Navigating Through Controversial Online Discussions: The Influence of Visualized Ratings

Jürgen Buder, Knowledge Media Research Center, Tübingen Germany, j.buder@iwm-kmrc.de
Christina Schwind, Knowledge Media Research Center, Tübingen Germany, c.schwind@iwm-kmrc.de
Anja Rudat, Knowledge Media Research Center, Tübingen Germany, a.rudat@iwm-kmrc.de
Daniel Bodemer, University of Duisburg-Essen, Media-Based Knowledge Construction, Germany, bodemer@uni-due.de

Abstract: Awareness tools using visualized ratings of other people can help recipients to find their way in crowded information spaces. This article reports on a laboratory experiment that investigates how the design of an awareness tool impacts the navigation behavior of recipients within an online discussion forum. In the experiment, 127 participants read through a forum discussion for which posts were rated by average agreement and average quality. Depending on the experimental condition, posts were visualized along continua indicating average agreement ratings (absent vs. present) and/or average quality ratings (absent vs. present). It was hypothesized that an agreement visualization will increase the salience of both high-agreement and low-agreement posts, whereas a quality visualization will only increase the salience of high-quality posts. Measuring reading times for discussion posts, these hypotheses could be confirmed. However, neither recipients' attitude strength nor recall of discussion arguments was affected.

Introduction

One of the hallmarks of computer technologies is their ability to create social situations over a distance. Email, online discussion forums, virtual environments, multiplayer games, social networks, microblogging, and wikis: All these technologies enable communication and collaboration among persons who are spatially separated. But communicating over a distance is not without its problems, as nearly 30 years of research has shown that computer-mediated communication is impoverished in comparison to the richness of face-to-face environments (Kiesler, Siegel, & McGuire, 1984). For instance, the lack of copresence implies that it is much more difficult in a computer-mediated environment to refer to spatial relations, for example, pointing at an object (Kramer, Oh, & Fussell, 2006). Similarly, other non-verbal cues like facial expressions, intonation or pitch, are not (or not sufficiently) transmitted (Walther & Tidwell, 1995). And finally, social cues about the person one is interacting with (e.g., gender, age) might be missing. In the early 1990s researchers in the field of computer-supported cooperative work began to address this problem by developing tools that try to re-create some of the richness of face-to-face environments (e.g., Heath & Luff, 1992). These so-called group awareness tools (Gutwin & Greenberg, 2002) were able to show who is present in an environment, indicate the activities that other persons are currently involved in, or provide background information about communicators through profile pages.

Starting in the mid-2000s, researchers in the field of computer-supported collaborative learning (CSCL) have started to adapt, develop, and test group awareness tools with the goal of supporting collaborative learning processes. In conjunction with this shift of application fields came a shift in the design of group awareness tools. Rather than trying to imitate the contextual richness of copresent scenarios, CSCL awareness tools provided information that would be difficult or even impossible to yield in face-to-face environments, thus lending an added value to computer-mediated communication (Buder, 2007). For instance, CSCL awareness tools provide information about the knowledge of learning partners (Engelmann, Dehler, Bodemer, & Buder, 2009), their degree of comprehension (Dehler, Bodemer, Buder, & Hesse, 2011), their assessments on tests (Sangin, Molinari, Nüssli, & Dillenbourg, 2011), perceived personality traits (Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011), or their opinions (Buder & Bodemer, 2008).

In particular, Buder and Bodemer (2008, 2011) conducted two experiments that investigated the use of group awareness tools for controversial online discussions. Small groups of learners (3-4 members) were required to discuss two conflicting hypotheses from various science topics. By providing differential access to learning material prior to discussion, it was ensured in all conditions of both experiments that one learner (the informed minority) of each group favored a "correct" hypothesis. However, this learner was confronted with a dissenting majority who favored an "incorrect" hypothesis. The task of the groups was to discuss the conflicting evidence before making a decision about the better hypothesis. Among other things, both experiments compared groups using an awareness tool and groups without a tool. In the tool conditions, group members were required to rate the discussion posts of their collaborators on two dimensions, namely, agreement and novelty. The awareness tool computed the average ratings for each contribution and fed these data into a two-dimensional visualization in real time where each dot represented a discussion post. On the horizontal axis of the visualization, dots were displayed along an agreement dimension. The rationale behind this was to visually

separate majority contributions (high agreement) from minority contributions (low agreement) in order to create an awareness that different opinions were available. On the vertical axis of the visualization, dots were displayed along a novelty dimension. The idea behind this was that minority contributions garner higher novelty ratings than majority contributions, and therefore become highly salient. The rationale behind these visualizations can be related to the notion of representational guidance (Suthers & Hundhausen, 2003). According to this concept, different representational formats lead to differences in how individuals and groups process information: Representations constrain the way in which learners think about an object (e.g., in terms of agreement and novelty), and they make some parts of a representation more salient than others (e.g., novel minority posts). The results of both experiments indicate that without awareness tool, post-discussion preferences of groups and individuals were leaning toward the incorrect majority viewpoint. However, in groups that were supported by the awareness tool, the social influence of minority members could be strengthened, resulting in a higher likelihood that the correct hypothesis was chosen by individuals and groups. In sum, these experiments showed that carefully designed awareness tools can have an impact on the behavior of collaborating groups.

The current work builds on these prior findings, but takes the project in two new directions. The first new direction is associated with a change of the learning setting. While the prior studies by Buder and Bodemer (2008, 2011) involved small groups of learners in formal contexts, we are now investigating potentially large groups in informal contexts, that is, controversial discussions in online forums that are part of the Web portals of newspapers or magazines. The rationale for this shift is that group awareness tools can act as filters for information seeking (Dourish & Chalmers, 1994). In small groups, learners can actually attend to all information that is produced by their collaborators. However, in a large online forum there can be literally thousands of discussion posts on a single topic, and participants can only attend to a small portion of the available information.

The second departure from the prior experiments has to do with a deeper understanding of the mechanisms that help explaining the effectiveness of awareness tools. While the experiments from Buder and Bodemer (2008, 2011) showed that visualizations by a rating-based awareness tool can be effective in shaping the behavior of a group, the present work is focused on the way in which these visualizations are actually used. In other words, the current study investigates how visualizations of rating dimensions impact the navigation of recipients. The general idea is that different rating dimensions exert different types of representational guidance (Suthers & Hundhausen, 2003).

In particular, we distinguish between two types of behavior termed bipolar and unipolar navigation. If discussion posts are visualized alongside a continuum, bipolar navigation means that both ends of the continuum become salient. This should be the case for the visualization of agreement ratings, as such a visualization gives insights into the pros and cons of a discussion. In contrast, unipolar navigation means that only one end of a continuum becomes salient. In previous studies, this was accomplished with the novelty dimension. However, as the current study involves discussions in online forums where participants can enter at different times, novelty becomes a relative concept. Therefore, the current study tries to trigger unipolar navigation through the use of a quality visualization. Visualization of quality ratings is likely to exert representational guidance in a way that high-quality posts become more salient than low-quality posts.

Hypotheses

The present experiment investigates these issues by confronting individuals with the content of a large online discussion, with only limited time to read all discussion posts. If the posts are visualized along an agreement rating continuum, it might become likely that recipients will read posts with both high and low agreement ratings (bipolar navigation) as this provides insights into the different viewpoints of the controversial discussion (pro vs. con). A good balance between attention towards the pros and cons of a controversial discussion should not only be found with regard to the frequencies of opened discussion posts, but also with regard to subsequent reading times. We measured navigation by reading times because they provide better insights into the processing of information. As a consequence of bipolar navigation, recipients might achieve a more balanced view on the controversial issue. This, in turn should be accompanied by an attenuation of a recipient's post-discussion attitude strength compared to the pre-discussion attitude strength. This leads to the following hypotheses:

Hypothesis 1: Participants who are provided with an agreement visualization should show smaller reading time differences between pro and con discussion posts than participants who are not provided with an agreement visualization.

Hypothesis 2: Participants who are provided with an agreement visualization should show a weakening in attitude strength, whereas participants who are not provided with an agreement visualization should show an enhancement in attitude strength.

If the discussion posts are visualized along a quality rating continuum, it might become likely that recipients will read high-quality posts at the expense of low-quality posts (unipolar navigation). As high-quality

contributions are most likely to contain the main arguments of a discussion, the availability of a quality visualization should also lead to better memory for these crucial arguments. This leads to the following hypotheses:

Hypothesis 3: Participants who are provided with a quality visualization should have longer reading times for the top quality contributions than participants who are not provided with a quality visualization.

Hypothesis 4: Participants who are provided with a quality visualization should better recall the crucial discussion arguments than participants who are not provided with a quality visualization.

Method

Participants

Data were collected from 127 participant volunteers (38 male). All participants were students, recruited via a university mailing list. For their participation in the experiment, which took 60 minutes, participants were paid 8 €. Alternatively, the students could get a certificate of their participation if needed for course requirements. Participants' age ranged from 18 to 58 years (M = 24.94 years, SD = 5.22, one missing value).

Design

This lab study used a 2 x 2 design to explore the effect of the two visualizations of agreement ratings and quality ratings. Thus, we set up four experimental conditions in which the visualizations of agreement ratings and/or quality ratings were either available or not (see Table 1): a condition without any visualization (*no Visualization*), a condition with only a visualization of the agreement ratings (*Agreement only*), a condition with only a visualization of the quality ratings (*Quality only*), and a condition with a combination of both visualizations, agreement ratings and quality ratings (*Combination*).

Table 1: 2 x 2 factorial design of the experiment.

Availability of rating visualization		Quality	
		NO	YES
Agreement	NO	*no Visualization*	*Quality only*
	YES	*Agreement only*	*Combination*

Material

Discussion forum: We used a discussion from an online forum that is part of the Web portal of a German news magazine. The discussion is about pro and cons of the three-tier school system, which is a controversially discussed topic in German public discourse. We took the first 137 original posts and coded them for their position (pro or contra the discussed three-tier school system, Interrater-Reliability Cohen's κ = .64) and the arguments contained therein.

Arguments: Based on the material we identified eleven main arguments, speaking for or against the three-tier school system (e.g., "Leveling down is bad and has negative consequences. It results in overload or underload." as an argument pro three-tier school system; "It is bad to select pupils too early because their potential might be unexploited." as an argument contra three-tier school system).

Ratings: In a pre-study (N = 67), all discussion posts have been rated on their quality on a five point Likert-scale (1 *very low* to 5 *very high*). Out of the pro-posts and out of the contra-posts we identified the ten qualitative best posts (so called *TOP 10* posts), five of pro and contra each. All main arguments can be found in these TOP 10 posts. For reasons of experimental systematicity, the ranges of ratings on agreement and quality have been adapted in a way that, for example, all posts could be identified separately in the visualization.

List display: The discussion was presented in a simulated online forum environment and the discussion posts were presented in all conditions in a list of their headings in their original chronological order. In the no Visualization condition only the list display was presented.

Visualizations: Participants in the three other conditions (Agreement only, Quality only, and Combination) were additionally provided with embedded visualized rating dimensions. The visualization represented all discussion contributions as dots along continua, with the position of each dot indicating the average ratings that a contribution received.

Measures

As dependent variables we measured reading times, attitude strength, and recall. Reading times include three different types: the reading time of the pro-posts and the contra-posts, and the reading time of the TOP 10 posts.

Reading time PRO-CON. Reading time PRO-CON was measured for testing the impact of different types of visualizations on bipolar navigation. We measured and compared reading time PRO-CON by calculating the difference of the reading time of pro-posts and the reading time of the contra-posts.

Attitude strength. The attitude indication was represented by a word pair ranging from *weak* (= 1) to *strong* (= 5) attitudes about the subject matter. The attitude strength was calculated as the difference between the pre-discussion attitude strength and the post-discussion attitude strength. Zero indicates that there was no difference in attitude strengths, whereas negative values indicate an enhancement of attitude strengths, and positive values indicate a weakening of attitude strengths.

Reading time TOP 10. We measured reading times of the TOP 10 posts. In this way, it was possible to test the impacts of different types of visualizations on unipolar navigation.

Recall. We measured free recall of arguments. For that, participants wrote down as many arguments out of the whole discussion they could remember. For measurement, we coded the number of recalled arguments with a self-developed coding scheme (Interrater-Reliability Cohen's $\kappa = .72$).

Procedure

We recruited participants from a database of all university students from a German university via mailing list and asked them to take part in an "online discussion forum - study" where they would have to read an online discussion forum. All instructions and materials were presented on a computer screen. Each participant was randomly assigned to one of the four experimental conditions. After having read the instructions, participants indicated their pre-attitudes about the three-tier school system. Then, all participants had 20 minutes time to navigate freely in the online discussion forum. To avoid them rushing through the posts without reading, they were told it is hardly possible to read all 137 posts within 20 minutes. Participants in the no Visualization condition saw an online forum without any ratings, but only with a list display. Participants in the Agreement only condition were additionally provided with visualized average ratings of agreement of all discussion posts. Each dot in the visualization represented one corresponding post (see Figure 1). In contrast, participants in the Quality only condition saw a visualization with quality ratings of all discussion posts. Again, each dot in the visualization represented the average rating the corresponding discussion post received. Participants in the Combination condition were provided with visualized ratings of both, agreement and quality (see Figure 2) additionally to the list display. Again, in the visualization each dot marked one corresponding discussion post. The visualization in the Combination condition was two-dimensional. That means, on the horizontal axis the agreement ratings were shown, and on the vertical axis the quality ratings were shown.

Participants could access forum posts either by scrolling through the list display, or by clicking on a dot in the visualization. Only by clicking, the corresponding post opened on the bottom of the screen. In all visualization conditions, the headings of the posts could be made visible by moving the mouse over a dot of the visualization. Posts that have been read were marked in orange in both the list display and the visualization.

After 20 minutes with the online forum, post-attitudes about the three-tier school system were measured. Then, participants performed the recall task. For that, they were told to keep in mind that arguments should not be confused with posts as one single post could include none, one or more arguments. At the end, participants were thanked and debriefed.

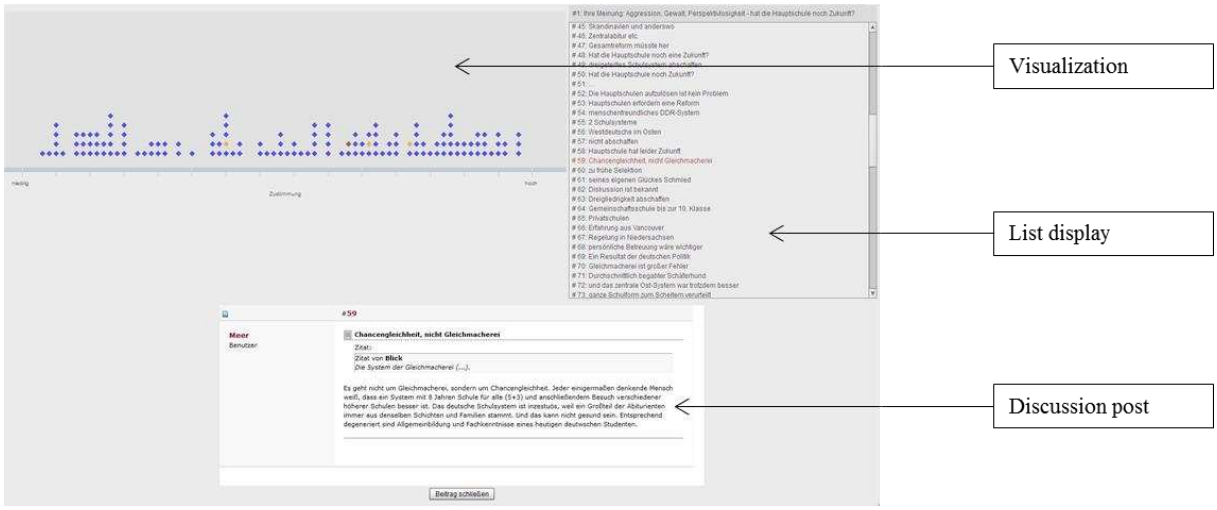

Figure 1. Screenshot of the presented online forum in the Agreement only condition.

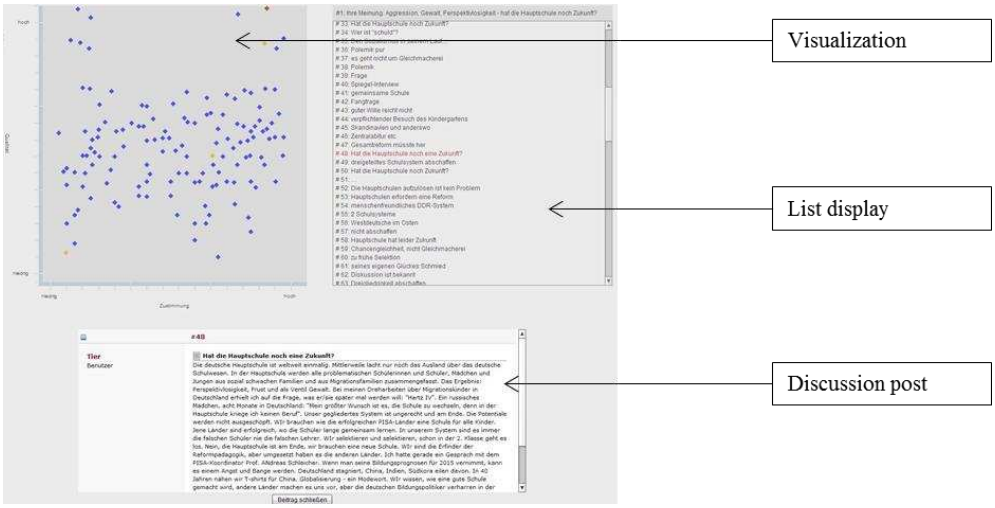

Figure 2. Screenshot of the presented online forum in the Combination condition.

Results

The preliminary analysis and the main results were based on logfile data (opening frequencies, reading time PRO-CON and reading time TOP 10). Navigation and recall measures were analyzed using analyses of variance (ANOVAs). Before addressing the main results, the preliminary analysis of the opening frequencies investigated in how far participants in the three conditions with visualizations (Agreement only vs. Quality only vs. Combination) used the visualization vs. the list display for navigation.

Preliminary analysis. We explored if participants show differences in using the visualizations or using the list display to navigate within the discussion. The measurement expressed the difference between the amount of post openings via the visualization and the amount of post openings via the list display. This means that positive values indicate that participants used mainly the visualization, whereas negative values indicate that participants used mainly the list display for navigation.

To explore the difference in opening frequencies via the visualization and via the list display, a one-factorial ANOVA with condition (Agreement only vs. Quality only vs. Combination) as independent variable and the difference of opening frequencies (opening frequencies for the visualization minus opening frequencies for the list display) as dependent variable was conducted. Results show a significant main effect for condition ($F(2, 92) = 5.75$, $p = .004$, partial $\eta^2 = .11$). Subjects in the Agreement only condition ($M = 5.81$, $SD = 16.94$) as well as subjects in the Quality only condition ($M = 1.87$, $SD = 16.02$) mainly used the visualization, whereas subjects in the Combination condition ($M = -8.19$, $SD = 18.02$) used the list display for navigation (see Figure 3).

Pairwise comparisons using Tukey-HSD revealed that there was no difference between the Agreement only condition and the Quality only condition ($p = .630$) regarding the opening frequencies via visualization and via list display. However, subjects in the Combination condition used the visualization less frequently compared to subjects in the Agreement only condition ($p = .004$) and marginally less frequently compared to subjects in the Quality only condition ($p = .055$).

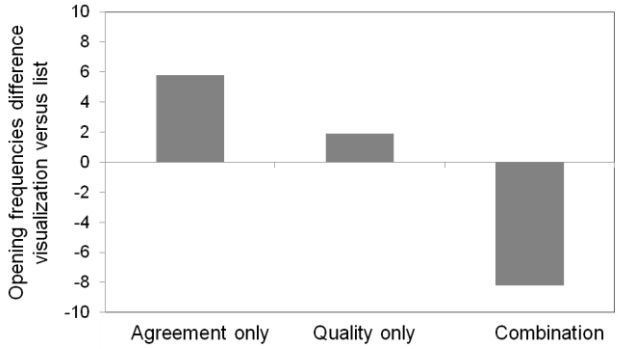

Figure 3. Opening frequencies difference (visualization versus list display) for the three conditions with visualization.

Reading time PRO-CON. In Hypothesis 1, it was predicted that the presence of an agreement visualization will support bipolar navigation: Participants in the conditions with agreement visualization will show more balanced reading times, meaning that they equally distribute their overall reading time to pro posts and to con posts. To test this prediction, we computed a two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and reading time PRO-CON as dependent variable. In line with the hypothesis, the analysis for the reading time PRO-CON revealed a significant main effect, $F(1, 123) = 8.37$, $p = .005$, partial $\eta^2 = .06$. As shown in Figure 4, participants in the conditions with agreement visualization ($M = 27.27$, $SD = 198.85$) showed more balanced reading time PRO-CON than participants in the conditions without agreement visualization ($M = 130.33$, $SD = 199.08$).

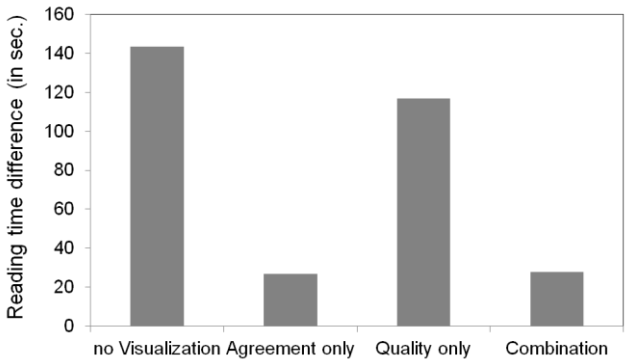

Figure 4. Reading time PRO-CON (in sec.) in the four conditions.

Attitude strength. We hypothesized that participants in the conditions with the agreement visualization will show a weakening in attitude strength, whereas participants in the conditions without the agreement visualization will show an enhancement in attitude strength (Hypothesis 2). We analyzed this hypothesis using a two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and the difference of pre-discussion attitude strength and post-discussion attitude strength as dependent variable. Contrary to our expectations, no difference between the four conditions was revealed: Participants with the agreement visualization ($M = -0.03$, $SD = 0.87$) showed the same attitude strength difference between pre-discussion and post-discussion as did participants in the conditions where the agreement visualization was not present ($M = -0.19$, $SD = 0.78$). There were no differences for agreement visualization ($F(1, 123) = 1.18$, *ns*), for quality visualization ($F(1, 123) < 1$, *ns*) and no interaction effect ($F(1, 123) < 1$, *ns*). This means that no change in attitude strength was found in any of the four conditions.

Reading time TOP 10. It was hypothesized that the quality visualization will support unipolar navigation. This means that for conditions with quality visualization, it was predicted that participants will show longer reading times for the TOP 10 posts than participants in the conditions without quality visualization (Hypothesis 3). A two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and reading time TOP 10 as dependent variable was conducted. The analysis yielded the expected highly significant main effect, $F(1, 123) = 56.91$, $p < .001$, partial $\eta^2 = .32$. Participants in the conditions with quality visualization ($M = 379.16$, $SD = 214.05$) read TOP 10 posts longer than participants in the conditions without quality visualization ($M = 157.58$, $SD = 103.08$); see Figure 5.

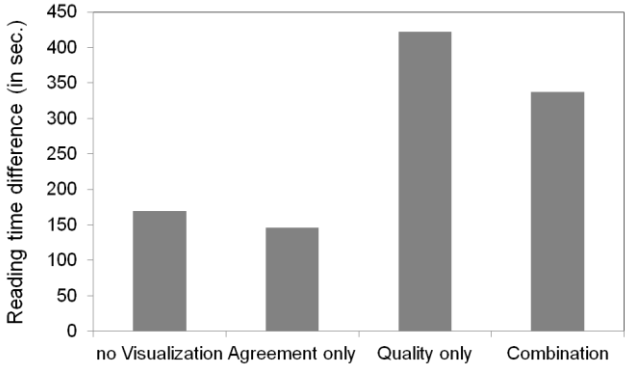

Figure 5. Reading time TOP 10 (in sec.) in the four conditions.

Recall. Since TOP 10 posts cover all crucial arguments in the discussion, we expected a similar result pattern for reading time TOP 10 and for recall. Therefore, it was hypothesized that participants in the conditions with quality visualization will show better recall for the crucial discussion arguments than participants in the conditions without quality visualization. To test Hypothesis 4, we computed a two-factorial ANOVA with Agreement (agreement visualization: no vs. yes) and Quality (quality visualization: no vs. yes) as independent variables and recall as dependent variable. Contrary to our expectations, no differences between the four conditions were found (no Visualization: $M = 3.56$, $SD = 1.32$; Agreement only: $M = 3.50$, $SD = 1.63$; Quality only: $M = 3.74$, $SD = 1.83$; Combination: $M = 3.46$, $SD = 1.67$). Participants with the quality visualization ($M = 3.46$, $SD = 1.76$) recalled the same amount of arguments as did participants in the conditions where the quality visualization was not present ($M = 3.53$, $SD = 1.47$). There were neither differences for agreement ($F(1, 123) = 1.15$, *ns*), nor for quality ($F(1, 123) < 1$, *ns*) and no interaction effect ($F(1, 123) < 1$, *ns*). This means that overall participants in all four conditions recalled the same amount of arguments.

Discussion

We have tested an awareness tool that employs visualized user ratings of discussion posts. In order to test whether different forms of visualizations give rise to different navigation styles, we conducted an experiment that varied the presence or absence of an agreement rating visualization and a quality rating visualization. A number of results provide insights into the design of awareness tools and its implications for learner navigation. First, results indicate that making agreement ratings available leads to bipolar navigation, that is, readers of the discussion forum tended to read both pro and con contributions of the controversial discussion to the same degree. Ensuring balanced attention to both viewpoints pertaining to a controversial issue is an important antecedent of unbiased opinion formation and critical thinking (Schwind, Buder, Cress, & Hesse, 2012). Hence, the current study gives some insight into how the design of awareness tools can contribute to an open-minded stance of learners towards a controversial topic. Second, the study could show that the availability of quality ratings leads to unipolar navigation, that is, readers of the discussion forum spent more time reading the TOP 10 posts once this information was made salient. Third, our results show that combining both rating visualizations in a two-dimensional display comes with a cost: This could be the main reason why learners in the Combination condition seemed to prefer a standard list display for navigation. Taken together, the navigation results show that the design of awareness tools can exert different types of representational guidance, but that the complexity of a graphical representation could be a roadblock that might prevent a group awareness tool from unfolding its full potential. For instance, in the present study the group awareness tool did *not* affect attitude strength and recall. The reason for the lack of effects on cognitive variables might be explained by the status of our participants as relatively passive recipients who could only read discussion posts. Our prior experiments involving active participants have shown that an awareness tool also impacts individual and perceived group preferences. This would be in line with the general notion that active participation is a key to effective collaboration (Cohen, 1994). However, the lack of cognitive effects can also be explained by the specific paradigm used in the study: The measurements for the cognitive variables (i.e. attitude strength and recall) were delayed measurements. Therefore, it might be interesting to investigate in follow-up studies if it is possible to affect cognitive variables when directly measured after the confrontation with the group awareness tool. Another question for further research is whether learners with higher involvement and/or interest in the topic show more effects on cognitive variables than the student population in the present study. Thus, topical involvement and interest might be moderators for the found results.

The current study employed an experimental approach that allowed us to carefully control for opening frequencies and reading times with regard to each discussion post. While the experimental control had the advantage that navigational behavior could be assessed in greater detail than before, it also gave rise to a number of limitations having to do with the selection of participants and the temporal nature of the study. First, participants were not self-selected. This is in contrast to real-world online forums in which it is likely that readers will have a personal interest in the topic under discussion. It would be interesting to see whether the results that we have found (and even more, the results that we haven't found) could be shown with a sample of participants who have a vested interest in the controversy. Second, participants were confronted with a static picture of a controversial group discussion. While it could be said that even in real-world online forum each reader only sees a static picture of the discussion when logging in, it cannot be denied that there is an ongoing dynamic in a discussion that unfolds over repeat visits in a forum. It would be fascinating to see how a rating-based group awareness tool would fare as a support mechanism of a large, controversial online forum, particularly at the early and formative stages of discussion. Will the distribution of agreement ratings lead to changes in the type of posts that are made? Will high-quality posts spur or inhibit further discussions? These questions can only be tapped into by employing a rating-based group awareness tool in an actual online forum. Consequently, transferring group awareness tools into the "real world" is an obvious step once the mechanisms of these tools are better understood.

Collecting user ratings has become a widespread phenomenon in Web 2.0 environments. However, most systems make use of relative primitive ratings (5-star ratings or "Like it" ratings). We set out to explore whether different forms of representational guidance can be achieved by using more specific rating dimensions. Our results lend some support to the idea that navigation of learners in large online environments can be shaped in educationally meaningful ways.

References

Buder, J. (2007). Net-based knowledge-communication in groups. *Zeitschrift für Psychologie, 215(4)*, 209-217.

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

Buder, J., & Bodemer, D. (2011). Group awareness tools for controversial CSCL discussions: Dissociating rating effects and visualized feedback effects. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings* (Vol. I, pp. 358-365). Hong Kong, China: International Society of the Learning Sciences.

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

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

Dourish, P., & Chalmers, M. (1994). Running out of space: Models of information navigation. In G. Cockton, S. W. Draper, G. R. S. Weir (Eds.), *Proceedings of the HCI Conference*. Cambridge, UK: Cambridge University Press.

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

Gutwin, C., & Greenberg, S. (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer-Supported Cooperative Work, 11*, 411-446.

Heath, C., & Luff, P. (1992). Collaboration and control: Crisis management and multimedia technology in London Underground Line control rooms. *Journal of Computer-Supported Cooperative Work, 1*, 24-48.

Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist, 39*, 1123-1134.

Kramer, A. D., Oh, L. M., & Fussell, S. R. (2006). *Using linguistic features to measure presence in computer-mediated communication* (Research Paper 53) Retrieved from Human-Computer Interaction Institute website: http://repository.cmu.edu/hcii/53

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*, 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*, 1059-1067.

Schwind, C., Buder, J., Cress, U., & Hesse, F. W. (2012). Preference-inconsistent recommendations: An effective approach for reducing confirmation bias and stimulating divergent thinking? *Computers & Education, 58*, 787-796.

Suthers, D., & Hundhausen, C. (2003). An empirical study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences, 12*, 183–219.

Walther, J. B., & Tidwell, L. C. (1995). Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *Journal of Organizational Computing*, *5*, 355-378.

Acknowledgments

This research was financed and supported by the ScienceCampus Tübingen "Informational Environments". We would like to thank Susanne Stoll and Tim Höfling for their support.