British Journal of Educational Psychology (2011)

The British Psychological Society

www.wileyonlinelibrary.com

Small-group, computer-mediated argumentation in middle-school classrooms: The effects of gender and different types of online teacher guidance

Christa S. C. Asterhan*, Baruch B. Schwarz and Julia Gil
Hebrew University of Jerusalem, Israel

Background. Research has shown the importance of careful teacher support during collaborative group work to promote productive discourse between students (Webb, 2009). However, this research has traditionally focused on face-to-face communication. The role of online teacher guidance of small-group computer-mediated discussions has received little attention, especially in secondary school classroom settings. Researchers of computer-supported collaborative learning (CSCL), on the other hand, have traditionally focused on software-embedded features, such as scripts, to support a-synchronous peer dialogue, and less so on human guidance of synchronous group discussions.

Aims. The main aim of the present *in vivo*, experimental study is to examine whether online teacher guidance can improve the quality of small-group synchronous discussions, and whether different types of guidance (epistemic or interaction guidance) affect these discussions differently, when compared to an unguided condition. The second goal of this study is to explore potential differences between all-female and all-male discussion groups.

Sample. Eighty-two 9th graders (three classrooms) and six teachers from a rural high school in Israel.

Results. Whereas epistemic guidance only improved aspects of the argumentative quality of the discussion, interaction guidance only improved aspects of collaboration. Discussions of all-girls groups scored higher on aspects of collaboration and argumentative quality, compared to all-boys groups.

Conclusions. The findings show that teacher guidance of synchronous, online discussions in classrooms is realizable and reasonably reaches its intended goals. Training should be focused on acquiring various guidance strategies to augment their beneficial effects. Furthermore, future research should pay more attention to potential gender differences in peer-to-peer argumentation.

**Correspondence should be addressed to Christa S. C. Asterhan, Hebrew University of Jerusalem, Mt. Scopus 91905, Jerusalem, Israel (e-mail: asterhan@huji.ac.il).*

DOI:10.1111/j.2044-8279.2011.02030.x

The potentially positive benefits of small-group peer interaction on individual learning and development have been repeatedly shown over the past four decades (for reviews, see a/o: Johnson & Johnson, 1989; O'Donnell, 2006; Slavin, 1996; Webb, 2009). Most researchers also agree, however, that simply placing students in small groups does not guarantee that they will engage in productive interaction. A substantive and ever-increasing amount of empirical evidence shows that the extent to which students benefit from peer interaction depends on the nature of participation and the type of dialogue they engage in (e.g., Asterhan & Schwarz, 2007; 2009b; Webb & Palincsar, 1996). However, learners often fail to spontaneously engage in these productive forms of interaction and dialogue (Cohen, 1984; Fischer, Kollar, Haake, & Mandl, 2007; Tolmie *et al.*, 2005). The question is how should students be supported to engage in productive peer-to-peer dialogue?

Many have argued that computer-mediated communication formats offer several advantages that may facilitate productive peer dialogue in educational settings (Kim, Anderson, Nguyen-Yahiel, & Archodidou, 2007). For example, students have been found to be more reflective (Guiller, Durndell, & Ross, 2008), interaction is often more egalitarian and democratic (Asterhan & Eisenmann, 2009; Herring, 2001), and students are more explicit in their communication (Kim *et al.*, 2007; Newman, Webb, & Cochrane, 1995).

The question is, however, whether the embedded support in the medium itself is sufficient, or can online teacher support *during* computer-mediated discussion improve the quality and effectiveness of these discussion activities? The main aim of the *in vivo* experiment we report on here is to explore this question in the context of synchronous online group argumentation in middle school classrooms. The school in which the study was conducted was committed to an educational initiative according to which certain disciplines were taught in same-gender classes. The reality of the school environment then enabled a second research question to explore: whether gender plays a role in online argumentation. The relevant literature for each of these two factors, human guidance and gender, will be reviewed separately, starting with the former.

Teacher guidance of online, peer-led dialogue

Teacher guidance of peer-led dialogue is a complex undertaking, which requires maintaining a delicate balance between supporting student autonomy while at the same time scaffolding reasoning and interaction. The following summary of the research shows that scholars from different research communities have chosen to focus on different aspects of this support. Moreover, it reveals that the role of human facilitation of synchronous computer-mediated discussions has received scant attention.

Human guidance in face-to-face settings

In a recent review, Webb (2009) summarized the empirical research on the teacher's role in promoting productive peer-to-peer dialogue in classrooms. In addition to preparing students, modelling dialogue practices, and designing tasks and settings, she identified teacher support *during* group work as an effective means to improve the productivity of children's peer-to-peer dialogue. However, the extent of effectiveness seems to be dependent on both the type, as well as the timing of teacher guidance. For example, several studies have shown that when teachers' interventions focus on doing the task for the group or providing the correct answer (direct instruction), students engage in less elaboration, less explanatory activities, and ask each other less questions (e.g., Dekker &

Elshout-Mohr, 2004; Gillies, 2006). The opposite is true for teacher interventions that probe student reasoning and thinking: students elaborate more, explain more, and ask each other more questions. Similar findings have been reported from studies on one-on-one tutoring activities (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001). As Webb points out, however, it is not only the content of the teacher intervention, but also its timing and calibration to group progress that distinguishes between effective and non-effective support (e.g., Chiu, 2004; Hogan, Nastasi, & Pressley, 2000).

In spite of its breadth and depth, this important review did not explicitly consider the affordances that are embedded in advanced communication technologies and the role of the instructor in such environments.

Support during computer-supported collaboration

Among researchers of computer-supported collaborative learning (CSCL), on the other hand, the role of guidance is often delegated to the design of computerized tools, for example, through representational affordances (Suthers, 2003), sentence openers (Cho & Jonassen, 2002), argumentative ontology (Schwarz & Glassner, 2007), and automated prompts from intelligent agents during students' learning activities (Koedinger & Corbett, 2006). The role of the human teacher in and during CSCL activities, on the other hand, has received little empirical attention, especially in primary and secondary school settings (e.g., Lentell & O'Rourke, 2004; McPherson & Nunes, 2004).

A particularly dominant approach in studies of CSCL support *during* group processes has been the collaboration script approach (Fischer *et al*., 2007). The main idea of computerized collaboration scripts is to promote productive, structured interaction by designing the environment such that they *force* collaborating students to engage in specific activities that might otherwise not occur (Kollar, Fisher, & Hesse, 2006; Rummel, Spada, & Hauser, 2009). This can be achieved by software design, for example, by blocking the ability to post a message until all group members have reacted, or by prompting certain sequences of dialogue moves (e.g., claim–counterclaim–rebuttal).

Several studies have shown that software-embedded CSCL scripts may improve online group functioning, and in some cases, also subsequent individual achievement (e.g., Rummel & Spada, 2005; Scardamalia & Bereiter, 1996; Stegmann, Weinberger, & Fischer, 2007; Weinberger, Ertl, Fischer, & Mandl, 2005). However, collaboration scripts have also been criticized: One critical issue is the coercive way in which scripts often dictate interaction. This coercion may dampen student motivation (Rummel *et al*., 2009), it may interfere with their personal, possibly well-functioning collaboration scripts (Hesse, 2007) and may prevent their independent, playful, and exploratory thinking (Dillenbourg, 2002).

However, arguably the most controversial feature of traditional software-embedded collaboration scripts is their rigidness. As aforementioned, the most crucial aspect of effective support *during* the collaborative task is calibration of support to the history and current state of group and individual functioning (Rogoff, 2003; Webb, 2009). Puntambekar and Hubscher (2005) noted that whereas many CSCL software tools provide novel techniques to support student learning, they often lack crucial features of 'scaffolding', such as ongoing diagnosis, calibrated support, and progressive fading. It is questionable whether traditional software-embedded scripting approaches, according to which the stages, prompts, activities, and roles of collaboration are preset and unalterable, can meet these goals. Indeed, several researchers have recently begun to incorporate some of these notions in the scripting approach (see e.g., Dillenbourg &

Tchounikine, 2007; Wecker & Fischer, 2007). The question remains, however, whether future progress in Computer Science and Artificial Intelligence may *ever* be expected to be of comparable levels of sensitivity, adaptivity, and flexibility that expert human guidance can offer?

We therefore propose that the study of support *during* CSCL should not be confined to scripting or other software-embedded tools only, but should also explore the possibility and potential promise of online *human* guidance. Moreover, the consideration of human support as a viable alternative seems to be particularly relevant for the immediate, dynamic, and simultaneous nature of *synchronous* modes of computer-mediated communication (CMC) and for learning tasks in ill-defined domains, such as small group discussions on social dilemmas and controversial issues. In a recent case study, Schwarz and Asterhan (in press) showed how human expertise and judgment is often called for in such tasks: (1) to adequately evaluate the social and motivational dimensions of these complex interpersonal situations; (2) to flexibly and instantly adapt support for individual and group processes in ways that were foreseen or unforeseen; and (3) to intervene in a matter that is sensitive to these subtleties.

Online human guidance of synchronous peer dialogue in classrooms

The role of the human instructor in distant learning, and in particular e-courses, has been extensively discussed in the e-moderation literature (e.g., Berge, 1995; Goodyear, Salmon, Spector, Steeples & Tickner, 2001; Mason & Kaye, 1989; Mazzolini & Maddison, 2003; Paloff & Pratt, 2001; Salmon, 2004). However, engaging in *synchronous* group communication is in multiple ways different from the asynchronous, distributed formats of communication that are common in e-courses (e.g., Cress, Kimmerle & Hesse, 2009; Veerman, Andriessen & Kanselaar, 2000) and these findings and insights may, therefore, prove to be of limited relevance. Indeed, a recent study showed that the findings on human guidance in other communication formats (F2F and asynchronous, distributed CMC) cannot be simply transferred to synchronous, co-located settings, and that a separate investigation is warranted (Asterhan & Schwarz, 2010).

Several studies have described the viability of guiding synchronous peer discussions in educational settings (Asterhan, 2011; Asterhan & Schwarz, 2010; De Groot, in press; Hlapanis, Kordaki, & Dimitrikapoulou, 2006; Schwarz & Asterhan, in press; Walker, 2004). However, these descriptive studies did not compare conditions of guided and unguided discussions, and it is therefore not possible to determine whether human guidance actually improves the quality of peer discussions, or not. This is important, since it is quite possible that the presence of a human instructor in the discussion environment may in fact be redundant or even interfere with group functioning.

Also, the research on teacher support in F2F settings clearly shows that one determinant of effectiveness is the *type* of guidance that instructors offer. In synchronous CMC, this has been investigated by Veerman, Andriessen and Kanselaar (2000), who compared the effect of two different types of guidance (a focus on improving argumentation structure vs. improving argumentation strength) on undergraduates' synchronous argumentation in a chat-based environment. When these two conditions were compared to an unguided control condition no significant differences in discussion quality were detected. It is possible however, that for this particular population human support was redundant, since undergraduate students can reasonably be expected to be capable of conducting a good argumentative discussion on a topic of their interest. Young teenagers, on the other hand, have been reported to have developed only rudimentary argumentative skills (Kuhn, 1991) and are therefore expected to benefit more from teacher support.

Interactional and epistemic guidance for dialogical argumentation

So as to define the different kinds of guidance to be studied, we first observed how in-service teachers intuitively guide synchronous small-group discussions in a natural, educational setting (Asterhan, 2011). With the help of a multi-dimensional methodology that triangulated different qualitative and quantitative discussion features, several distinctively different human guidance styles were identified. Of these, two were of particular interest, since they seemed to balance active involvement with an unobtrusive but caring form of guidance: The first was characterized by interventions that aimed at organizing and structuring the interactional and the task-related aspects of the activity (e.g., making sure that people respond to each other and that they adhere to the instructions), whereas the second focused more on the argumentative elements of the discussion (e.g., scaffolding individual and group reasoning with prompts).

These findings were used to formally define two different types of guidance for peer argumentation: Argumentation can be regarded from an epistemological perspective, by specifying the different components (e.g., claims, reasons, rebuttals) of either an individual argument or the collective argument constructed by the group. Support that focuses on this aspect of dialogical argumentation should then assist the group in presenting clear, sound arguments and counterarguments and in considering different perspectives.

Conversely, an argumentative activity can also be regarded from an interactional, social perspective, by specifying the constellation of different opinions represented by different discussants and mapping the interaction between these ideas. Support that focuses on this aspect of an argumentative interaction should make sure that students reveal their personal standpoints, interact with other students that think differently from them, and relate to these ideas (i.e., by expressing agreement and/or disagreement). We refer to the former as *epistemic guidance for argumentation* and to the latter as *interaction guidance for argumentation*[1].

It should be noted that in spite of their different emphases, these two types of guidance have a common aim, namely improving collaborative group argumentation such that students consider different perspectives in a collaborative reasoned manner. The latter seeks to achieve this goal by exploiting the social situation: By encouraging students to listen and respond to students with different viewpoints, it is assumed that they will engage in reasoned argument to settle their differences.

The effects of these different types of guidance were tested in a randomized, *in vivo* experiment conducted in one middle school. The quality of small-group argumentation was assessed according to two different discussion dimensions: the extent to which individuals as well as the group engage in reasoned argumentation (argumentative

---

[1] Our distinctions are similar but somewhat different from those by Weinberger, Stegmann, Fischer and Mandl (2007): They used the term "epistemic scripts" to refer to scripting that guides the individual learner to focus on task characteristics and the cognitive actions needed to complete the task successfully, whereas "argumentative scripts" aim to help individual collaborating learners construct formally adequate arguments. We, on the other hand, use the term epistemic guidance to refer to human support that focuses on the epistemic nature of argumentation, both on the group as well as the individual level. Weinberger et al.'s conceptualization of social collaboration scripts, on the other hand, is very similar to what we named interaction support for argumentation. In both, discussants are encouraged to critique and respond to each others' contributions.

dimension), the extent to which students actively participate, interact, and respond to each other (collaboration dimension). It is expected that, compared to an unguided control condition each type of teacher guidance will improve different aspects of the collaborative dialogue: teacher support that focuses on the epistemic aspects of peer argumentation is expected to improve the argumentative quality of the computer-mediated discussion, but will not improve the extent of participation and interactivity. In contrast, teacher support that focuses on the interactional aspects of peer argumentation is expected to increase rates of participation and interactivity. However, since it does not *directly* address the quality of student argumentation, it is not expected to be successful in improving group discussions on this dimension.

Gender and online group argumentation

As aforementioned, the middle school in which the present study was situated participated in an educational pilot project, unrelated to ours, according to which certain disciplines were taught in same-gender classes. The data were therefore collected from same-gender discussion groups. Even though the study of gender differences was not the main goal of the study, it can nevertheless not be ignored. In spite of the large number of scholarly works on dialogical argumentation in education, however, to our best knowledge the role of gender has thus far not been considered (but see Asterhan, Butler, & Schwarz, 2011, for a recent study). This is surprising, since findings from a number of adjuvant research fields provide indications that female and male discussants may behave quite differently during argumentative activities:

First of all, research on cognitive styles has shown that, on average, girls tend towards a different cognitive style than boys, also referred to as 'connected way of knowledge' (Belenky, Clinchy, Goldberger, & Tarule, 1986). Accordingly, girls emphasize understanding, empathy, acceptance, cooperation (Clinchy, 1989; Galotti, Clinchy, Ainsworth, Lavin, & Mansfield, 1999; Galotti, Drebud, & Reimer, 2001), and interaction with others through verbal conversation (Zohar, 2006). Moreover, girls are more socialized in collaborative problem-solving tasks and discussion practices and they tend to take into consideration their own personal knowledge more than boys do (e.g., Baxter-Magolda, 1992; Miller, 2005).

Secondly, communication research shows that the degree to which people tend towards confrontational or consensual discourse differs consistently among men and women (Cameron, 1998; Stokoe, 2000; Weatherall, 2000). In their early analysis of same-sex peer groups, Maltz and Borker (1982) distinguished between the competitive, adversarial speech of boys aimed at asserting and maintaining dominance and the collaborative, affiliative speech of girls, which aims to 'create and maintain relationships of closeness and equality' (p. 424).

Finally, gender differences have also been reported in the literature on online communication in higher education learning settings (see Caspi, Chajut, & Sapporta, 2008, for an overview): Even though there is some variability with regards to the findings, the majority of studies report on higher participation by female students, both in terms of posting as well as of reading messages (e.g., Caspi *et al.*, 2008; Gunn, McSporran, Macleod, & French, 2003).

Based on these combined findings, it was expected that all-girls discussion groups would show higher measures of participation and interaction than all-boys groups. Our expectations with regards to the argumentative dimension of discussions were less well formulated: On the one hand, and based on discourse differences found in F2F settings, boys could be expected to show a higher tendency towards critical dialogue. On the other hand, computer-mediated communication tools have been reported to decrease the effects of social cues and expectations (Kiesler, Siegel, & McGuire, 1984; Herring, 2004). As a result, such differences may be less apparent or even disappear in e-discussions.

Method

Participants

Eighty-two 9th graders (44 male and 38 female) and three female homeroom teachers from three intact 9th grade classrooms of a rural, secular school in central Israel participated in this study. The homeroom teachers participated in an in-service teacher training provided by the Kishurim program (Schwarz & DeGroot, 2007), which is an educational initiative developed in Israel to foster argumentation and dialogic thinking in schools that has been active since 1998. They participated in a total of 32 hr of training (eight meetings of 4 hr each) and received in-class support to implement dialogic activities and argumentation in their teaching practices. The teachers were also trained to implement such activities with the help of computer-mediated discussion tools, and in particular with Digalo (see below). In addition, three adult female teachers who were not part of the school faculty served as 'guest teachers' in this study to act as moderators in the two experimental conditions. Each of these guest teachers was part of the Kishurim pedagogical team, and one of them (Julia Gil) was also part of the research team. They then had prior experience with argumentative discussions in classrooms as well as with the software environment. However, they were not particularly experienced with e-moderation in this or any other online discussion environment (less than three times prior to the experiment) and their moderation behaviour was tightly controlled and scripted to avoid interpersonal and/or inter-condition differences (see Procedure). Other than acting as moderators in the online discussion sessions they did not take part in any of the other everyday classroom activities.

Design

Students were assigned to one of four discussion groups within each same-sex classroom, based on teacher considerations that safeguarded heterogeneous groups with regard to academic ability and social position. Each discussion group was then randomly assigned to condition (EGA [Epistemic Guidance for Argumentation], IGA [Interactional Guidance for Argumentation], control or homeroom teacher – see the procedure further on) within classroom.

Tools and materials

The discussion environment

The discussions were all conducted in the diagram-based discussion environment *Digalo* (freely available at http://www.argunaut.org). Participants in a Digalo discussion post textual contributions inside tagged geometrical shapes, place these in a two-dimensional space, and link the textual contributions with different types of arrows. Each discussant works on a personal computer and sees the display of the collaboratively constructed discussion map of his/her own group. The different tagged geometrical shapes constitute the ontology that specifies and constraints the kinds of dialogue moves discussants can choose from during their discussions. The palette of tagged shape options is specified beforehand by the teacher. In the present study, for example, the different options

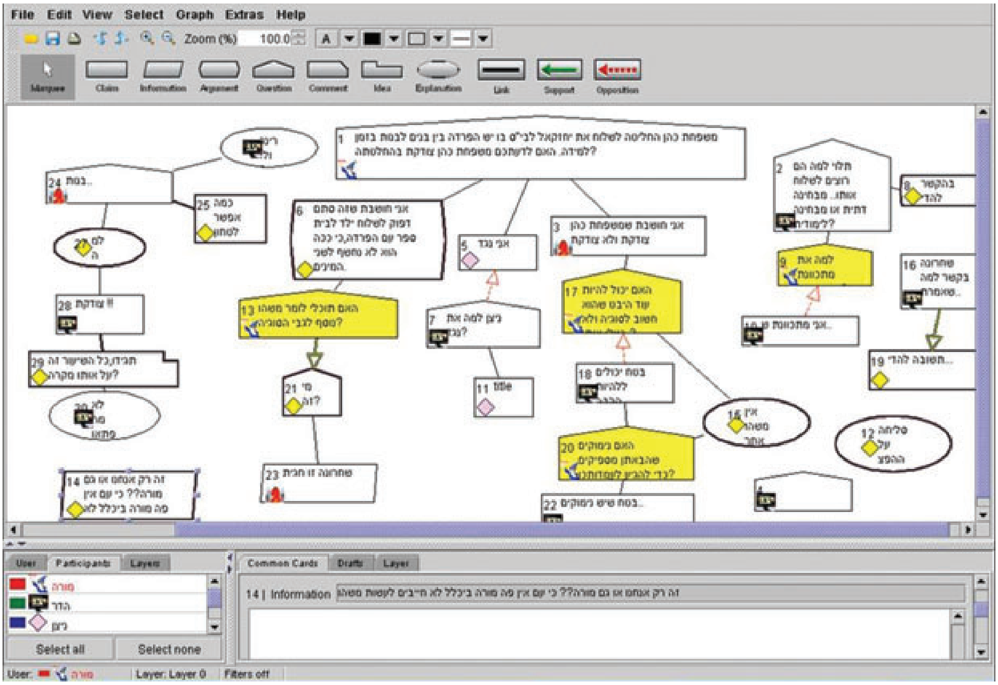

Figure 1. An example of a moderated Digalo discussion map, with yellow-coloured moderator interventions.

students could chose from were ‘idea’, ‘claim’, ‘explanation’, ‘argument’, ‘comment’, and ‘question’. Together with the three different types of arrows (opposing, supporting, neutrally linking), a variety of different argumentative moves are covered. The output from this activity is then a collaboratively constructed argumentative diagram. Figure 1 shows an e-discussion between four male students, for illustration.

In this map, the upper bar displays the pallet of tagged shapes to be chosen from. The lower left window displays the icons for each of the four discussants that are attached to each shape in the map. Discussants write the title of their contribution in the title rubric (visible at all times). The content of their contribution, on the other hand, is visible when hovering over a shape or by opening a shape by double-clicking it. The yellow shapes in the map are the teacher-posted interventions.

The advantages of and different uses of diagram-based tools in learning is not the focus of the present study. However, it is important to point out that in contrast to several other approaches (e.g., Janssen, Erkens, Kirschner, & Kanselaar, 2010; Lund, Mollinari, Séjourné, & Baker, 2007; Van Amelsvoort, Andriessen, & Kanselaar, 2007) students in the present study *communicate* through the argumentative map, and therefore, the argumentative map is both the representation as well as the communication mode.

Educational materials

Each of the three participating homeroom teachers developed three different learning cases on student-relevant social dilemmas of their choice and implemented them during the weekly 45-min long homeroom class. In the present study we report on the second case, implemented during the winter semester. The three winter semester cases were on the following topics: (1) sex-segregation in secondary schools; (2) sex education; and (3) dieting behaviours. Each case was based on the same sequence of learning activities, which included among others a preparatory phase of guided knowledge-gathering on the topic and one session of teacher-moderated, online group discussions. Materials in this phase were various and included, among others, movies, presentations, school book texts, and conflicting newspaper articles and were infused with face-to-face instruction on argumentation. Following this knowledge-gathering phase, the teacher asked a question to be discussed in groups of three to four discussants in the Digalo discussion environment. The question had not been directly addressed in the preliminary stage. Place limitations do not allow a full description of all the materials for each of the three cases. As an illustration, we present here the main materials for the learning activities on the topic of sex-segregation in secondary schools:

(1) Three newspapers articles in favour of or against sex-segregation in schools.
(2) A TV broadcast that presented a simulation of a discussion in a legislative commission about sex-segregation in science education.
(3) A classroom activity on how to extract arguments from the three written articles.

Following the preparatory stages, students were asked to discuss the following question in small-group Digalo discussions: 'The Cohen family decided to send their son Ezekiel to a school in which girls and boys are separated into same-sex science classes. Do you think that the Cohen family has taken the right decision?' It is worth noting that the 9th grade science and homeroom classes in the participating school were sex-segregated as part of a pilot project initiated by the Ministry of Education. The school administration was openly in favour of sex-segregated science instruction as a means to increase academic involvement amongst girls and boys, and this rationale was also communicated to the parents. Even though this policy decision could have been presented as part of a feminist agenda, it was not done so in the classroom activities: the different materials that were read and discussed in class presented a range of reasons for and against sex-segregated education and considered the (dis-)advantages for female and male students alike in a balanced manner. Our interpretation is further corroborated by the fact that the discussions showed similar opinion distributions for boys and for girls, that is, they were almost unanimously against sex segregation.

Procedure

In the weekly 45-min session, students not only learned about student-relevant social dilemmas, but were also taught about argumentation. Argumentation instruction was semi-controlled across classrooms and homeroom teachers: All three teachers adhered to the content and lesson plans that they collaboratively designed for this purpose together with the Kishurim pedagogical team. Among others, students learned the basics of argumentation (e.g., the distinction between a reasoned and an unreasoned argument), were trained to extract arguments from texts and conducted several computer-mediated, small-group discussions throughout the year. As part of an on-going project in the school, classes were sex-segregated for this weekly homeroom class. Classroom size was therefore relatively small (between 12 and 16 students per class), but of regular size for computer lab sessions. The discussion sessions that we report on here, took place in the school's computer lab, and lasted between 25 and 35 min. Each student was individually seated at and communicated through a personal computer. Students had already mastered the use of the Digalo environment during the first case activities, implemented in the Fall, but had not experienced teacher moderation yet. In each of the six different sex-segregated classes, students were organized in four groups of three to four students. Thus, in each class, four different groups of students participated in four different, yet parallel discussions at the same time. Group members were physically spread out over the computer lab facility to insure that two discussants in the same group did not sit next to one another. The groups were formed by the teachers according to social and academic heterogeneity.

In each class, the homeroom teacher was randomly assigned one discussion group and experimented for the first time with guiding an online discussion as part of her in-service training. The focus of the homeroom teacher moderation activities was to gain first experience with navigating the system, monitoring student and group progress and sending interventions in an intuitive manner. As is common among most teachers, the experience of being able to read and closely monitor student–student group discussions and to intervene in an online environment was a novel experience for each of the three homeroom teachers, one that requires quite some getting used to, both in logistic as well as pedagogical terms. The three remaining groups in each classroom sessions were randomly assigned to one of the three following conditions: (1) no guidance; (2) human guidance with a clear focus on the epistemic actions and components of argumentation (Epistemic Guidance for Argumentation, or: EGA); and (3) human guidance with a clear focus on interactional aspects of collective argumentation (Interactional Guidance for Argumentation, or: IGA).

Assignment of the three guest teachers to the different guidance conditions was counterbalanced within condition and gender. Teachers were co-located in the same computer lab as students, but students did not know which of the three present adult teachers was moderating the group session they participated in. Potential differences with regards to personal style were minimized with the help of a set of general instructions and pre-determined prompts to be used by all three teachers.

Experimental intervention

The general instructions were as follows: 'Please, read the contributions of each of the students and identify opportunities for intervention when (a) one of the claims is not reasoned; (b) no counterargument has been raised; (c) no additional perspectives have been raised; (e) an argument is not clear; (f) a student is idle'.

In each of the two experimental conditions, guest teachers then received different instructions on the goal of moderation and the type of pre-defined prompts they could use. For the EGA condition, instructions were as follows:

> 'In this experiment, our goal is to guide discussions and encourage students to raise counterarguments and multiple perspectives. Please use any of the following prompts for your interventions: (1) Prompts that target clarification of reasoning*:* What do you mean? Can you clarify yourself? Can you please elaborate this issue further?; or (2) Prompts that aim at broadening the discussion space with different perspectives*:* What other alternative exist? Is there another point of view that hasn't been considered yet?'

For the IGA condition, instructions were as follows:

> 'In this experiment, our goal is to guide discussions and encourage students to react to each other. Please use any of the following prompts for your interventions: Please relate to what X said. Does anyone like to respond to the contribution of X? Does anyone want to oppose

to X? Does anyone want to question the contribution made by X or add to it? Does anyone want to strengthen X's view or ask X a question?'

In both conditions, guest teachers were furthermore instructed to intervene at least three times but to refrain from being too intrusive, in order to avoid that they would interrupt the flow of the discussion. If students would not react, they were instructed to add the following comment outside the chain of reasoning: 'Boys/Girls, please consider my suggestions!'

In summary, the experiment involved 12 groups of three to four boys and 12 groups of three to four girls. Instructors posted 4.17 (EGA) and 4.5 (IGA) interventions on average, and the content of all the interventions was in compliance with the instructions. No significant differences in number of moderation moves were found between IGA and EGA condition ($p = .622$) or between the three guest moderator teachers ($p = .879$). Because of the clear training purpose of the homeroom teachers involvement in e-moderation and the lack of clear guidelines for its content, the six discussion maps that were assigned to the homeroom teachers were not included for analyses that targeted questions concerning the experimental effect of epistemic and interaction guidance. Only the discussions that received epistemic guidance ($N = 19$; $K = 6$), interaction guidance ($N = 22$; $K = 6$), and no guidance ($N = 21$; $K = 6$) were included for these analyses, where N is the number of student participants and K is the number of groups. Since analyses that examined questions of gender focused on existing differences in all-boy and all-girl groups, these analyses included all 24 discussion maps.

Coding procedures

Twenty-four discussion maps were collected and analysed. To analyse the maps, we developed a methodology inspired by Rourke, Anderson, Archer, and Garrison (1999), Lotan (2006), and by dimensions proposed by Lund (2004). We focused on the *argumentative* quality of the individual contributions and the collective dialogue outcome, as well as on the *collaborative* aspects of discussions (participation and interaction).

Three different grains of analysis were discerned: The micro-level concerns argumentative moves, which are concretized within the discussion environment as the choice of a tagged shape, its title and textual content, and the arrow linking it to another contribution. The meso-level concerns a chain of argumentative moves linked with arrows. This level helps scrutinizing the development of co-constructed arguments. The macro-level concerns features of the map as a whole. It helps to determine aspects of the collective product, that is, the dialogue as an outcome.

Concerning the dimension of *argumentative quality*, only on-task contributions were considered: At the micro-level, a message was identified as a *claim* when it expresses only an unreasoned claim (e.g., 'I oppose') or viewpoint and as an *argument* when it expresses a claim and (a) reason(s) supporting it. We also considered verbal content as an argument when it expressed a conclusion and (a) justification(s) or an assumption and (a) conclusion(s) (Kuhn & Udell, 2003; Scriven, 1976). An argument was identified as a *simple* argument if it included one reason/justification only (e.g., 'I believe that it is not a good idea to send the kid to a segregated school, because that way he will never be exposed to children from both sexes'). When it included more than one reason/justification or a conditional argument it was identified as a *complex argument* (e.g., 'I believe that the family is right and is not right. On the one hand, it is

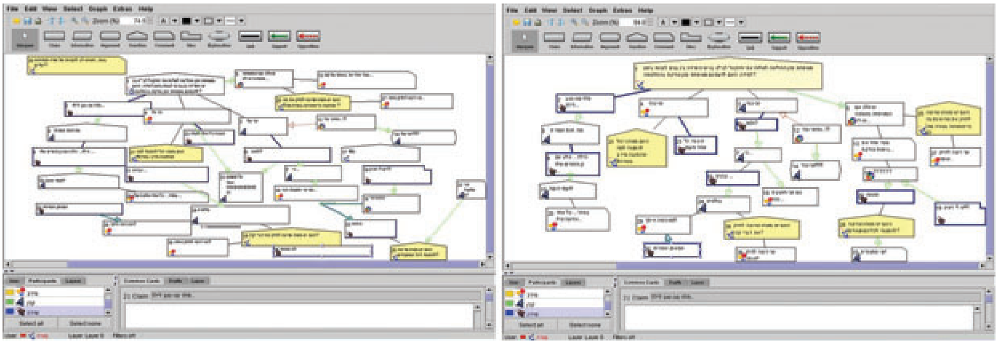

Figure 2. An original Digalo discussion map (on the left) and the same map after its re-organization according to chains (on the right).

good, because it will give him the opportunity to express himself better in class. (. . .) On the other hand, the boy will need to learn how to cope with women later on in his life').

At the meso-level, we searched for chains of reasoning including a continuum of reasons/justifications and other argumentative moves (Asterhan & Schwarz, 2009a) in which more than one discussant contributed. This continuum may include links of any kind (supporting, opposing, or neutral) and should include at least two on-task contributions by different students.

To identify chains of reasoning, we first expurgated from the map all the off-task and the not content-related messages. Figure 2 shows graphically how chains of reasoning were identified. This enabled us to identify the number of chains of shared reasoning. A distinction was made between chains in which a collaborative argument was developed (when a claim was collectively developed into an argument or into a complex argument), to those in which no argument was developed, and only counted the former. On the map level, the number of different perspectives on the topic of discussion raised by the group was assessed.

One third of the maps were analysed by two independent raters, blind to condition. The inter-rater reliability for those categories that required human interpretation of dialogue content ranged from Kappa Cohen's coefficient $\kappa = .77$ to $\kappa = .84$.

As for the *collaborative dimension* of the peer discussions, we focused on student participation and interactivity. Once more, only on-task contributions were considered. The degree of *participation* was measured by the average number of textual contributions and links per student. *Interactivity*, on the other hand, refers to the extent to which student discussants interacted with each other. Operational measures of interactivity at the individual student (micro-)level are the mean absolute number of links a student made to contributions of fellow students and the mean number of contributions per student that received no response from anyone. Two additional measures were added to assess interactivity on the map (macro-) level: *Connectivity*, which is an indicator of the overall connectedness of the discussion map (independent of the connection creators' identity), is operationally defined as the ratio of links and shapes. *Interaction density*, on the other hand, is an indicator often used in social network analyses and refers to the frequency with which any two students within a group interacted with each other, corrected for the possible number of different student–student interaction combinations (Scott, 1991). For example, in a group of four, the number of possible interaction combinations between

Table 1. Argumentative features of online Digalo discussions, by type of human guidance condition (epistemic, interaction, or none) and discussion dimension

	Interaction guidance for argumentation ($N = 22$; $K = 6$)		No human guidance ($N = 21$; $K = 6$)		Epistemic guidance for argumentation ($N = 19$; $K = 6$)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Individual level variables						
Unreasoned claims p.p.	1.45	1.30	1.05	1.12	0.63	0.69
Simple reasoned arguments p.p.	0.86	0.77	0.90	0.77	1.74	1.20
Complex reasoned arguments p.p.	0.68	0.99	0.48	0.87	0.32	0.75
Group level variables						
Extent of critical referencing (supportive/ opposing links)	3.31	4.78	2.13	2.44	1.33	1.54
Chains with argument construction	1.50	1.05	1.17	1.17	1.67	1.21
Number of different perspectives	4.50	1.22	4.83	1.72	4.33	1.63

any two students is six. Interaction density was operationally defined as the number of actual interactions (links) between any two discussants, divided by the maximum number of possible student–student combinations (links) in the group.

Results

Overview of the analyses

To test our hypotheses concerning the effect of human guidance that focuses either on the epistemic or the interactional dimension of collaborative argumentation, we compared the discussion features of the maps in each of these two experimental conditions with the control condition (no human moderation) (see Tables 1 and 2). Gender differences were explored by comparing the 12 all-female discussion maps with the 12 all-male discussion maps (see Table 3).

In line with previous work by others (e.g., Arvaja, Salovaara, Hakkinen, & Jarvela, 2007), a combined analysis approach, separately analyzing the individual and group characteristics of interaction behaviour: comparisons on the dialogue variables that were computed at the level of the individual student (e.g., number of textual contributions per person, number of simple claims per person, and so on, see Tables 1 and 2) were conducted with single-factor MANOVA analyses for the composite of the Argumentative Dimension variables and that of the Collaborative Dimension variables separately, for each of the three comparisons (EGA vs. Control, IGA vs. Control, and Female vs. Male groups). To further explore which of the different dependent variables contributed to a multivariate effect within a dimension, step-down F analyses (Tabachnick & Fidell,

Table 2. Collaboration features of online Digalo discussions, by type of human guidance condition (epistemic, interaction, or none) and discussion dimension

	Interaction guidance for argumentation ($N = 22$; $K = 6$)		No human guidance ($N = 21$; $K = 6$)		Epistemic guidance for argumentation ($N = 19$; $K = 6$)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Individual level variables						
Textual contributions created p.p.	6.14	3.14	4.19	2.09	4.63	2.63
Links created p.p.	7.38	3.86	3.77	1.91	5.08	2.62
Links to fellow student contributions p.p.	4.33	1.25	2.67	1.93	2.58	1.84
Unlinked shapes p.p.	0.45	0.80	0.71	0.90	0.58	0.84
Group level variables						
Connectivity (shapes / links)	0.95	0.28	1.08	0.18	0.97	0.25
Interaction density	3.44	1.46	2.08	1.06	2.44	1.77

2006) were conducted after each MANOVA analysis. Step-down *F* analyses are a series of analyses of covariance (ANCOVAs) in which the dependent variables are evaluated in terms of their unique overlap with the independent variable in order of their theoretical importance.

Since five of the dependent variables in this study referred to discussion behaviour characteristics that are inherently *group* characteristics (i.e., critical referencing, chains

Table 3. Discussion characteristics of online Digalo discussions by gender (all-girls or all-boys groups)

	All-female ($N = 38$; $K = 12$)		All-male ($N = 44$; $K = 12$)	
	M	*SD*	*M*	*SD*
Argumentative dimension				
Unreasoned claims p.p.	1.32	1.32	0.89	0.84
Simple reasoned arguments p.p.	1.26	0.95	0.95	0.88
Complex reasoned arguments p.p.	0.89	1.09	0.32	0.60
Extent of critical referencing (supportive / opposing links)	2.98	4.14	2.10	1.99
Chains with argument construction	1.67	0.98	1.00	0.95
Number of different perspectives	5.25	1.76	3.91	1.31
Collaborative dimension				
Textual contributions created p.p.	6.13	2.61	3.84	2.19
Links created p.p.	5.95	3.36	4.36	4.26
Links to fellow student contributions p.p.	3.47	2.29	2.52	2.31
Unlinked shapes p.p.	0.71	0.87	0.52	0.88
Connectivity (shapes / links)	1.06	0.26	1.03	0.24
Interaction density	3.15	1.21	1.88	1.38

of argument construction, number of perspectives, interaction density, and interconnectivity) analyses of effects for these variables were conducted with separate models that used the group, as opposed to the individual student, as the unit of analysis. Due to the decrease in number of observations on these group variables, comparisons were conducted with one-tailed *t*-tests (EGA vs. Control, IGA vs. Control) and two-tailed *t*-tests (Female vs. Male groups), with Holm-Bonferroni corrections for alpha values (Holm, 1979) within each dimension.

It should be noted that on four of the different measures (namely, number of simple claims, ratio of consensual/critical referencing, number of unlinked shapes, and connectivity) a lower numerical value indicates a higher measures of argumentative quality or interactivity, respectively.

Students taught by different teachers were not found to significantly differ on any of the composite or discrete dialogue variables. The identity of homeroom teacher was therefore omitted from further analyses.

The effect of Epistemic Guidance for Argumentation

The data in Tables 1 and 2 show that students who received guidance with a focus on the epistemic nature of argumentation posted more reasoned arguments, were overall more critical towards ideas, and more often collaboratively constructed chains of reasoned argument. In contrast, they did not differ very much on the interactional and participation dimensions. Indeed, a significant multivariate main effect was found for the composite of the three individual student level variables that relate to the Argumentative Dimension of discussion participation, Wilks' $\lambda = .803$, $F\ (3, 36) = 2.94$, $p = .046$, with a large effect size of partial $\eta^2 = .197$, but not for the composite of the four Collaborative Dimension's variables, Wilks' $\lambda = .887$, $F\ (4, 35) = .36$, *ns.* Since EGA focuses first and foremost on providing reasoned arguments, subsequent step-down *F* analyses were prioritized in the following order: number of reasoned arguments, followed by number of unreasoned claims and complex arguments. It was found that students that received EGA posed more reasoned arguments than those in the control condition, $F\ (1, 38) = 7.00$, $p = .012$, partial $\eta^2 = .156$. Step-down ANCOVA analyses on the other two dependent variables in this category did not reveal any significant differences, $F\ (1, 37) = 1.33$, ns and $F < 1$ for unreasoned claims and complex arguments, respectively. In addition, the observed differences on the three Argumentative Dimension group level variables did not reach statistical significance at the omnibus alpha level of .05.

The effect of Interactional Guidance for Argumentation

An opposite pattern was found for the IGA condition: a significant multivariate main effect was found for the composite of the four individual variables of the Collaborative Dimension, Wilks' $\lambda = .945$, $F\ (4, 38) = 2.92$, $p = .034$, partial $\eta^2 = .235$ but not for the composite of the three Argumentative Dimension variables, Wilks' $\lambda = .887$, $F\ (4, 39) = .76$, *ns.* Table 2 shows that compared to the control condition, student discussion behaviour in the IGA condition was characterized by higher indices of participation and interactivity on all four individual measures of the Collaborative Dimension. Since the IGA prompts first and foremost encouraged students to interact and to relate to the contributions of fellow students, subsequent step-down *F* analyses were conducted in the following order: Number of links to fellow student contributions, number of unlinked shapes, number of links, and number of textual contributions. Students in the

IGA condition were indeed found to have reacted more often to the contributions of fellow students, $F\,(1, 41) = 4.88$, $p = .033$, partial $\eta^2 = .106$. After controlling for number of links to other students and unlinked shapes, the observed difference in the overall number of links posted was found to be marginally significant, $F\,(1, 39) = 3.97$, $p = .056$, partial $\eta^2 = .090$. Differences on the remaining two dependent variables in this category failed to reach significance after controlling for inter-collinearity between the dependent variables.

As for the group level variables, discussion groups in the IGA condition were overall less critical towards ideas, but interlinked their contributions more (interconnectivity) and interacted more with each of the group members (interaction density). Only the observed difference in interaction density was found to be statistically significant, $t\,(10) = 1.86$, $p = .046$, $d = 1.07$, after Holm-Bonferroni corrections for multiple comparisons within each dimension.

Gender differences

We then turned to a comparison between the discussion characteristics of all-male and all-female groups (see Table 3).

The data in Table 3 show that congruent with our expectations, the discussion behaviour of girls was overall characterized by higher indices of participation and interaction: they participated more frequently (number of links and of textual contributions) and interacted more frequently with their fellow discussants (links to fellow students, interaction density), even though all-boys discussion maps were slightly better connected and had less unlinked shapes.

A significant multivariate effect was found for the composite of the four individual student level Collaborative Dimension variables, Wilks' $\lambda = .792$, $F\,(4, 77) = 5.05$, $p = .001$, with a strong effect size (partial $\eta^2 = .208$). Since the literature review on gender differences in online environments most frequently refers to the participation dimension, subsequent step-down F tests were conducted in the following order of dependent variables: Number of textual contributions, links, links to fellow student contributions and unlinked contributions. Step-down F tests showed that the observed gender differences on the four individual Collaborative dimension variables were significant for the number of textual contributions only, $F\,(1, 80) = 18.68$, $p < .001$, with a large effect size (partial $\eta^2 = .189$).

As for the observed gender differences on the group level characteristics, of the two Collaborative Dimension variables only the difference in interaction density reached significance, $t\,(22) = 3.04$, $p = .006$, with a large effect size of $d = .98$.

As for gender differences with regards to the Argumentative Dimension of student online discussion behaviour, the data in Table 3 show a slightly more complex picture: girls were observed to have posted a larger number of simple and complex reasoned arguments, co-constructed more collective argument chains and considered a larger number of perspectives. Boys, on the other hand, posted a smaller number of unreasoned claims and were overall slightly more critical.

A significant multivariate main effects was found for the composite of the three individual level Argumentative Dimension variables, Wilks' $\lambda = .792$, $F\,(3, 78) = 6.821$, $p < .001$, with a large effect size (partial $\eta^2 = .208$). Step-down F test comparisons were conducted for number of unreasoned claims, simple reasoned arguments, and complex arguments, in that order. Even when controlled for the other two variables, girls were found to have posted a statistically significant larger number of complex arguments than boys, F (1, 78) = 13.09, p = .001, partial η^2 = .144. The observed differences on the other two dependent variables did not reach statistical significance.

As for the gender differences observed on the three group level variables of argumentative quality, after Holm-Bonferroni correction of alpha values for multiple comparisons only the difference in the number of perspectives was found to be statistically significant, t (22) = 2.35, p = .028, with a large effect size of d = .86

Discussion

The blending of discussion tools is becoming increasingly more common in secondary school settings. Questions on how to effectively monitor and guide group reasoning in these environments have then become pertinent. By means of the present study, we hope to have contributed a first step towards answering that question in the specific case of online human guidance of group argumentation on societal dilemmas in middle-school classrooms. The findings presented here show that online human guidance of synchronous discussions is a feasible practice for which different moderation goals and prompts have an impact that reasonably fits their intentions. The different types of human guidance addressed in this study focused either on the epistemic or the interactional dimension of collaborative argumentation. Compared to a control condition, we found that each type of human guidance improved discussion features with regards to the dimension they focused on, but did not affect the other dimension.

The finding that interactional guidance of argumentation did not result in better argumentative quality of the discussion is rather surprising. Previous studies have shown that software-embedded collaboration scripts that focus on these very same features were successful in improving argumentative quality of small-group CMC (Weinberger *et al*., 2005; Weinberger, Stegmann, Fischer, & Mandl, 2007). To some extent, we believe that this difference can be explained by the human factor as well as the communication format used in the different studies: In the CSCL scripting approach, implicit guidance is embedded in the computer software, whereas in our approach a human being explicitly instructs students to interact in a certain way. Discussants may interpret the latter as too intrusive or as interfering with their ongoing discussion. Secondly, unlike in the present study, the mode of group communication in the Weinberger *et al*. studies was a-synchronous. In the high pace of a synchronous group discussion, guidance efforts that focus on regulating the interaction may go by unnoticed and may be disregarded easily. It is possible that instructor prompts in synchronous environments will have to be more salient and more explicit with regards to their intention (Asterhan & Schwarz, 2010). If the goal of guidance is to improve the quality of student argumentation, then direct prompts that target these aspects may be more effective. Further research on the difference between CSCL scripting and human guidance will have to show whether the different findings we reported on here can be generalized, or not.

Our findings concerning gender difference in computer-mediated small-group argumentation are intriguing and open up potentially new research venues in the study of argumentation. Gender differences in favour of girls were found both on the argumentative as well as the collaborative dimension of the discussions. Whereas the advantage of girls over boys on the collaborative dimension was expected, the finding that the all-girls discussions were also of an overall better argumentative quality is surprising. More research is needed to see whether this finding can be replicated and generalized to other settings and populations and if so, what the reasons behind this difference could be. One particular interesting venue, which we have recently begun to explore, concerns the question whether male and female students differ in the extent to which they endorse collaborative or competitive interaction goals and how this in turn affects online argumentation (Asterhan *et al.*, 2011).

Limitations of the current research

The current study tested the effectiveness of an innovational educational practice in authentic, *in vivo* research settings. The introduction of novel educational practices should be done carefully and preferably be tested in small-scale, authentic settings first. However, this type of research is also characterized by small sample sizes and several other statistical limitations. Most importantly, the statistical analyses presented in this study could not take into account the occurrence of nested effects of the individual student within the discussion group. Researchers of collaborative learning have become increasingly aware of the need for statistical solutions that can adequately cope with this issue. Even though different researchers have chosen to address it in different ways (see Strijbos & Fischer, 2007, for an overview), none of these fully address the issue in a satisfying way. Ideally, the application of Hierarchical Linear Modelling (HLM) techniques should be preferred whenever possible. However, statisticians have warned against using HLM for data sets that include small-group sizes that have less than 10 observations per group (e.g., Clarke, 2008; Clarke & Wheaton, 2007). Since in collaborative learning settings group sizes typically range between two and six students, these models are then currently irrelevant for analyses in this area of research[2] . Moreover, these models require sample sizes that are unrealizable for studies that test the effectiveness of a novel educational innovation in authentic settings. Since some of the dialogue characteristics assessed in this study were inherently group variables (e.g., interaction density, number of collaboratively constructed chains of reasoning) and others focused on characteristics of individual behaviour, we then chose to adopt a 'combined approach' (e.g., Arvaja *et al*., 2007) and analysed each separately. Hopefully, future developments in statistical modelling will provide an adequate solution for dealing with nested effects in small-group research, and allow for multi-level analyses in one single model.

This study focused on a new practice, namely online teacher guidance of synchronous student discussions. It was argued that the rigidity of software-embedded collaborative scripts imposes limitations on its ability to provide in-time, adaptive support, especially in synchronous communication formats. We believe that online human guidance presents a promising and feasible alternative to the scripting approach that is worthy of further research. However, in order to become practically *viable in classrooms*, instructors should be able to monitor and guide several on-going discussions in parallel. Recent development initiatives, such as the ARGUNAUT project have focused on designing environments that support the instructor in this endeavour, by providing him/her with, among others, awareness indicators at-a-glance, an alerting system, and a user-friendly intervention panel with predefined messages (De Groot, 2010; McLaren, Scheuer, & Mikšátko, in press; Schwarz & Asterhan, 2011). In a recent case study (Schwarz & Asterhan, 2011), we have shown how one teacher could capitalize on this support and successfully moderate multiple synchronous discussions, by deploying various

[2] *Unless collaboration partners are replaced by confederates whose actions are tightly controlled under laboratory conditions (e.g., Asterhan & Schwarz, 2007).*

strategies in a flexible and adaptive manner that was contingent upon student and group needs. We are then confident that future developments in educational and instructional technology will provide the tools necessary for scaling up and enabling online human guidance in genuine classroom settings. We hope that the present study has contributed to our understanding of how different types of online guidance may affect discussion quality, so that ultimately instructors may be better informed when making decisions about which guidance strategy to use to achieve the desired goal.

Acknowledgment

The research reported in this paper was conducted within the framework of the ARGUNAUT project, funded by European Community grant IST-2005-027728.

References

Arvaja, M., Salovaara, H., Hakkinen, P., & Jarvela, S. (2007). Combining individual and group-level perspectives for studying collaborative knowledge construction in context. *Learning and Instruction*, *17*(4), 448–459. doi:10.1016/j.learninstruc.2007.04.003

Asterhan, C. S. C. (2011). Assessing e-moderation behavior from synchronous discussion protocols with a multi-dimensional methodology. *Computers in Human Behavior*, *27*, 449–458. doi:10.1016/j.chb.2010.09.008

Asterhan, C. S. C., Butler, R., & Schwarz, B. B. (2011). Peer argumentation that supports conceptual learning: The role of individual achievement goals, goal instructions and gender. Manuscript submitted for publication.

Asterhan, C. S. C., & Eisenmann, T. (2009). Online and face-to-face discussions in the classroom: A study on the experiences of 'active' and 'silent' students. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Computer-supported collaborative learning practices: CSCL2009 conference proceedings* (pp. 132–136).

Asterhan, C. S. C., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, *99*, 626–639. doi:10.1037/0022-0663.99.3.626

Asterhan, C. S. C., & Schwarz, B. B. (2009a). The role of argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialogue. *Cognitive Science*, *33*, 373–399. doi:10.1111/j.1551-6709.2009.01017.x

Asterhan, C. S. C., & Schwarz, B. B. (2009b). Transformation of robust misconceptions through peer argumentation. In B. B. Schwarz, T. Dreyfus, & R. Hershkowitz (Eds.), *Guided transformation of knowledge in classrooms* (pp. 159–172). New York: Routledge, Advances in Learning & Instruction series.

Asterhan, C. S. C., & Schwarz, B. B. (2010). Online moderation of synchronous e-argumentation. *International Journal of Computer-Supported Collaborative Learning*, *5*, 259–282. doi:10.1007/s11412-010-9088-2

Baxter-Magolda, M. B. (1992). *Knowing and reasoning in college – gender related patterns in students' intellectual development*. San Francisco: Jossey-Bass Publishers.

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986) *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.

Berge, Z. L. (1995). Facilitating computer conferencing: Recommendations from the field. *Educational Technology*, *15*, 22–30.

Cameron, D. (1998). Language, gender and discourse: A review. *Signs: Journal of Women, Culture and Society*, *23*(4), 945–973. doi:10.1086/495297

Caspi, A., Chajut, E., & Sapporta, K. (2008). Participation in class and in online discussions: Gender differences. *Computers & Education*, *50*, 718–724. doi:10.1016/j.compedu.2006.08.003

Chi, M. T. H., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science*, *25*, 471–534. doi:10.1207/s15516709cog2504_1

Chiu, M. M. (2004). Adapting teacher interventions to student needs during cooperative learning: How to improve student problem solving and time on-task. *American Educational Research Journal*, 41, 365–399. doi:10.3102/00028312041002365

Cho, K.L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology: Research and Development*, *50*, 5–22. doi:10.1007/BF02505022

Clarke, P. (2008). When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. *Journal of Epidemiology and Community Health*, *62*,752–758. doi:10.1136/jech.2007.060798

Clarke, P., & Wheaton, B. (2007). Addressing data sparseness in contextual population research using cluster analysis to create synthetic neighborhoods. *Sociological Methods & Research*, *37*, 311–351. doi:10.1177/0049124106292362

Clinchy, B. M. (1989). The development of thoughtfulness in college women: Integrating reason and care. *American Behavioral Scientists*, *32*, 647–657. doi:10.1177/0002764289032006005

Cohen, P. R. (1984). The pragmatics of referring and the modality of communication. *Computational Linguistics*, *10*(2), 97–124.

Cress, U., Kimmerle, J., & Hesse, F. W. (2009). Impact of temporal extension, synchronicity, and group size on computer-supported information exchange. *Computers in Human Behavior*, *25*, 731–737. doi:10.1016/j.chb.2009.01.010

De Groot, R. (in press). Teachers' use of the Argunaut system in the classroom. In N. Pinkwart & B. M. McLaren (Eds.), *Educational technologies for teaching argumentation skills*. Sharjah: Bentham Science Publishers.

Dekker, R., & Elshout-Mohr, M. (2004). Teacher interventions aimed at mathematical level raising during collaborative learning. *Educational Studies in Mathematics*, *56*, 39–65. doi:10.1023/B:EDUC.0000028402.10122.ff

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 61–91). Heerlen: Open Universiteit Nederland.

Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for computer-supported collaborative learning. *Journal of Computer Assisted Learning*, *23*, 1–13. doi:10.1111/j.1365-2729.2007.00191.x

Fischer, F., Kollar, I., Mandl, H., & Haake, J. (Eds.). (2007). *Scripting computer-supported communication of knowledge: Cognitive, computational, and educational perspectives*. New York: Springer.

Galotti, K.M., Clinchy, B.M., Ainsworth, K.H., Lavin, B., & Mansfield, A.F. (1999) A new way of assessing ways of knowing: The Attitudes Toward Thinking and Learning Survey (ATTLS). *Sex Roles*, *40*, 745–766. doi:10.1023/A:1018860702422

Galotti, K.M., Drebud, D.W., & Reimer, R.L. (2001). Ways of knowing as learning styles: Learning MAGIC with a partner. *Sex Roles*, *44*, 419–436. doi:10.1023/A:1011978011991

Gillies, R. M. (2006). Ten Australian teachers' discourse and reported pedagogical practices during cooperative learning. *Elementary School Journal*, *106*, 429–451. doi:10.1086/505439

Goodyear, P., Salmon, G., Spector, J. M., Steeples, C., & Tickner, S. (2001). Competences for online teaching: a special report. *Educational Technology, Research and Development*, *49*(1), 65–72. doi:10.1007/BF02504508

Guiller, J., Durndell, A., & Ross, A. (2008). Peer interaction and critical thinking: Face-to-face or online discussion? *Learning and Instruction*, *18*, 187–200. doi:10.1016/j.learninstruc.2007.03.001

Gunn, C., McSporran, M., Macleod, H., & French, S. (2003). Dominant or different? Gender issues in computer supported learning. *Journal of Asynchronous Learning Networks*, 7, 14–30.

Herring, S. (2001). Computer-mediated discourse. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 612–634). Oxford: Blackwell Publishers.

Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In: S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for Virtual Communities in the Service of Learning* (pp. 338–376). New York: Cambridge University Press.

Hesse, F. W. (2007). Being told to do something or just being aware of something? An alternative approach to scripting in CSCL. In F. Fischer, I. Kollar, H. Mandl, & J. Haake (Eds.), *Scripting computer-supported collaborative learning. Cognitive, computational and educational perspectives* (pp. 91–98). New York: Springer-Verlag.

Hlapanis, G., Kordaki, M., & Dimitrikapoulou, A. (2006). Successful e-courses: The role of synchronous communication and e-moderation via chat. *Campus-wide Information Systems*, *23*, 171–181.

Hogan, K., Nastasi, B. K., & Pressley, M. (2000). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction*, *17*, 379–432.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.

Hron, A., Hesse, F. W., Cress, U. & Giovis, C. (2000). Implicit and explicit dialogue structuring in virtual learning groups. *British Journal of Educational Psychology*, *70*, 53–64. doi:10.1348/000709900157967

Janssen, J., Erkens, G., Kirschner, P., & Kanselaar, G. (2010). Effects of representational guidance during computer-supported collaborative learning. *Instructional Science*, *38*, 59–88. doi:10.1007/s11251-008-9078-1

Johnson, D. W., & Johnsson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book Co.

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

Kim, I.-H., Anderson, R. C., Nguyen-Jahiel, K., & Archodidou, A. (2007). Discourse patterns during children's collaborative online discussions. *The Journal of the Learning Sciences*, *16*(3), 333–370. doi:10.1080/10508400701413419

Koedinger, K., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–78). Cambridge MA: Cambridge University Press.

Kollar, I., Fisher, F., & Hesse, F. W. (2006). Collaboration scripts – a conceptual analysis. *Educational Psychology Review*, *18*, 159–185. doi:10.1007/s10648-006-9007-2

Kuhn, D. (1991). *The skills of argument*. Cambridge, MA: Cambridge University Press.

Kuhn, D. & Udell, W. (2003). The development of argument skills. *Child Development*, *74*, 1245–1260. doi:10.1111/1467-8624.00605

Lentell, H., & O'Rourke, J. (2004). Tutoring large numbers: An unmet challenge. *International Review of Research in Open and Distance Learning*, *5*. Retrieved from: http://www.irrodl.org/index.php/irrodl/article/view/171/253

Lotan, E. (2006). Analyzing graphic-based electronic discussions: Evaluation of student's activity in Digalo. In W. Nejdl & K. Tochtermann (Eds.), *Proceedings of the First European Conference on Technology Enhanced Learning (EC-TEL 2006)* (pp. 652–659). Berlin, Germany: Springer.

Lund, K. (2004). Human support in CSCL: What, for whom and by whom? In J-W. Strijbos, P.A. Kirshner, R. L. Martens, & P. Dillenbourg (Eds.), *What we know about CSCL and implementing it in higher education*, CSCL (Vol. 3, pp.167–198). Norwell, MA: Kluwer Academic Publishers.

Lund, K., Mollinari, G., Séjourné, A. & Baker, M. (2007). How do argumentation diagrams compare when student pairs use them as a means for debate or as a tool for representing debate? *International Journal of Computer-Supported Collaborative Learning*, *2*, 273–295. doi:10.1007/s11412-007-9019-z

Maltz, D. & Borker (1982). A cultural approach to male-female miscommunication. In J. Gumperz (Ed.), *Language and social identity*. (pp.281–312). Oxford: Cambridge University Press.

Mason, R., & Kaye, A. (Eds.). (1989). *Mindweave: Communication, computers and distance education*. Oxford: Pergamon Press.

Mazzolini, M., & Maddison, S. (2003). Sage, guide or ghost? The effect of instructor intervention on student participation in online discussion forums. *Computers & Education*, *40*, 237–253. doi:10.1016/S0360-1315(02)00129-X

McLaren, B.M., Scheuer, O., & Mikšátko, J. (in press). Supporting collaborative learning and e-discussions using artificial intelligence techniques. *International Journal of Artificial Intelligence in Education*.

McPherson, M., & Nunes, M. (2004). The role of tutors as an integral part of online learning support. *European Journal of Open and Distance Learning*. Retrieved from: http://www.eurodl.org/materials/contrib/2004/Maggie_MsP.html

Miller, P. H. (2005). Gender and information technology: Perspectives from human cognitive development. *Frontiers: A Journal of Women Studies*, *26*, 148–167. doi:10.1353/fro.2005.0007

Newman, D. R., Webb, N., & Cochrane, B. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology*, *3*, 56–77.

O'Donnell, A. M. (2006). The role of peers and group learning. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Paloff, R. M., & Pratt, K. (2001). *Lessons from the cyberspace classroom – the realities of online teaching*. San Francisco: Jossey-Bass.

Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, *40*(1), 1–12. doi:10.1207/s15326985ep4001_1

Rogoff, B. (2003). *The cultural nature of human development*. Oxford: Oxford University Press.

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

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences*, *14*, 201–241. doi:10.1207/s15327809jls1402_2

Rummel, N., Spada, H, & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*, 69–92. doi:10.1007/s11412-008-9054-4

Salmon, G. (2004). *E-moderating: The key to teaching and learning online*. London: Taylor & Francis.

Scardamalia, M., & Bereiter, C. (1996). Computer support for knowledge-building communities. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 249–268). Mahwah, NJ: Lawrence Erlbaum Associates.

Schwarz, B. B., & Asterhan, C. S. C. (2011). E-moderation of synchronous discussions in educational settings: A nascent practice. *Journal of the Learning Sciences*. Advance online publication. doi: 10.1080/10508406.2011.553257

Schwarz, B. B., & De Groot, R. (2007). Argumentation in a changing world. *International Journal of Computer-Supported Collaborative Learning*, *2*(2–3), 297–313. doi:10.1007/s11412-007-9020-6

Schwarz, B. B., & Glassner, A. (2007). The role of floor control and of ontology in argumentative activities with discussion-based tools. *International Journal of Computer Supported Collaborative Learning*, *2*, 449–478. doi:10.1007/s11412-007-9024-2

Scriven, M. (1976). *Reasoning*. New York: McGraw-Hill.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, *21*, 43–69. doi:10.1006/ceps.1996.0004

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, *2*, 421–447. doi:10.1007/s11412-007-9028-y

Stokoe, E. H. (2000). Towards a conversation analytic approach to gender and discourse. *Feminism & Psychology*, *10*, 552–563. doi:10.1177/0959353500010004018

Strijbos, J-W., & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and Instruction*, *17*, 389–393. doi:10.1016/j.learninstruc.2007.03.004

Suthers, D. D. (2003). Representational guidance for collaborative inquiry. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 27–46). Dordrecht: Kluwer Academic Publishers.

Tabachnick, B. G., & Fidell L. S. (2006). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn and Bacon.

Tolmie, A., Thomson, J. A., Foot, H. C., Whelan, K., Morrison, S., & McLaren, B. (2005). The effects of adult guidance and peer discussion on the development of children's representations: Evidence from the training of pedestrian skills. *British Journal of Educational Psychology*, *96*,181–2004.

Van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (2007). Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *Journal of the Learning Sciences*, *16*, 485–521.

Veerman, A. L., Andriessen, J. E. B., & Kanselaar, G. (2000). Learning through synchronous electronic discussion. *Computers & Education*, *34*, 269–290.

Walker, A. (2004). Socratic strategies and devil's advocacy in synchronous CMC debate. *Journal of Computer Assisted Learning*, *20*, 172–182. doi:10.1016/S0360-1315(99)00050-0

Weatherall, A. (2000). Re-visions for gender and language research in the 21st century. In: J. Holmes (Ed.), *Gendered Speech in Social Context: Perspectives From Town and Gown* (pp. 39–51). Wellington, NZ: Victoria University Press.

Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *British Journal of Educational Psychology*, *79*, 1–28. doi:10.1348/000709908X380772

Webb, N. M., & Palincsar, A. S. (1996). Group processes in the classroom. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 841–873). New York, NY: Macmillan.

Wecker, C., & Fischer, F. (2007). Fading scripts in computer-supported collaborative learning: The role of distributed monitoring. In C. A. Chinn, G. Erkens & S. Puntambekar (Eds.), Mice, minds and sociecty. Proceedings of the Computer Supported Collaborative Learning (CSCL) Conference 2007, Vol. 8, Part 2 (pp. 763–731). International Society of the Learning Sciences.

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science*, *33*(1), 1–30. doi:10.1007/s11251-004-2322-4

Weinberger, A., Stegmann, K., Fischer, F., & Mandl, H. (2007). Scripting argumentative knowledge construction in computer-supported learning environments. In F. Fischer, I. Kollar, H. Mandl, & J. Haake (Eds.), *Scripting computer-supported communication of knowledge – cognitive, computational and educational perspectives* (pp. 191–211). New York: Springer-Verlag.

Zohar, A. (2006). Connected knowledge in science and mathematics education. *International Journal of Science Education*, *28*, 1579–1600. doi:10.1080/09500690500439199

Received 19 April 2010; revised version received 14 March 2011