Group Awareness Tools for Controversial CSCL Discussions: Dissociating Rating Effects and Visualized Feedback Effects

Jürgen Buder, Knowledge Media Research Center, j.buder@iwm-kmrc.de
Daniel Bodemer, University of Tübingen, Applied Cognitive Psychology and Media Psychology, Konrad-Adenauer-Str. 40, 72072 Tübingen, Germany, d.bodemer@iwm-kmrc.de

Abstract: An experimental study investigated how a group awareness tool impacts the social influence of a minority faction in controversial online group discussions. The awareness tool involved mutual ratings of contributions on dimensions that make minority opinions more salient. In order to dissociate between the potentially facilitative functions of rating interfaces and visualized feedback, a control condition was compared to a rate-only condition and a tool condition using visualized feedback about ratings. Results indicated that rating without feedback did not strengthen minority viewpoints, but rather widened the differences between minority and majority factions. The group awareness tool that used ratings and visualized feedback yielded moderate effects on minority influence, strong effects on the perceived group preference, and a more pronounced task focus. The results are discussed with regard to the relation of group awareness to different types of social influence and different types of self-awareness.

Introduction

Collaborative learning is first and foremost an activity among peers. Through communication, peers try to achieve shared understanding and to collaboratively construct or re-construct knowledge. Consequently, the role of a teacher fundamentally changes in contexts of collaborative learning. In face-to-face (FTF) collaborative learning, teachers often provide subtle cues to peer communication by providing guiding questions, by eliciting participation, or by applying meta-cognitive strategies of planning and monitoring the interaction. In contexts of purely computer-supported collaborative learning, however, it is not uncommon that teachers do not participate at all in the interaction among learners (e.g. in informal learning scenarios). In these cases, technologies must be designed in order to guide collaboration. There are several ways of how this can be accomplished, e.g. through the use of scaffolds (Scardamalia, 2002) or collaboration scripts (Kollar, Fischer, & Hesse, 2006).

In the last few years the repertoire of CSCL technologies that provide guidance to peer activities has been extended by so-called group awareness tools. The notion of group awareness and the development of tools to foster group awareness have originated in the field of computer-supported cooperative work (Gutwin & Greenberg, 1995). Group awareness tools are technologies that register information about a group, its members and its products, aggregate this information and feed it back to the group members. Originally, group awareness tools were designed to address shortcomings of spatially or temporally distributed group activities. For instance, they sought to re-create the richness of FTF interaction by providing information about the presence of group members in a shared workspace, or by indicating which group member is currently working on which document. Mimicking FTF, however, has become less important once group awareness tools began to be explored in CSCL contexts. One reason for this shift in the conceptualization of CSCL group awareness tools was the general consensus that technologies should provide an added value over FTF scenarios in order to make CSCL justifiable (Buder, 2007). That is, group awareness tools for CSCL should do more than just passively register and feed back "what's going on"; ideally, they should provide guidance to a group and its members. These motivations led to the general idea to develop tools that provide information about a group that would be difficult or even impossible to yield in face-to-face contexts.

For instance, some group awareness approaches in CSCL are based on informing learners about their levels of participation (Janssen, Erkens, & Kirschner, 2011), or require learners to rate the behavior of their collaborators (Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011). However, the most frequently used method involves providing information about the knowledge of collaborators (knowledge awareness tools; Engelmann, Dehler, Bodemer, & Buder, 2009). Knowledge awareness can be conceptualized in very different ways: e.g. by requiring learners to externalize their knowledge by creating concept maps prior to collaboration (Engelmann & Hesse, 2010); by requiring learners to explicitly rate their level of understanding with regard to pieces of learning material (Dehler, Bodemer, Buder, & Hesse, 2011); by making the results of a prior knowledge test available to collaborators at the beginning of interaction (Sangin, Molinari, Nüssli, & Dillenbourg, 2011); or by constraining ongoing collaborative interaction in ways that make differences among learner knowledge visible (Bodemer, 2011). In all these cases, an element that can only be indirectly inferred in FTF interaction (viz. knowledge) is made explicit and salient, thereby guiding collaborative processes. A common finding in the CSCL studies on group awareness is that tools unfold their power by feeding back information about differences among learners. If levels of understanding are different, issues can be resolved by one learner providing explanations to a collaborator; if collaborators differ in how they understand elements of a learning material, they can focus on these conflicting issues and negotiate on a shared understanding.

However, in some learning scenarios resolving conflict might become more complex. This is the case for differences in viewpoints or opinions. An opinion does not relate to isolated arguments, but rather emerges from evaluating and forming a social judgment on a whole set of arguments. This lends a social psychological dimension to collaboration. Conflicts cannot be resolved through explanations or through negotiation of single arguments, but rather involve persuasive communication on how to evaluate a whole set of arguments. Resolving differences in opinion is tricky, as it is likely that individuals process information in a somewhat biased way. For instance, individuals exhibit confirmation bias, a tendency to disregard dissenting information (Jonas, Schulz-Hardt, Frey, & Thelen, 2001). Similarly, in collaborative contexts there is a general tendency for groups not to take the full variability among members into account (Hinsz, Tindale, & Vollrath, 1997). For instance, groups favor shared over unshared information even in contexts where the consideration of unshared information would lead to better group performance (Stasser, 1992). Another facet of variability reduction in groups appears in conflicts between factions of different size, i.e. majorities and minorities. Asch (1951) was the first to show that minority members tend to conform to incorrect viewpoints when they are advocated by a majority.

In order to address these issues, a group awareness tool for resolving different opinions in a group must touch on social psychological dimensions. This idea was at the core of a study about so-called augmented group awareness tools (Buder & Bodemer, 2008). They developed and tested a group awareness tool that tried to resolve learners' different opinions on two conflicting physics hypotheses. Their study was modeled after the informed minority paradigm (Stewart & Stasser, 1998), and it involved 4-person groups where three learners were in favor of a scientifically incorrect hypothesis, and one learner (the informed minority) was in favor of a scientifically correct hypothesis. Learners were required to come to a consensual decision on one of the hypotheses after a 30-minute online forum discussion. The results of this study have shown that minorities tended to conform to the scientifically incorrect majority opinion. However, half of the groups in this study were supported by a group awareness tool that made minority contributions particularly salient. The tool required learners to rate their discussion contributions on two dimensions, viz. agreement with a contribution and perceived novelty of a contribution during discussion. The tool aggregated these ratings, computed average ratings for each contribution, and fed back these results in a visualization where each contribution was represented as a dot in a two-dimensional coordinate system. It was shown that this augmented group awareness tool strengthened minority influence, as groups arrived more often at the scientifically correct minority viewpoint.

While this tool was effective in strengthening minority influence, the underlying mechanisms of the effects deserve further investigation. The augmented group awareness tool rests on two components, both of which might have contributed to its effectiveness. The first component is related to the rating activities. It can be argued that rating a contribution requires reflective thought that would not have occurred to the same degree without rating. In this case, the requirement to rate might have served as a valuable meta-cognitive prompt (Kramarski & Mevarech, 2003) inducing group awareness. The second component that might have contributed to the effectiveness of the tool is the visualized feedback. Only through the visualization, learners could see how the group as a whole thought about the discussion contributions, and only through the visualization minority contributions were becoming salient (lower average agreement and higher average novelty ratings than majority contributions). In order to dissociate between these two components, the study presented here attempted to extend the findings of Buder and Bodemer (2008) by including an experimental condition where learners were asked to rate contributions, but these ratings were neither aggregated, nor fed back to the group members.

It was hypothesized that the components of a rating interface and of visual feedback about ratings affect performance additively. That is, individuals in rate-only groups (with rating, without visualization) should outperform individuals in unsupported groups (no rating, no visualization), and individuals in groups using the complete awareness tool (with rating, with visualization) should outperform individuals in rate-only groups. The ordering of conditions was expected with regard to the following dependent variables:

Hypothesis 1a. Post-discussion preferences leaning towards the minority viewpoint
Hypothesis 1b. Strong minority influence for majority participants (pre- vs. post-discussion)
Hypothesis 1c. Weak majority influence for minority participants (pre- vs. post-discussion)
Hypothesis 1d. Perceived group preferences learning towards the minority viewpoint
Hypothesis 2a. Higher performance in a knowledge test
Hypothesis 2b. Higher knowledge test performance on preference-inconsistent items
Hypothesis 3. Higher rates of discussion focusing on the exchange of arguments
Hypothesis 4. Higher salience of minority contributions through lower agreement ratings, but higher novelty ratings.

To test these predictions, a laboratory experiment was conducted.

Method

The basic setup of this study was adapted from Buder and Bodemer (2008), but the learning domain and the group size were altered. Additionally, this study contained a knowledge test.

Design and Participants

For testing the hypotheses, a one-factorial design with three conditions was employed. In the control condition, groups were using the discussion environment without ratings or visualizations. In the rate-only condition, group members rated discussion posts, but these ratings were not aggregated or fed back to the entire group. In the tool condition, the fully functional group awareness tool was available.

87 student participants (60 female, 27 male; *M* = 24.84 years) were randomly assigned to experimental conditions, and within three-person groups they were randomly assigned to the minority position or one of the two majority positions. Eventually, 29 groups were taking part in the experiment (9 in the control condition; 10 in the rate-only condition; 10 in the tool condition). Subjects were paid for participation, or received course credit. Students of biology or geology were excluded from participation. Prior knowledge of the learning domain was uniformly low across participants (*M* = 2.67, *SD* = 1.24, on a scale ranging from 1 through 7).

Materials

Instructional Material

Instructional material consisted of an introductory text and sets of arguments covering the event that caused the extinction of dinosaurs about 65 million years ago. The introductory text described some basics about the event, and introduced two competing hypotheses about the cause of extinction (meteorite impact vs. long-term volcanism). A pool of ten arguments was created that provided evidence for these hypotheses. Arguments were printed on separate index cards and had an average length of about 160 words. Six of these arguments were in favor of the volcanism hypothesis, whereas only four arguments were in favor of the meteorite hypothesis, thus making the volcanism hypothesis superior. At the time of testing, the actual merits of both hypotheses were still hotly debated among geologists and paleontologists. However, popular scientific accounts had much stronger coverage of the meteorite hypothesis, thereby ensuring a conservative testing of social influence effects. Prior tests had shown that nine out of ten subjects who received the whole set of arguments were favoring the volcanism hypothesis afterwards. For the experiment, two sets of arguments were created. One set was identical for all majority members and consisted of all four arguments favoring the meteorite hypothesis plus two arguments favoring the volcanism hypothesis. The set for the minority members consisted of all six arguments favoring the volcanism hypothesis plus two arguments favoring the meteorite hypothesis.

Online Discussion Environment

The online discussion environment was developed at the Knowledge Media Research Center in Tübingen. The environment consists of a temporally ordered list of separate posts. Author names are anonymized. Depending on experimental conditions, a rating interface and an awareness tool were available. The rating interface was implemented as two sliders attached to each discussion post except one's own. One slider was designated to express novelty ratings; the other was designated to express agreement ratings. The ratings on the sliders were expressed as numbers between 0 and 100. The awareness tool contained a visualization of the discussion posts represented as dots on a two-dimensional graph, where the x-axis represented the average agreement rating, and the y-axis represented the average novelty rating that a given contribution received. The visualization was personalized in that learners could distinguish their own contributions from other group members' contributions, and by indicating contributions a learner hadn't rated yet (see Figure 1). By clicking on a particular dot in the visualization the discussion window automatically scrolled to the selected discussion post.

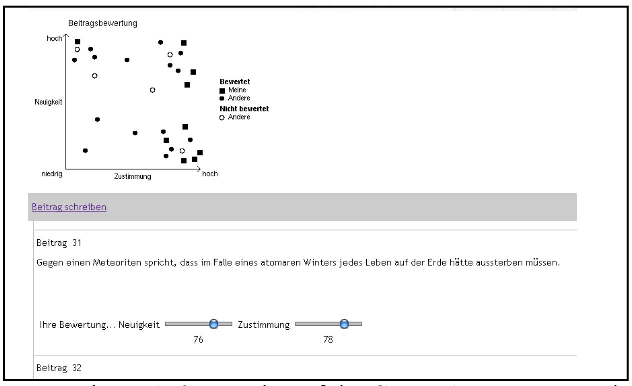

Figure 1. Screenshot of the Group Awareness Tool.

Measurements

In order to test the general predictions concerning the effectiveness of ratings and visualizations, four classes of dependent measures were analyzed. The first two classes refer to outcome variables (preference data, knowledge test data); the third and fourth classes refer to process variables (participation data, rating data).

Preferences were captured by measuring learner ratings on corresponding sliders ranging from -100 ("complete agreement with meteorite hypothesis") to +100 ("complete agreement with volcanism hypothesis"). Sliders were used to capture individual pre-preference, individual post-preference and perceived group preference. Minority influence and majority influence were measured by computing the difference from pre- to post-preference in the direction of the scale midpoint.

Knowledge test performance was measured as number of correct responses in a 10-item multiple-choice test (with one target and three distractors, respectively).

Participation data were captured both through objective data (number of posts) and through content analysis. For the content analysis, all posts were coded by two independent raters. Among other things, it was coded for each post whether it contained an explanation of an argument (Cohen's *kappa* = .88).

Finally, *rating* data on average agreement and average novelty were captured for each post in the rate-only condition and the tool condition.

Procedure

The experiment consisted of three phases: an individual learning phase, a group discussion phase, and an individual knowledge test phase. During the entire experiment subjects of a group were seated in separate rooms. In the first phase learners individually worked through the learning material on the dinosaur extinction event (20 minutes). After that, individual pre-discussion preferences were measured. After the learning phase, group members were given the opportunity to test the online discussion environment by writing some contributions. In the rate-only and tool conditions, participants were asked to rate test contributions by other learners. In the tool condition, the functionality of the awareness tool was explained during this stage.

In the second phase index cards containing the arguments were removed, and groups were instructed to discuss the conflicting hypotheses using the online environment. All learners were made aware that other group members might have received different pieces of evidence. Groups were asked to come to an agreement about the conflicting hypotheses within the allotted discussion time (30 minutes). According to the experimental design of the study groups in the control condition were only provided with the online discussion environment. Groups in the rate-only condition were additionally asked to rate contributions of their collaborators on agreement and novelty (e.g. low ratings for arguments that are mentioned repeatedly) by using two sliders ranging from 0 to 100. Groups in the tool condition also used the rating interface, but were additionally provided with visualized feedback. After the discussion phase individual learners were asked to indicate their post-discussion preference and their perceived overall group preference.

In the third phase, participants individually worked through the multiple-choice knowledge test. Subjects were briefed about the study at the end of the experiment.

Results

Means and standard deviations for *preference data* are listed in Table 1. The experimental manipulation was checked using a two-factorial analysis of variance (ANOVA) with condition (control, rate-only, tool) and status (majority, minority) as factors and *pre-discussion preference* as dependent variable. As expected, it yielded no main effect for condition ($F = 0.24$, $p = .78$, $\eta^2 = .00$); a main effect for status: $F(1, 81) = 399.90$, $p < .01$, $\eta^2 = .73$; and no interaction effect: $F(2, 81) = 1.55$, $p = .22$, $\eta^2 = .01$. This indicates that the manipulation worked properly.

Table 1: Means and standard deviations for condition and status with regard to preference data.

Condition	Status	Pre-preference	Post-preference	Influence	Group preference
Control	Majority	-52.3 (24.2)	-38.6 (49.7)	13.8 (53.3)	-39.9 (52.6)
Rate-only	Majority	-61.1 (25.0)	-64.6 (39.8)	-3.5 (31.0)	-46.4 (39.0)
Tool	Majority	-61.7 (29.7)	-40.5 (45.2)	21.3 (47.2)	-11.7 (45.6)
Control	Minority	46.8 (16.8)	30.9 (28.8)	15.9 (27.7)	-44.8 (49.5)
Rate-only	Minority	62.2 (19.3)	69.5 (22.8)	-7.3 (27.8)	-28.3 (25.5)
Tool	Minority	53.6 (24.3)	53.8 (26.3)	-0.2 (17.6)	-3.5 (52.7)

Post-discussion preferences indicate in how much learners tended towards the correct minority viewpoint (Hypothesis 1a). These analyses were conducted using the same 3 (condition) x 2 (status) analysis of variance as for pre-discussion preferences. The ANOVA did not yield the expected main effect for condition: $F(2,81) = 0.47$, $p = .63$, $\eta^2 = .00$); the strong main effect for status remained: $F(1, 81) = 126.05$, $p < .01$, $\eta^2 =$

.54. However, there was a significant interaction effect: $F(2, 81) = 4.50$; $p = .01$, $\eta^2 = .04$. Additional analyses revealed that this interaction was due to the large majority-minority spread in the rate-only condition. It appears that rating without subsequent visualized feedback actually enforces initial preferences of both minorities and majorities.

Minority influence (Hypothesis 1b) was measured using a one-factorial ANOVA for majority participants with regard to difference between post- and pre-preferences. The data reveal that minority influence was strongest in the tool condition, but the effect failed to reach significance: $F(2, 55) = 1.62$; $p = .21$, $\eta^2 = .05$.

A similar analysis was conducted for (undesirable) *majority influence* (Hypothesis 1c), based on minority participants. Majority influence was strongest in the control condition, but again, the effects were not strong enough to yield significance: $F(2, 26) = 2.17$; $p = .13$, $\eta^2 = .14$.

Participants were also required to express *perceived group preference* (Hypothesis 1d). The corresponding analyses were conducted with a 3 (condition) x 2 (status) analysis of variance. Here, a significant main effect for condition could be obtained: $F(2, 81) = 4.41$, $p = .02$, $\eta^2 = .07$. A post-hoc test revealed that participants in the tool condition perceived the group decision to be closer to the minority viewpoint than participants in the other two conditions. Neither a status effect ($F(1,81) = 0.46$, $p = .50$, $\eta^2 = .00$) nor a significant interaction ($F(2,81) = 0.39$, $p = .68$, $\eta^2 = .00$) were obtained.

Table 2: Means and standard deviations for condition and status with regard to knowledge test performance, participation, and rating data.

Measurement	Status	Control	Rate-Only	Tool
Knowledge test overall	Majority	.73 (.21)	.63 (.14)	.68 (.11)
	Minority	.73 (.10)	.77 (.15)	.79 (.11)
Knowledge on preference-inconsistent items	Majority	.66 (.24)	.54 (.19)	.58 (.18)
	Minority	.52 (.29)	.70 (.19)	.63 (.11)
Number of contributions	Majority	17.9 (9.8)	12.8 (5.6)	11.2 (5.8)
	Minority	22.0 (5.9)	13.9 (5.3)	11.1 (3.2)
Relative number of explanations	Majority	.29 (.14)	.44 (.19)	.56 (.21)
	Minority	.43 (.19)	.57 (.18)	.74 (.22)
Received agreement ratings	Majority		61.8 (14.4)	68.1 (14.7)
	Minority		43.1 (11.3)	50.7 (14.0)
Received novelty ratings	Majority		56.3 (15.8)	50.3 (16.4)
	Minority		67.9 (11.8)	65.0 (12.1)

A secondary outcome variable was the *knowledge test* (see Table 2). For *overall performance* (Hypothesis 2a), no significant main effect for condition could be found: $F(2,81) = 0.38$, $p = .68$, $\eta^2 = .00$. There was a very small, but significant main effect for status: $F(1, 81) = 5.74$, $p = .02$, $\eta^2 = .00$, with minority members showing better test performance, but this was probably due to the fact they initially received more information about the arguments than majority members. The condition x status interaction was $F(2, 81) = 1.42$, $p = .25$, $\eta^2 = .00$.

As to subsets of the knowledge test, additional analyses were conducted with the performance on *preference-inconsistent items* (Hypothesis 2b). While no main effects for condition ($F(2, 81) = 0.16$, $p = .85$, $\eta^2 = .00$) or status ($F(1, 81) = 0.31$, $p = .58$, $\eta^2 = .00$) were observed, the data revealed a significant interaction: $F(2, 81) = 3.43$, $p = .04$, $\eta^2 = .01$. In the control condition, majority members outperformed minority members, whereas this pattern was reversed in the rate-only condition. The tool condition yielded similar performance levels for majorities and minorities. As the control condition had the strongest majority influence, and the rate-only condition did not show any minority influence (see Table 1), the results hint at the possibility that performance on preference-inconsistent items might be related to patterns of social influence. Additional analyses revealed that over all three conditions, majority influence was negatively correlated with performance on preference-inconsistent items ($r = -.10$, $p = .60$), whereas minority influence was positively correlated with performance on preference-inconsistent items ($r = .24$, $p = .07$). While these correlations were non-significant, a tendency could be observed that being influenced by a majority was associated with less learning of majority-related concepts, whereas being influenced by a minority led to higher learning of minority-related concepts.

The *participation data* indicated that subjects in the control condition wrote more *contributions* than group members in the other conditions: $F(2, 81) = 12.11$, $p < .01$, $\eta^2 = .05$. Neither a main effect for status – $F(1, 81) = 1.34$, $p = .25$, $\eta^2 = .05$ - nor a condition x status interaction was found ($F(2, 81) = 0.66$, $p = .52$, $\eta^2 = .00$). A more detailed result was revealed through the analysis of the *relative amount of explanations* (Hypothesis 3). Here, a significant main effect for condition was found: $F(2, 81) = 15.07$, $p < .01$, $\eta^2 = .04$. Post-hoc tests revealed that in the tool condition, significantly more explanations occurred than in the rating-only condition, and in both these conditions, more explanations occurred than in the control condition. Moreover, minority members produced a higher rate of explanatory posts than majority members; $F(1,81) = 12.09$, $p < .01$, $\eta^2 = .02$. The interaction was non-significant: $F(2,81) = 0.17$, $p < .84$, $\eta^2 = .00$.

Finally, *rating data* for the rating-only and the tool condition were analyzed (Hypothesis 4). About 90% of contributions were rated by participants in both conditions. For the *agreement ratings*, a marginally significant main effect for condition was found: $F(1, 56) = 3.28$, $p = .08$, $\eta^2 = .00$, indicating that participants in the tool condition expressed slightly higher agreement with the discussion posts than participants in the rate-only condition. As expected, majority contributions generally received higher agreement ratings than minority contributions; $F(1, 56) = 22.29$, $p < .01$, $\eta^2 = .02$. No interaction effect was obtained: $F(1, 56) = 0.03$, $p = .87$, $\eta^2 = .00$. As for *novelty ratings*, conditions did not differ significantly from each other: $F(1, 56) = 1.18$, $p = .28$, $\eta^2 = .00$; however, minority contributions received higher novelty ratings than majority contributions: $F(1, 56) = 10.44$, $p < .01$, $\eta^2 = .01$. Again, the interaction was non-significant: $F(1, 56) = 0.15$, $p = .70$, $\eta^2 = .00$.

Discussion

The present study investigated the influence of rating interfaces and visualized feedback of an awareness tool on individual preferences, perceived group preferences, and individual knowledge test performance. It was expected that ratings serve as meta-cognitive prompts that improve outcome variables. Moreover, a visualized feedback about average ratings through an awareness tool was expected to enhance these positive effects. The predictions were tested in an experimental design involving three conditions (control groups, rate-only groups, tool groups).

The preference data yielded rather mixed results. For post-discussion preferences it was not found that participants in the tool condition were leaning stronger towards the minority viewpoint. Patterns of social influence revealed that unsupported groups exhibited strong majority influence and a moderate minority influence. In the tool condition, majority influence was non-existent whereas minority influence was highest among conditions. While this result is in line with predictions, the effects were too weak to reach significance. In contrast, rate-only groups showed neither minority influence nor majority influence. Rather, requesting repeated ratings without any feedback appeared to strengthen initial individual preferences. For this reason, the explanatory mechanism of meta-cognitive stimulation through ratings must be ruled out for this scenario. The data on perceived group preferences are more encouraging. Participants that were supported by an awareness tool estimated the average preference of their groups stronger in favor of the volcanism hypothesis than participants in the other two conditions. This can be interpreted as a higher awareness for the minority opinion. However, it should be noted that in this scenario the meteorite hypothesis was generally deemed much stronger. Only in one out of nine control groups, one out of ten rate-only groups, and three out of ten tool groups the averages of perceived group preferences tended towards the volcanism hypothesis. This might be due to the fact that in popular scientific accounts of the extinction event the meteorite hypothesis has received much higher coverage.

The knowledge test data did not yield overall differences among the three conditions, thus the hypothesis that ratings and rating visualizations increase performance cannot be confirmed. The detailed results for test performance on preference-inconsistent items revealed some interesting effects: in conjunction with the data on social influence pattern it was found that majority influence was associated with lower performance on majority test items. In other words, minorities might shift towards the majority opinion, but learn relatively little about that opinion. Such a pattern could be interpreted as normative social influence, an unthinking adoption of the majority viewpoint. In contrast, minority influence was associated with higher performance on minority test items. This indicates that majority members who shift towards the minority opinion learn more about that opinion. This higher performance might be due to informational social influence, an effect that causes majorities to scrutinize minority viewpoints more carefully (Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994).

The participation data can be interpreted in a way that ratings and visualized feedback about ratings both improve the learning process. The overall reduction of posts in the rate-only and tool conditions was probably due to the fact that rating and/or using the visualization takes time, thereby leading to lower productivity. However, this was offset by a higher task focus. In the tool condition, the rate of explanatory posts was significantly higher than in the rate-only condition, and the latter condition in turn yielded a higher rate of explanatory posts than the control condition. Additional analyses revealed that control groups generated a much higher amount of posts pertaining to task coordination and off-topic talk. A possible interpretation for this pattern is that rating activities subtly structure individual learning processes, thereby reducing the need for explicit coordination among collaborators.

Rating data might help to illuminate the social influence and deliberation processes of rate-only groups vs. tool groups. While the rating behavior of rate-only groups did not differ much from tool groups, the data indicate that minority contributions were rated lower on agreement, but much higher on novelty than majority contributions. As a consequence, minority contributions became visually salient in the tool condition which is exactly what the augmented group awareness tool tried to capture. Moreover, it was found that tool groups expressed slightly higher agreement with the discussion posts which might have contributed to an atmosphere where majority members scrutinized minority contributions more carefully and thus experienced stronger minority influence. It might be the case that additional analyses on subsets of posts will yield even stronger effects. In the current study, participants were asked to rate each discussion post, and it can be argued that agreement with a question cannot be interpreted in the same way as agreement with an answer. This also begs the question of whether the effectiveness of awareness tools and the clarity of experimental results can be improved by instructing collaborators to rate only those contributions that they regard as essential for the learning process.

Taken together, the results indicate that a rating interface per se is not sufficient to improve collaborative learning. The fact that social influence was virtually non-existent in the rate-only condition suggests that rating without feedback does not foster reflection on other group members and their products. On the contrary, it might be the case that repeated ratings direct attention to the self, thereby strengthening initial preferences. In other words, rating without feedback might not evoke group awareness, but *private self-awareness*, a tendency to adhere to *personal* standards of behavior (Froming, Walker, & Lopyan, 1982).

The results also indicate that visualized feedback about ratings can lead to social influence. In a scenario where the minority opinion is associated with better outcomes, the tool condition led to a moderate minority influence, and practically no majority influence. Moreover, participants in the tool condition perceived the group preference to be more shifted towards the minority opinion, exhibited more explanatory behavior, and expressed slightly higher agreement with the contributions of others. In this regard, it can be said that group awareness was achieved through the tool. However, since the visualization also provided feedback about oneself, it can be argued that the awareness tool also increased *public self-awareness*, the tendency to adhere to *social* standards of behavior (Froming et al., 1982).

Conclusions

The objectives of the present study were twofold. First, further evidence for the effectiveness of group awareness tools to foster computer-supported collaborative learning should be obtained. Social influence patterns were weaker than in the study by Buder and Bodemer (2008). This might be due to the selected learning domain, as argumentation skills can vary considerably across different domains (Mason & Scirica, 2006). Moreover, inducing preference change is difficult in a limited time frame. Another explanatory mechanism for the differences between the original study and the replication study could be subtle differences in the framing of the task that led to a differentiation-focused debate mode rather than an integration-focused controversy mode (Johnson & Johnson, 1979). Future studies could address how the awareness tool used here would work in a setting where the task focus is on open discussion rather than on joint decision making. Moreover, it would be interesting to see how the tool influences collaborative learning in a scenario where majority and minority factions are either non-existent or randomly distributed.

There have been many empirical studies showing the effectiveness of group awareness tools in CSCL scenarios. However, relatively little work has been done to uncover the mechanisms of group awareness (Buder, 2011). The second objective of the present study was to contribute to our understanding of *how* these tools actually work. This was attempted through the inclusion of an experimental rate-only condition that was not inspired by principles in the learning sciences, but served mainly to dissociate rating effects from visualized feedback effects. The results do not only indicate that isolated ratings are detrimental to collaborative learning, but also help to establish links between normative majority influence, informational minority influence, private self-awareness, public self-awareness, and group awareness.

References

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership, and men* (pp. 177-190). Pittsburgh: Carnegie Press.

Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior, 27*(3), 1079-1086.

Buder, J. (2007). Net-based knowledge-communication in groups: Searching for added value. *Zeitschrift für Psychologie/Journal of Psychology 215*(4), 209-217.

Buder, J. (2011). Group awareness tools for learning: Current and future directions. *Computers in Human Behavior, 27*(3), 1114-1117.

Buder, J., & Bodemer, D. (2008). Supporting controversial CSCL discussions with augmented group awareness tools. *International Journal of Computer-Supported Collaborative Learning, 3*(2), 123-139.

Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior, 27*(3), 1068-1078.

Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior 25*(4), 949-960.

Engelmann, T., & Hesse, F.W. (2010). How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning, 5*(3), 299-320.

Froming, W. J., Walker, G. R., & Lopyan, K. J. (1982). Public and private self-awareness: When personal attitudes conflict with societal expectations. *Journal of Experimental Social Psychology, 18*(5), 476-487.

Gutwin, C., & Greenberg, S. (1995). Support for group awareness in real-time desktop conferences. In: *Proceedings of the Second New Zealand Computer Science Research Students' Conference.* (University of Waikato, Hamilton, New Zealand), April 18-21, 1995.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualizion of groups as information processors. *Psychological Bulletin, 121*, 43-64.

Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It's what you do with it that matters. *Computers in Human Behavior, 27*(3), 1046-1058.

Johnson, D. W., & Johnson, R. T. (1979). Conflict in the classroom: Controversy and learning. *Review of Educational Research*, *49*(1), 51-69.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, *80*(4), 557-571.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts - a conceptual analysis. *Educational Psychology Review, 18*(2), 159-185.

Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: Effects of cooperative learning and metacognitive training. *American Educational Research Journal, 40*(1), 281-310.

Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning & Instruction, 16*(5), 492-509.

Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior, 27*(3), 1087-1102.

Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior, 27*(3), 1059-1067.

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal education in a knowledge society* (pp. 67-98). Chicago: Open Court.

Stasser, G. (1992). Pooling of unshared information during group discussions. In S. Worchel, W. Wood. & J. A. Simpson (Eds.), *Group process and productivity* (pp. 48-67). Newbury Park: Sage

Stewart, D.D. & Stasser, G. (1998). The sampling of critical, unshared information in decision-making groups: The role of an informed minority. *European Journal of Social Psychology, 28*, 95-113.

Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin, 115*(3), 323-345.