (e.g., Bales, 1988; Brok, Brekelmans, & Wubbels; Kenny, 1994; Salas, Sims, & Burke, 2005).

Influence is directly derived from Wubbels, Créton, and Hooymayers’ (1985) influence dimension (i.e., dominance vs. submissiveness) in their model for interpersonal teacher behavior. This dimension is also used by Bales (1988) and represents the prominence, status, power, and personal influence that the individual is seen to have in relation to other group members. The variable is labeled ‘influence’, and not ‘dominance’ or ‘submissive’, because those labels can be perceived as negative traits.

Friendliness is one of the eight behavior categories from Wubbels, Créton, and Hooymayers’ (1985) model for interpersonal teacher behavior. Bales (1988) used a similar dimension (i.e., friendliness vs. unfriendliness). Bales and Cohen (1979) defined this as the extent to which individual members are friendly and respectful to each other.

Cooperation, which denotes the degree to which someone is willing to work with others, is derived directly from Wubbels et al’s (1985) dimension Proximity (i.e., opposition vs. cooperation) They defined proximity as the property of being close together, or in group settings as the feeling of being a group (i.e., group cohesiveness).
Reliable is considered a trait reflecting ‘trust’ which has been identified as an important precursor for successful collaboration, both in face-to-face teams (Castleton_Parterns/TCO, 2007) and in CSCL (Jarvenpaa & Leidner, 1999). According to Emans, Koopman, Rutte, and Steensma (1996) trust can be seen as the cognitive and affective assurance of group members that they respect each other’s interests and, therefore, can orient themselves towards each other’s words, actions, and decisions with an easy conscience.

Productivity is the extent to which individual members contribute to tasks or duties, central to group performance or group efficiency (Salas, Sims, & Burke, 2005). This trait, which represents cognitive or task-related behavior, was selected because research has shown that group members monitor the performance of their other group members in comparison to their own performance (Salas, Sims, & Burke).

In Radar, all group members are both assessor and assessee. In the role of assessor, the to-be-assessed peer in the group can be selected and her/his profile will appear as dotted lines in the centre circle of the radar diagram. Each group member is represented by a specific color. The assessor rates her/himself and all of the other group members on each of the five traits using a continuous scale ranging from 0 to 4 (0 = none - 4 = very high). To make sure that all assessors interpret the five traits in the same way, assessors saw a text balloon with content information and definitions when they moved the cursor across one of the five traits in the tool. For example, when the assessor moves the cursor across ‘influence’ a balloon pops up with the text ‘A high score on influence means that this person has a big influence on what happens in the group, other group members behavior, and the form and content of the group product (the essay)’.

The ratings are automatically saved in a database. To simplify data-analysis, the ratings are transformed to a scale from 0 to 100 by multiplying the ratings by 25. The assessment is anonymous; group members can see the output of the assessments of the other group members, but cannot see who entered the data.

In order to stimulate students to complete the Radar, they can only gain access to the
individual and average assessments of their peers after they have completed the assessment themselves. When all group members have completed their self assessments and peer assessments, two modified radar diagrams become available. The first - Information about yourself - shows the output of the self assessment (e.g., Chris about Chris) along with the average scores of the peer assessments of her/him (e.g., Group about Chris). The self assessment is not taken into account for computing the average scores. To provide more information about the variance in the average score of their peer assessment, students can also choose to view the individual peer assessments about their own behavior (e.g., Group members about Chris). The second - Information about the group (see Figure 2) - represents the average scores of the group members, so that group members can get a general impression about the functioning of the group.

All group members are represented as a solid line in the diagram, each with a different color. The student can include or exclude group members from the diagram by clicking a name in the legend.

--- INSERT FIGURE 2 ---

Reflection tool (Reflector). VCRI was also augmented with a reflection tool (Reflector) in order to stimulate group members to reflect and/or co-reflect on their individual behavior and overall group performance. This tool contained the five reflective questions discussed earlier:

1. What is your opinion on how the group functioned? Give arguments to support this.
2. What do you contribute to the functioning of the group? Give examples.
3. What do other group members think about your functioning in the group? Why do you think this?
4. What is your opinion on how you functioned in the group? Give arguments to support this.
5. What does the group think about its functioning in general? Discuss and formulate a
conclusion that is shared by all group members.

The first four questions are answered in Reflector, and completion is indicated by clicking an ‘Add’-button. This allows students to share their answers with the rest of the group and allows them to see the others’ answers. Students can only gain access to the answers of their peers after they have added their own answers, so as not to be influenced by one another. The fifth question is completed in the Co-Writer, which allows writing a shared conclusion. The responses made by the students in the Reflector are not scored or evaluated.

Data Analyses

First, to examine whether groups with Radar and Reflector show larger differences for self assessments and peer assessments than groups with only Radar between T1, T2 and T3, a paired samples t-test (one-tailed) with the dependent variables influence, friendliness, cooperation, reliability and productivity, are used to (1) compare the self assessment scores at T1, T2 and T3, and (2) compare the peer assessment scores at T1, T2 and T3. Differences between the self- and peer assessments at T1, T2 and T3 are analyzed using an independent t-test (two-tailed).

Second, to examine whether groups with Radar and Reflector show more congruency between self assessments and peer assessments than groups with only Radar, a Pearson product-moment correlation coefficient is used. We expect that peer assessments at T1 and T2 will affect self assessments at T2 and T3, therefore correlations will be calculated between peer assessments at T1, T2 and T3 and self assessments at T2 and T3.

Third, to examine whether groups with Radar and/or Reflector perceived better social and cognitive behavior than groups without these tools, a two way between-groups analysis of variance (ANOVA) (two-tailed) is conducted to explore the effect of Radar and/or Reflector on influence, friendliness, cooperation, reliability and productivity, as measured at T3 for both self assessment as peer assessment.
Fourth, to examine whether Radar and/or Reflector lead to higher social performance, a two way between-groups analysis of variance (one-tailed) is conducted with the dependent variables ‘team development’, ‘group satisfaction’, ‘level of group conflicts’, and ‘attitude towards collaborative problem solving’, as measured by the questionnaire at the end of the experiment.

Fifth, to examine whether Radar and/or Reflector lead to higher cognitive performance, a two way between-groups analysis of variance (one-tailed) is conducted with the grade on the essay as dependent variable.

Results

Self assessment scores. Table 3 shows the mean scores and standard deviations of self assessments at T1, T2 and T3 per condition. At T1 and T2 only groups with a Radar (condition +Ra~Rf and +Ra+Rf) could complete a self assessment. At T3 all conditions received and completed a self assessment. Except where noted, tests were one-sided. The rule of thumb (Kittler, Menard & Phillips, 2007) for effects sizes (η²) was small ≥ .01, medium ≥ .06, and large ≥ .14.

--- INSERT TABLE 3 ---

To examine whether groups with Radar and Reflector show higher discrepancies for self assessments between T1, T2 and T3, than groups with only a Radar, a paired samples t-test was used to compare the average self assessment scores at T1, T2 and T3, with respect to perceived social and cognitive behavior (influence, friendliness, cooperation, reliability and productivity). No significant differences between the first and the second assessment were found for groups with only a Radar (+Ra~Rf). In comparison with the second assessment, students perceived at T3 significantly more Reliability, \( t(7) = 2.53, p = .02, \eta^2 = .48 \). Compared with the first assessment, students perceived at T3 more Friendliness, \( t(7) = 3.10, p = .009, \eta^2 = .58 \), and more
Productivity, $t(7) = 2.55, p = .02, \eta^2 = .48$. For groups with both Radar and Reflector (+Ra+Rf) no significant differences were found between the average self assessment scores at T1, T2 and T3.

Independent $t$-tests comparing self assessment scores of conditions +Ra~Rf and +Ra+Rf at T1, T2 and T3 revealed that groups with only Radar (+Ra~Rf) perceived significantly more Influence at T3 than groups with both Radar and Reflector (+Ra+Rf); $t(13) = -2.33, p = .04$ (two tailed). The magnitude of the difference in means (mean difference = 9.71, 95% CI: .72 to 18.69) was large ($\eta^2 = .48$). No other significant differences were found.

**Peer assessment scores.** Table 4 shows the mean scores and the standard deviations of peer assessments at T1, T2 and T3 per condition. It was assumed that peer feedback provided by Radar would make group members aware of, possible, unrealistic positive perceptions of the performance of their peers. The expectation was that peer assessment scores would decrease at T2.

--- INSERT TABLE 4 ---

To examine whether groups with Radar and Reflector show greater differences than groups with only a Radar for peer assessments between T1, T2 and T3, a paired samples $t$-test (one-tailed) was used to compare average peer assessment scores at T1, T2 and T3 with respect to perceived social and cognitive behavior (i.e., influence, friendliness, cooperation, reliability and productivity).

No significant differences were found between the first and the second assessments for groups with only Radar (+Ra~Rf). Compared to the second assessment, students at T3 perceived significantly more Friendliness, $t(23) = 2.80, p = .01, \eta^2 = .25$, more Cooperativeness, $t(23) = 2.29, p = .02, \eta^2 = .19$, more Reliability, $t(23) = 2.62, p = .008, \eta^2 = .23$, and higher Productivity, $t(23) = 2.38, p = .01, \eta^2 = .20$. Compared with the first assessment, students
perceived at T3 significantly more Friendliness, \( t(23) = 3.27, p = .002, \eta^2 = .32 \), and more Productivity, \( t(23) = 4.33, p = .00, \eta^2 = .45 \).

Significant differences were found between the first and the second assessments for groups with both Radar and Reflector (+Ra+Rf). Compared to T1, students with Radar and Reflector at T2 perceived significantly less Influence, \( t(29) = -2.00, p = .03, \eta^2 = .06 \), less Friendliness, \( t(29) = -4.40, p = .00, \eta^2 = .25 \), and less Reliability, \( t(29) = -1.81, p = .04, \eta^2 = .05 \).

In comparison with the second assessment, students at T3 perceived significantly more Friendliness, \( t(29) = 2.05, p = .03, \eta^2 = .07 \), and more Reliability, \( t(29) = 1.88, p = .04, \eta^2 = .06 \). Compared to T1, students at T3 perceived significantly less Influence, \( t(29) = -2.15, p = .02, \eta^2 = .07 \).

**Comparing self assessments and peer assessments for groups with Radar.** An independent \( t \)-test (one-tailed) was used to examine the differences between self assessments and peer assessments at T1, T2 and T3 with respect to perceived social and cognitive behavior (i.e., influence, friendliness, cooperation, reliability and productivity). Tables 3 and 4 show the mean scores and standard deviations of self assessments and peer assessments per condition.

Students with only Radar (+Ra~Rf) perceived their peers at T1 as significantly more Reliable than themselves, \( t(30) = -2.02, p = .03 \). The magnitude of the difference in means (mean difference = -6.88, 95% CI: -13.82 to .07) was moderate (\( \eta^2 = .12 \)).

In comparing the other students in their team with themselves, students with both Radar and Reflector (+Ra+Rf) perceived their peers at T1 as being significantly more Friendly \( t(39) = -1.80, p = .04 \), with a moderate (\( \eta^2 = .08 \)) magnitude of the difference in means (mean difference = -6.22, 95% CI: -13.21 to .77) and as significantly more Reliable, \( t(39) = -2.05, p = .02 \), with a moderate (\( \eta^2 = .10 \)) magnitude of the difference in means (mean difference = -8.46, 95% CI: -16.80 to .13). No other significant differences were found between self- and peer assessments for condition +Ra~Rf and +Ra+Rf at T1, T2 and T3, or for condition ~Ra+Rf and
Examining congruency between self assessments and peer assessments. A Pearson product-moment correlation coefficient was used to test congruency between peer assessments at T1, T2, T3 and self assessments at T2 and T3 with respect to perceived social and cognitive behavior (i.e., influence, friendliness, cooperation, reliability and productivity).

Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity and homoscedasticity. Table 5 shows the Pearson correlations for peer assessments at T1, T2, T3, and self assessment at T2, T3.

--- INSERT TABLE 5 ---

There was a strong negative correlation for groups with only Radar (+Ra~Rf) between self assessment and peer assessment scores for Influence at T3, $r = -.81, n = 8, p = .01$, and between peer assessment scores at T2 and self assessments at T3 for Friendliness, $r = -.72, n = 8, p = .04$. A strong positive correlation was found between the self assessment scores for Reliability at T2 and peer assessments at T3, $r = .71, n = 8, p = .05$.

For groups with both Radar and Reflector (+Ra+Rf), there was a strong positive correlation between self assessment and peer assessment scores for Influence at T2, $r = .64, n = 11, p = .03$, and also at T3, $r = .69, n = 11, p = .02$. Peer assessment scores for Influence at T2 correlate strongly with self assessments at T3, $r = .81, n = 11, p = .00$, indicating a convergence of self and peer perceptions. Peer assessment scores for Friendliness at T3 correlate strongly with self assessments at T2, $r = .62, n = 11, p = .04$, and self assessments at T3, $r = .67, n = 11, p = .03$.

Peer assessment scores for Productivity at T2 correlate strongly with self assessments at T2, $r = .73, n = 11, p = .01$, and with self assessments at T3, $r = .62, n = 11, p = .04$, indicating a convergence of self and peer perceptions. A strong positive correlation was also found between
self assessment scores at T2 and peer assessments at T3, \( r = .72, n = 11, p = .01 \).

**Comparing peer assessment scores for all conditions at T3.** It was expected that at the end of the task (T3), groups with both Radar and Reflector (condition +Ra+Rf) would perceive more social behavior (e.g., less influence, more friendliness) and better cognitive behavior (e.g., more productivity), than groups with only Radar (+Ra~Rf), only Reflector (~Ra+Rf) or without either (~Ra~Rf). A two-way between-groups ANOVA was conducted to explore the effect of Radar and/or Reflector at T3 for both peer assessment and self assessment. Analysis of peer assessments showed no significant interaction between Radar and Reflector and no significant main effects. Analysis of self assessments showed no significant interaction or main effect for Radar, but did show a statistically significant main effect for Reflector on Influence, \( F(1, 35) = 4.54, p = .04 \) (two-tailed), partial \( \eta^2 = .12 \).

In an independent t-test comparing the self assessment scores on Influence for groups with and without Reflector, the ~Ra+Rf and +Ra+Rf conditions were combined. Groups with Reflector scored significantly higher on Influence \( (M = 68.08, SD = 11.48) \), than groups without, \( M = 76.00, SD = 10.90; t(37) = -2.16, p = .04 \). The magnitude of the difference in means (mean difference = -7.91, 95% CI: .49 to 15.33) was moderate \( (\eta^2 = .11) \).

**Impact of tools on social performance.** A two-way between-groups ANOVA was conducted to explore the effect of Radar and Reflector on social performance with respect to team development, group satisfaction, group conflicts and attitude towards collaborative problem solving. Participants were divided into four groups according to their condition (~Ra~Rf, +Ra~Rf, ~Ra+Rf, +Ra+Rf). There were no significant interaction effects between Radar and Reflector and no significant main effects for Reflector. There was a main effect for Radar on team development, \( F(1, 30) = 4.19, p = .05 \), partial \( \eta^2 = .12 \), level of group conflict, \( F(1, 31) = 4.49, p = .04 \), partial \( \eta^2 = .13 \), and attitude towards collaborative problem solving,
\[ F(2, 31) = 1.44, p = .04, \text{ partial } \eta^2 = .13. \]

An independent \( t \)-test was conducted to examine the main effects of Radar on team development, group conflict and attitude towards problem based collaboration. Conditions +Ra~Rf and +Ra+Rf were combined into a new group named ‘with Radar’, and conditions ~Ra+Rf and ~Ra~Rf were combined into group ‘without Radar’ (see Table 6).

--- INSERT TABLE 6 ---

The results in Table 6 show that groups with Radar (+Ra~Rf and +Ra+Rf) scored significantly higher on team development, \( t(32) = 1.79, p = .04 \), experienced significantly less group conflicts \( t(36) = -2.03, p = .03 \), and had a significantly more positive attitude towards collaborative problem solving, \( t(29) = 1.84, p = .04 \), than groups without Radar (~Ra+Rf and ~Ra~Rf).

**Impact of tools on cognitive performance.** A two way between-groups ANOVA was conducted to explore the effect of Radar and Reflector on group cognitive performance, as measured by the grade given to their essays. There were no significant interaction effects between Radar and Reflector, and no significant main effects for Radar or Reflector. Table 7 shows mean and standard deviations for cognitive performance per condition.

--- INSERT TABLE 7 ---

Discussion and Conclusion

The first aim of this study was to examine whether groups with peer feedback tool (Radar) and reflection tool (Reflector) showed larger differences for self assessments and peer assessments between T1, T2 and T3, than groups with only Radar. Based on Stroebe, Diehl, and Abakoumkin (1992), we assumed that group members would generally form unrealistically
positive perceptions of self performance and peer performance. Therefore, we expected that peer feedback provided by Radar at the first assessment (T1) would make group members aware of these perceptions, resulting in a decrease of self assessment and peer assessment scores at the second assessment (T2). Analysis of self assessment scores showed no significant decrease in scores at T2, but, as expected, analyses of peer assessment scores at T2 for groups with both Radar and Reflector (+Ra+Rf) showed a decrease in scores for Influence, Friendliness, and Reliability as compared to the first assessment (T1).

The second aim of this study was to determine whether the self assessments and peer assessments scores of groups with Radar and Reflector would be more similar (be more congruent) than the scores of groups with only Radar. We assumed that group members would adjust their unrealistic positive self perceptions towards more realistic perceptions between T2 and T3. Therefore, positive correlations were expected between the peer assessments at T2 and the self assessments at T3. As expected, the peer assessments of groups with both Radar and Reflector (+Ra+Rf) at T2 correlated strongly with the self assessments at T3 for Influence and Productivity, indicating a convergence of self and peer perceptions. For groups with only Radar (+Ra~Rf), a strong negative correlation was found between peer assessment scores at T2 and self assessments at T3 for Friendliness. This suggests that a combination of the two tools will make students more aware of their social and cognitive behavior during collaboration.

The third aim of this study was to enhance students’ social and cognitive behavior with Radar and/or Reflector. We assumed that at T3, groups with both Radar and Reflector (+Ra+Rf) would perceive more social behavior (e.g., less influence, more friendliness) and cognitive behavior (e.g., more productivity) than groups without Radar (~Ra+Rf), without Reflector (+Ra~Rf), or without both (~Ra~Rf). A two-way ANOVA of the self assessment scores showed a significant main effect for Reflector on Influence but not on other variables. Independent t-tests between groups with and without Reflector showed that groups with Reflector scored significantly higher on Influence than groups without Reflector, but no differences for the other
variables. This may be because Reflector’s questions make stress the individual *contribution* to group functioning and, thus, make group members especially aware of their influence rather than their friendliness, reliability or other characteristics. How one contributes (i.e., amount and value) can, possibly, be perceived as an influence on group functioning.

For the peer assessment scores, a two-way ANOVA did not show the expected main effects of both Radar and Reflector on Influence, Friendliness, Cooperativeness, Reliability, and Productivity at T3. This might be due to a tendency for all group members to assess their peers more positively than they normally would, because of the high level of group satisfaction based upon successful task completion. This is in line with Locke and Latham (1990) who found that group performance is strongly related to group satisfaction. This would also explain the increase of self assessment scores and peer assessment scores at T3 as compared to T2.

The fourth and fifth aims of this study were to examine the effects of Radar and/or Reflector on students’ perceived team development, group satisfaction, level of group conflict, attitude towards collaborative problem solving, and grade given for their essay. As expected, main effects were found for Radar on team development, group conflict, and attitude towards collaborative problem solving. However, no effects were found for group satisfaction and grade.

The lack of a significant main effect for Radar on group satisfaction is probably due to the short period of time in which the groups had to collaborate in order to accomplish the task. Deadlines and the task at hand can influence group development (Gersick, 1988). The short amount of time could ‘force’ group members to fulfill a role or task in which they do not feel comfortable or satisfied with. Changing circumstances, such as desired role-changes or disappointing level of task accomplishment, may cause group development to revert to the stage of *storming* (Bales & Cohen, 1979), which can be contentious, unpleasant, and even painful to group members who do not like conflicts (Tuckman & Jensen, 1977).

In the same vein, the period of time may be too short to find effects of the tools on cognitive performance. Therefore, further studies will examine the effects of Radar and Reflector...
during a longer period (i.e., three months) during which students collaborate on a complex learning task.

Several limitations of this study should be kept in mind. First, the statistical power of this study is low because of the relatively small sample size ($N = 39$). However, even with this small sample, significant main effects were found for Radar on team development, level of group conflict and attitude towards collaborative problem solving. Second, in this study Radar is both an intervention and a measurement tool for the two of the dependent variables (i.e., Influence, Friendliness). Therefore, the design did not allow us to determine whether the decrease of self assessment and peer assessment scores halfway collaboration at T2 was caused by Radar or Reflector, or whether this also occurred in the control group. In future studies, an extra control group will be added for which Radar will become available at T2.

Although the effects of this study are mainly ascribed to the Radar, we still assume that a combination of Radar and Reflector will be most effective. An explanation why no significant main effects for the Reflector on social group performance were found could be that Reflector focused here on past and present group functioning and not on future functioning, which might have resulted in superficial reflections which do not take future group behavior into account. In further studies, the Reflector will also focus on future group functioning. That is, it will also stimulate group members to formulate plans and set goals for improving social and cognitive group performance. Research has shown, for example, that outcome feedback can increase individual and group performance, especially when it is combined with goal setting (Mento, Steel, & Karren, 1987; Neubert, 1998; Tubbs, 1986). There is no reason that this should not also be the case for process feedback.

In conclusion, the effects of Radar on group functioning are very promising. They show that social group performance in CSCL environments, such as team development, level of group conflicts and attitude towards collaborative problem solving, can be enhanced by adding this easy to complete and easy to interpret peer feedback tool. For Reflector, it was argued that the
focus of the questions should be directed towards future group performance and goal setting.
References


Appendix: reviewers’ comments and authors’ revisions

Manuscript Number: CHB-D-09-00115
Manuscript Title: Awareness of Group Performance in a CSCL Environment: Effects of Peer Feedback and Reflection
First submission: May 4th, 2009

<table>
<thead>
<tr>
<th>Comment of the reviewer</th>
<th>Revisions of the authors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General comment</strong></td>
<td>I would like to express our gratitude for the helpful suggestions for revision given by the reviewer. The reviewer made a number of insightful recommendations that helped us improve the clarity and direction in the conceptual framework as well as the presentation of the results. For a detailed explanation and explication of the corrections based upon the comments of the reviewer see the following.</td>
</tr>
</tbody>
</table>

1. **Conceptual framework**

1b Intro (p. 2 line 10): Do Tuckman and Jensen say that groups must go through these stages in order to perform well…? Or do they just describe the observed stages in the group process? They write in 1977; did they speak specifically to CSCL, or is this the authors’ extrapolation?

   We agree with the reviewer that our text implies that group members must go through these stages of group development in order to become a well performing team. What we meant to say was that group development in face-to-face groups can be described in several stages, and it seems that the observed stages of Tuckman and Jensen also apply to CMC groups. In the section ‘Introduction’ (p. 2, lines 5-12) we have changed the text to reflect this.

1c The argument for this study hinges on the benefits of incorporating social process tools as a means of achieving “full potential”; it is important to define that term carefully. E.g., p. 3 line 55 says that “groups in CSCL environments do not always reach their full potential.”

   We have revised the conceptual framework by dividing the section ‘Social and Cognitive Processes in CSCL’ (see p. 3) into three sections, namely 1) ‘Social and Cognitive Processes in CSCL’, 2) ‘Social and Cognitive Benefits of Groups in CSCL-environments’ and 3) ‘Effects of CSCL Design on Social and Cognitive Behavior’. The revised section ‘Social and Cognitive Benefits of Groups in CSCL-environments’ (p. 4) describes not only the social and cognitive benefits of CMC groups compared to F2F groups, but also the basis for the claim that CMC groups do not always reach
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What is the basis for this claim? Who says so? In what way? Do non-computer-supported groups do better at this?</strong></td>
<td>their full potential, which is the fact that CMC groups seem to experience the same problems and difficulties that sometimes occur in FTF groups, such as social loafing or the free rider effect.</td>
</tr>
<tr>
<td><strong>1d</strong> Similarly, need to be more specific on what is meant by “learning processes” and “higher performance” (p. 2, lines 10-12)</td>
<td>We revised the section ‘Social and Cognitive Processes in CSCL’ (p. 3) in which we 1) describe the benefits of collaborative learning from a social and cognitive point of view, 2) describe what is meant with social and cognitive learning processes, 3) explain what is meant by social and cognitive performance, and 4) describe why it is important that CSCL environments support both cognitive and social processes.</td>
</tr>
<tr>
<td><strong>1e</strong> Need to expand somewhat on the specific benefits of computer-based collaborative learning (as opposed to non-computer-based) p. 3, lines 32-26</td>
<td>We have added a new section ‘Social and Cognitive Benefits of Groups in CSCL-environments’ (p. 4) in which we describe 1) social and cognitive benefits of CMC groups, and 2) problems that occur in both F2F groups and CMC groups, such as social loafing and free riding.</td>
</tr>
<tr>
<td><strong>1f</strong> You state that &quot;most CSCL-environments focus primarily on the support of cognitive or task-related processes, and limit possibility for social processes.&quot; (p. 4, line 5ff) Clarify why this is a problem.</td>
<td>At the end of the revised section ‘Social and Cognitive Processes in CSCL’ (p. 3) we describe why it is important that CSCL environments support cognitive processes as well as social processes. Additionally, in the added section ‘Effects of CSCL Design on Social and Cognitive Behavior’ (p. 5) we describe that most CSCL environments use text-based communication systems which limit the possibility for social processes to take place.</td>
</tr>
<tr>
<td><strong>1g</strong> (p. 5, line 55)ff: &quot;It is assumed that enhancement ...performance.&quot; This is a very broad assumption, but a very important one to this study. Given all the sources on group work you have cited elsewhere in the paper, you can make a much stronger statement by substantiating this statement with cites rather than the weaker &quot;it is assumed.&quot;</td>
<td>We agree that our statement could be more strongly formulated by substantiating this statement with cites. In section ‘Peer Feedback’ (p. 7) we have reformulated the sentence and added citations as requested.</td>
</tr>
<tr>
<td><strong>2. Research questions and measures</strong></td>
<td>In the section ‘Research questions’ (p. 10) we have changed the research questions accordingly.</td>
</tr>
</tbody>
</table>
2b Social performance measures (p. 11, line 39): are these instruments the same as the 60-item questionnaire referenced on p. 12, line 55?

This is an erroneous keystroke. In section ‘Task and procedure’ (p. 14) we have changed the text into: “Finally, all participants completed a 30-item questionnaire measuring the social performance of the group.”

2c In general, do the questions reflect how the users feel about collaborative problem solving in general (as Table 1 suggests)? Do other questions address how well they felt that they were able to approach the collaborative problem solving process? And/or how well the product turned out?

We agree with the reviewer that the ‘Attitude Towards Collaborative Problem-Solving Scale’ is not necessarily "self-evident" and we thank the reviewer for his reflective remarks. In section ‘Measures’ (p. 13) we have changed the text into: “The Attitude towards Collaborative Problem Solving scale (cf. Strijbos et al.) provides information on the perceived level of group effectiveness and how group members felt about working and solving problems in a group.”

3. Tasks and procedure

3a P. 12, line 36-48: groups with Radar used it at T1, T2, T3; "time on task was equal for all 4 conditions." Clarify to what task this refers. At T3 all groups used Radar, but what were the non-Radar groups doing at T1 and T2 while the Radar groups were using the tool?

In section ‘Task and procedure’ (p. 14) we have changed the text into: “While groups with Radar and/or Reflector completed the tools, groups without a Radar and/or Reflector continued working on their collaborative writing task. Groups with Radar and/or Reflector received extra time for their collaborative writing task so that time-on-task was equal for all four conditions.”

3b Figure 1 shows the VCRI interface with both Radar and Reflection. What did the interface look like for ~Radar, ~Refl, ~Radar~Reflection? What replaced those tools in the screen - blank space? Other tools? Larger versions of the other tools?

The interface of groups without Radar and/or Reflector displayed larger versions of the tools available.
In the section ‘Tools’ (p. 14) at the bottom of sub-section ‘Virtual Collaborative Research Institute (VCRI)’ we have added the text: “Windows of the available tools are automatically arranged on the screen, when students log in to the VCRI.”

3c How did the students using Radar learn about the categories on which they were to evaluate themselves or others? E.g., how did they know what "influence" was? When did instruction take place? How?

In the section ‘Task and procedure’ (p. 13) we have added a clarifying text. Additionally, in the section ‘Peer feedback tool (Radar)’ (p. 16) we have also added a clarifying text.

3d Were the responses made in Reflector evaluated/scored in any way?

Responses made in the Reflector were not evaluated or scored in any way. In the section ‘Reflection tool (Reflector)’ (p. 18) we have added the text:
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3e</strong></td>
<td>Did all students post feedback/respond to the questions/read feedback?</td>
<td>“The responses made by the students in the Reflector are not scored or evaluated.”</td>
</tr>
<tr>
<td><strong>4.</strong></td>
<td><strong>Data analysis</strong></td>
<td></td>
</tr>
<tr>
<td><strong>4a</strong></td>
<td>Did all students post feedback/respond to the questions/read feedback?</td>
<td>All students completed the tools because their teacher instructed them to do so. An extra stimulation for students to complete the Radar, is that they could only gain access to the (average) assessments of their peers after they have completed the Radar themselves. In the section ‘Peer feedback tool (Radar)’ (p. 17) we added a clarifying text.</td>
</tr>
<tr>
<td><strong>4b</strong></td>
<td>The data analysis is quite complex, perhaps unnecessarily so; some analyses are offered that do not seem to follow from the research questions, and having so many analyses actually obscures the results that do address the research questions. The analyses should be carefully examined and those that do not address the research questions should be omitted.</td>
<td>To help the reader keep things straight we added in section ‘Measurement’ (p. 12) a table titles ‘Overview of Scales, Subscales, and Instruments’. We have also revised the section ‘Results’ (p. 19 - 25) by consistently referring to the sub-scales at every measurement.</td>
</tr>
<tr>
<td><strong>4c</strong></td>
<td>The correlation between self and peer assessments is quite interesting (p. 21, line 43; table 7), but it is not an analysis that is addressed in the research questions (p. 9) or in the plan of analysis (p. 17). That being said, perhaps it should be added.</td>
<td>By rephrasing the research questions (see comment 2a) all analyses follow from the research questions. Nevertheless, we have deleted the following tables from the original manuscript: Table 3, ‘Paired Samples t-test between Self Assessments at T1, T2 and T3 (+Ra<del>Rf)’ Table 5, ‘Paired Samples t-test between Peer Assessments at T1, T2 and T3 (+Ra</del>Rf)’ Table 6, ‘Paired Samples t-test between Peer Assessments at T1, T2 and T3 (+Ra+Rf)’</td>
</tr>
<tr>
<td><strong>4d</strong></td>
<td>The correlation between self and peer assessments is quite interesting (p. 21, line 43; table 7), but it is not an analysis that is addressed in the research questions (p. 9) or in the plan of analysis (p. 17). That being said, perhaps it should be added.</td>
<td>In section ‘Research questions’ (p. 10) we have added research question to address this.</td>
</tr>
</tbody>
</table>
Awareness of Group Performance in a CSCL Environment:
Effects of Peer Feedback and Reflection

Chris Phielix¹, Frans J. Prins¹, Paul A. Kirschner¹²
¹Utrecht University, The Netherlands
²Open University of the Netherlands, The Netherlands

Corresponding author:
Chris Phielix
Research Centre Learning in Interaction
Utrecht University
PO Box 80140
3508 TC Utrecht, The Netherlands
c.phielix@uu.nl
Telephone: +31 (30) 253 4941
Fax: +31 (30) 253 2352
Acknowledgements

The authors would like to thank Jos Jaspers for the software programming and his technical assistance.
Figure captions:

Figure 1: Screenshot of VCRI with the six tools used in this experiment.

Figure 2. Radar - output group assessment
Figure 1. Screenshot of VCRI with the six tools used in this experiment.

Note: this figure is intended for black & white reproduction in print and color reproduction on the Web.
Awareness of Group Performance in CSCL

Note: this figure is intended for black & white reproduction in print.

See figure 2 on the next page for color reproduction on the Web.

Figure 2. Radar - output group assessment

Note: this figure is intended for black & white reproduction in print.

See figure 2 on the next page for color reproduction on the Web.
Note: this figure is intended for color reproduction on the Web

See figure 2 on the previous page for black & white reproduction in print.

Figure 2. Radar - output group assessment

Note: this figure is intended for color reproduction on the Web

See figure 2 on the previous page for black & white reproduction in print.
<table>
<thead>
<tr>
<th>Scale</th>
<th>Subscales</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social behavior</td>
<td>Influence, Friendliness, Cooperation, Reliability</td>
<td>Radar</td>
</tr>
<tr>
<td>Cognitive behavior</td>
<td>Productivity</td>
<td>Radar</td>
</tr>
<tr>
<td>Social performance</td>
<td>Team development, Group process Satisfaction, Intra-group Conflicts, Attitude towards Collaborative Problem Solving.</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>Cognitive performance</td>
<td>-</td>
<td>Essay grade</td>
</tr>
</tbody>
</table>
Table 2

*Examples of Social Performance Scales*

<table>
<thead>
<tr>
<th>Scale</th>
<th>$k$</th>
<th>Example</th>
<th>Cronbach’s $\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Development</td>
<td>10</td>
<td>Group members contribute ideas and solutions to problems.</td>
<td>.77</td>
</tr>
<tr>
<td>Group-process Satisfaction</td>
<td>6</td>
<td>I felt that my group worked very hard together to solve this problem.</td>
<td>.71</td>
</tr>
<tr>
<td>Intra-group Conflicts</td>
<td>7</td>
<td>I found myself unhappy and in conflict with members of my group.</td>
<td>.84</td>
</tr>
<tr>
<td>Attitude towards Collaborative Problem Solving</td>
<td>7</td>
<td>Collaborating in a group is challenging.</td>
<td>.74</td>
</tr>
</tbody>
</table>
Table 3

*Mean and Standard Deviations of Self Assessments per Condition*

<table>
<thead>
<tr>
<th>T</th>
<th>Condition</th>
<th>N</th>
<th>Influence</th>
<th>SD</th>
<th>Friendliness</th>
<th>SD</th>
<th>Cooperative</th>
<th>SD</th>
<th>Reliability</th>
<th>SD</th>
<th>Productivity</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+Ra~Rf</td>
<td>8</td>
<td>73.62</td>
<td>9.97</td>
<td>66.00</td>
<td>4.60</td>
<td>69.38</td>
<td>8.88</td>
<td>62.38</td>
<td>10.94</td>
<td>62.25</td>
<td>6.63</td>
</tr>
<tr>
<td></td>
<td>+Ra+Rf</td>
<td>11</td>
<td>73.09</td>
<td>9.18</td>
<td>68.91</td>
<td>9.69</td>
<td>70.09</td>
<td>16.59</td>
<td>62.27</td>
<td>11.33</td>
<td>67.18</td>
<td>13.42</td>
</tr>
<tr>
<td>2</td>
<td>+Ra~Rf</td>
<td>8</td>
<td>75.88</td>
<td>7.34</td>
<td>68.38</td>
<td>8.23</td>
<td>66.25</td>
<td>7.96</td>
<td>59.88</td>
<td>18.60</td>
<td>66.50</td>
<td>11.25</td>
</tr>
<tr>
<td></td>
<td>+Ra+Rf</td>
<td>11</td>
<td>69.27</td>
<td>13.45</td>
<td>64.00</td>
<td>8.61</td>
<td>67.09</td>
<td>12.64</td>
<td>60.00</td>
<td>21.37</td>
<td>61.82</td>
<td>19.10</td>
</tr>
<tr>
<td>3</td>
<td>+Ra~Rf</td>
<td>8</td>
<td>76.25</td>
<td>4.71</td>
<td>73.63</td>
<td>6.97</td>
<td>69.88</td>
<td>3.83</td>
<td>70.50</td>
<td>9.93</td>
<td>72.13</td>
<td>7.70</td>
</tr>
<tr>
<td></td>
<td>+Ra+Rf</td>
<td>11</td>
<td>66.54</td>
<td>12.68</td>
<td>68.46</td>
<td>16.98</td>
<td>69.63</td>
<td>13.43</td>
<td>62.09</td>
<td>17.84</td>
<td>65.09</td>
<td>15.29</td>
</tr>
<tr>
<td></td>
<td>-Ra~Rf</td>
<td>8</td>
<td>75.75</td>
<td>15.24</td>
<td>76.50</td>
<td>14.14</td>
<td>77.63</td>
<td>12.62</td>
<td>69.00</td>
<td>16.61</td>
<td>76.13</td>
<td>14.14</td>
</tr>
<tr>
<td></td>
<td>-Ra+Rf</td>
<td>12</td>
<td>69.50</td>
<td>10.62</td>
<td>67.58</td>
<td>22.15</td>
<td>70.92</td>
<td>5.81</td>
<td>68.33</td>
<td>16.81</td>
<td>73.33</td>
<td>11.69</td>
</tr>
</tbody>
</table>
## Table 4

*Mean and Standard Deviations of Peer Assessments per Condition*

<table>
<thead>
<tr>
<th>T</th>
<th>Condition</th>
<th>n</th>
<th>Influence</th>
<th>M</th>
<th>SD</th>
<th>Friendliness</th>
<th>M</th>
<th>SD</th>
<th>Cooperation</th>
<th>M</th>
<th>SD</th>
<th>Reliability</th>
<th>M</th>
<th>SD</th>
<th>Productivity</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>+Ra−Rf</td>
<td>24</td>
<td>74.17</td>
<td>6.70</td>
<td></td>
<td>66.42</td>
<td>7.22</td>
<td></td>
<td>69.88</td>
<td>6.27</td>
<td></td>
<td>69.25</td>
<td>7.35</td>
<td></td>
<td>63.33</td>
<td>8.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+Ra+Rf</td>
<td>30</td>
<td>75.17</td>
<td>11.51</td>
<td>75.13</td>
<td>9.85</td>
<td>73.07</td>
<td>12.77</td>
<td>70.73</td>
<td>11.81</td>
<td></td>
<td>68.27</td>
<td>13.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>+Ra−Rf</td>
<td>24</td>
<td>75.04</td>
<td>7.67</td>
<td></td>
<td>67.08</td>
<td>8.83</td>
<td></td>
<td>66.58</td>
<td>9.18</td>
<td></td>
<td>66.00</td>
<td>13.05</td>
<td></td>
<td>66.79</td>
<td>10.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+Ra+Rf</td>
<td>30</td>
<td>68.80</td>
<td>13.48</td>
<td>63.77</td>
<td>11.15</td>
<td>69.40</td>
<td>7.93</td>
<td>64.30</td>
<td>15.32</td>
<td></td>
<td>69.50</td>
<td>11.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>+Ra−Rf</td>
<td>24</td>
<td>75.29</td>
<td>6.95</td>
<td></td>
<td>71.63</td>
<td>7.81</td>
<td></td>
<td>71.04</td>
<td>5.04</td>
<td></td>
<td>72.08</td>
<td>6.98</td>
<td></td>
<td>71.00</td>
<td>7.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+Ra+Rf</td>
<td>30</td>
<td>70.07</td>
<td>8.80</td>
<td></td>
<td>69.60</td>
<td>17.09</td>
<td></td>
<td>69.40</td>
<td>11.36</td>
<td></td>
<td>69.83</td>
<td>12.23</td>
<td></td>
<td>66.73</td>
<td>12.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>−Ra−Rf</td>
<td>24</td>
<td>71.79</td>
<td>18.87</td>
<td></td>
<td>66.62</td>
<td>18.17</td>
<td></td>
<td>71.04</td>
<td>17.89</td>
<td></td>
<td>71.46</td>
<td>16.99</td>
<td></td>
<td>68.25</td>
<td>18.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>−Ra+Rf</td>
<td>36</td>
<td>70.28</td>
<td>12.87</td>
<td></td>
<td>66.14</td>
<td>17.41</td>
<td></td>
<td>68.19</td>
<td>16.87</td>
<td></td>
<td>69.53</td>
<td>17.87</td>
<td></td>
<td>68.28</td>
<td>14.02</td>
<td></td>
</tr>
</tbody>
</table>
Table 5

Pearson correlations for Peer Assessments at T1, T2, T3 and Self Assessment at T2, T3

<table>
<thead>
<tr>
<th>Condition</th>
<th>Assessment</th>
<th>Influence</th>
<th>Friendliness</th>
<th>Cooperative</th>
<th>Reliability</th>
<th>Productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>+Ra-Rf</td>
<td>Peer-1</td>
<td>-.47</td>
<td>-.04</td>
<td>.22</td>
<td>-.15</td>
<td>-.27</td>
</tr>
<tr>
<td></td>
<td>Peer-2</td>
<td>-.23</td>
<td>-.43</td>
<td>-.52</td>
<td>-.72*</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td>Peer-3</td>
<td>-.64</td>
<td>-.81*</td>
<td>.11</td>
<td>-.33</td>
<td>-.65</td>
</tr>
<tr>
<td>+Ra+Rf</td>
<td>Peer-1</td>
<td>.41</td>
<td>.50</td>
<td>-.01</td>
<td>.29</td>
<td>.47</td>
</tr>
<tr>
<td></td>
<td>Peer-2</td>
<td>.64*</td>
<td>.81**</td>
<td>.37</td>
<td>.53</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Peer-3</td>
<td>.51</td>
<td>.69*</td>
<td>.62*</td>
<td>.67*</td>
<td>-.05</td>
</tr>
</tbody>
</table>

* p < .05 (2-tailed)
** p < .01 (2-tailed)
Table 6

*Independent Samples t-test Between Groups With and Without Radar*

<table>
<thead>
<tr>
<th>Scale</th>
<th>Treatment</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Mean difference</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team development</td>
<td>with radar</td>
<td>16</td>
<td>4.08</td>
<td>.35</td>
<td>.26*</td>
<td>.04</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>without radar</td>
<td>18</td>
<td>3.82</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group satisfaction</td>
<td>with radar</td>
<td>17</td>
<td>3.95</td>
<td>.55</td>
<td>.00</td>
<td>.49</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>without radar</td>
<td>18</td>
<td>3.95</td>
<td>.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of group conflict</td>
<td>with radar</td>
<td>17</td>
<td>1.79</td>
<td>.37</td>
<td>-.38*</td>
<td>.03</td>
<td>.11</td>
</tr>
<tr>
<td></td>
<td>without radar</td>
<td>18</td>
<td>2.17</td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attitude towards collaborative problem solving</td>
<td>with radar</td>
<td>17</td>
<td>3.89</td>
<td>.39</td>
<td>.32*</td>
<td>.04</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>without radar</td>
<td>18</td>
<td>3.57</td>
<td>.62</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05 (1-tailed)
### Table 7

Mean and Standard Deviations for Cognitive Performance per Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>$M$</th>
<th>$SD$</th>
<th>$Min$</th>
<th>$Max$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Ra-Rf</td>
<td>6.00</td>
<td>.71</td>
<td>5.5</td>
<td>6.5</td>
</tr>
<tr>
<td>-Ra+Rf</td>
<td>5.83</td>
<td>1.04</td>
<td>5.0</td>
<td>7.0</td>
</tr>
<tr>
<td>+Ra-Rf</td>
<td>6.25</td>
<td>2.47</td>
<td>4.5</td>
<td>8.0</td>
</tr>
<tr>
<td>+Ra+Rf</td>
<td>5.17</td>
<td>1.61</td>
<td>4.0</td>
<td>7.0</td>
</tr>
</tbody>
</table>