T. Hirashima et al. (Eds.) (2011). Proceedings of the 19th International Conference on Computers in Education. Chiang Mai, Thailand: Asia-Pacific Society for Computers in Education

Scripting Collaboration: What Affects Does it Have on Student Argumentation?

Oliver SCHEUER[a*], Bruce M. MCLAREN[a, b], Maralee HARRELL[b], & Armin WEINBERGER[a]
[a]*Saarland University, Saarbrücken, Germany*
[b]*Carnegie Mellon University, Pittsburgh, PA, U.S.A.*
*o.scheuer@mx.uni-saarland.de

Abstract: Computer-mediated environments provide an arena for learning to argue. We investigate to what extent student dyads' online argumentation can be facilitated with collaboration scripts that (1) prompt learners to prepare individually, (2) create conflict, and (3) encourage productive collaboration and argumentation. A process analysis of the chats of the dyads showed that the scripted treatment group used significantly more words and broadened and deepened their discussions significantly more than the unscripted group. Qualitative analysis indicates that scripted learners engaged in more critical and objective argumentation than non-scripted learners.

Keywords: computer-supported collaborative learning, collaboration scripts, argumentation

Introduction

Researchers have been increasingly interested in studying how to use technology to help students learn argumentation skills [17]. For instance, computer-based argument mapping tools have been developed and used to teach argument analysis skills in philosophy classes [8]. Computer-Supported Collaborative Learning (CSCL) researchers have focused on the *interpersonal* dimension of argumentation, including how to compose effective learning groups, how to distribute resources, and how to structure the discussion [1, 26]. Finally, the Artificial Intelligence In Education (AIED) community has investigated how to analyze arguments and how to provide feedback to support learning or argumentation [18] in domains such as science [21], the law [14], and applied ethics [12].

In this paper, we present initial results of our approach to engage student dyads in critical debate in a computer-mediated setting. Their task was to critically review argumentation texts on a controversial issue (global warming ethics) and to jointly agree on a reasoned position. Our main research question is: Will structured student collaboration lead to higher quality argumentation? This work is in the tradition of CSCL work and follows from others who have investigated similar questions of the effect of scripts [26] and structuring [19] on the learning of argumentation.

Our goals are twofold: First, we want to extend the body of empirical results on conflict-oriented argumentation scripts (e.g., [3], [9]) in a discussion domain that focuses on ethical rather than scientific issues. Second, we aim to develop and investigate an effective baseline scripting approach that we will eventually support in *adaptive* fashion in future studies. In this paper we focus on the first of the two goals.

1. Problem and Approach

There is little doubt that collaborative argumentation is a valuable educational activity. Yet, past research has shown that it is not sufficient to simply assign a task to a group of students with no guidance or structuring [26]. Students often lack sufficient argumentation skills to engage in collaborative argumentation and productively resolve conflicts [10]. On a macro level, process losses due to inefficient task coordination often outweigh the advantages of combining forces [20]. On a micro level, students often avoid taking a critical stance towards peers' contributions, instead aiming for quick consensus [24]. To tackle these problems, we devised a structured computer-mediated collaboration approach with three key elements:

E1: Prompting individual preparation. Past research has shown that successful collaboration usually involves a combination of individual and collaborative activities [16]. Individual preparation gives students time to make up their own minds about a controversial issue without social pressure [3]. It also allows students to come up with their own ideas before the ideas of others influence their thinking. Thus, more diverse knowledge resources can be activated and contributed to collaborative argumentation [25]. With a clear picture on a given topic in mind, one gained from careful individual deliberation, students are better prepared to engage in fruitful interaction with others.

E2: Creating conflict. Proponents of the socio-cognitive conflict theory see the attempt to resolve social disagreements as a key component of cognitive development and conceptual learning [7]. To create conflict, we let students first make a decision between two alternatives in the individual preparation phase; in the collaborative decision phase, we pair up students with opposite opinions. To emphasize initial disagreement, we make students aware of their different decisions. Conflicting opinions call for explanations, justifications and collaborative conflict resolution – activities that have been shown to be supportive of learning [13]. Similar tactics aimed at inducing and emphasizing conflict to promote discussions and learning have been used, for instance, in teaching physics [3] and instructional design [9, 23].

E3: Encouraging productive collaboration and discussion norms. The final element is a set of guiding instructions students are asked to read before collaborating with one another. We try to promote the following "productive" behaviors: (1) *a mutual commitment to the starting point* [22]. Students should be aware of what their partners know, believe and argue for. Thus, we ask students to peer-review and discuss the results of the individual phase. (2) *Willingness to criticize the position of others*. "Consensual" groups often achieve only suboptimal results compared to "critical" groups [15]. We encourage students to take a critical position by identifying and discussing possible weaknesses in their partner's contributions. (3) *Constructive synthesis*. We attempt to scaffold collaborative writing by encouraging the following: agreeing on the main thesis, agreeing on the main points supportive of the thesis, agreeing on the distribution of work, and finally, in iterative cycles, writing, peer-reviewing and discussing answer components. In sum, our goal is to test the following hypothesis with our instructional design:

H: Students' collaborative argumentation will be of higher quality when
- students have time to prepare individually **(E1)**,
- a conflict of opinion exists and is emphasized **(E2)**, and
- students receive instructional guidance to encourage productive collaboration and discussion norms **(E3)**.

Aspects of collaboration quality we are interested in include the level of student activity (i.e., number of contributions), a broader and deeper elaboration of content, and a reduction

in rapid, uncritical consensus building [24]. We also intend to explore the learning effects of scripted collaboration – this is, in fact, the ultimate goal of this work – but for this initial study we confine our analysis to hypothesis H.

2. Study Context, Design and Methods

The study was carried out as part of an "Introduction to Philosophy" course at Carnegie Mellon University in Pittsburgh, PA (U.S.A.). Three sessions (with required attendance) were conducted between November 10 and December 3, 2010. A quasi-experimental pretest-intervention-posttest design with two conditions was employed. Students in an early recitation class constituted the comparison group; students in a late recitation class constituted the treatment group.

2.1 Sample

The initial set of participants consisted of 54 students. Since not all students attended all sessions, the final process analysis is based on 38 students: 8 comparison group dyads (16 students; 38% female) and 11 treatment group dyads (22 students; 55% female). The treatment group had a higher percentage of freshmen and sophomores (86%), while a majority of the comparison group students were juniors and seniors (63%). However, the groups were similar in terms of midterm course grades (91% of the treatment group students and 94% of the comparison group students scored A or B), so we assume homogenous abilities across the groups.

2.2 Materials

To encourage interest and a lively debate, we identified two source argumentation texts with conflicting perspectives on a controversial topic: global warming. Further, we selected texts that focus on the *ethical* dimension of global warming rather than the purely scientific dimension (i.e., "Who should be responsible for global warming and what should they do?" vs. "Is global warming a scientific reality?").

Brown [5] argues for large reductions of greenhouse gas emission levels by developed countries. He outlines moral and legal obligations for (in particular) the United States to act, even if developing countries do not, based largely on the principle of distributive justice (summary: "The U.S. has an ethical obligation to act in resolving global warming, since it produces a disproportionate amount of the gases that have led, and will continue to lead, to global warming.").

Lomborg [11] argues for moderate reductions using a cost-benefit argument. He argues that there are more pressing global ethical issues, such as banishing poverty, that could be addressed with the money that might be used to resolve global warming (summary: "Developed, first-world countries could do more good addressing other problems with the money saved by not addressing global warming.").

Although the authors of these articles do not argue directly against one another's position, they are clearly at odds about the ethical issues related to global warming and thus these articles promote exactly the type of cognitive conflict we are interested in exploring.

2.3 Procedure

Fig. 1 depicts the experimental procedure. The data was collected on November 19 and December 3, 2010. In preparation for the experimental sessions, students read the two source texts. The task environment consisted of Google Documents (https://docs.google.com/) that contained instructions, input fields to answer essay questions, and a chat tool.

The comparison group worked collaboratively and in a self-organized manner on both days (unscripted collaboration). On November 19 students were asked to paraphrase the Brown (Q1) and Lomborg (Q2) arguments, and to decide jointly which argument was more compelling (Q3). They were allowed (and encouraged) to consult the two source texts. On December 3 students were asked to argue for and justify the text they considered to be more compelling, without access to the source texts. Instead, they received their answers from the November 19 session. We expected livelier discussion when students use their own interpretations rather than skimming through the source texts again.

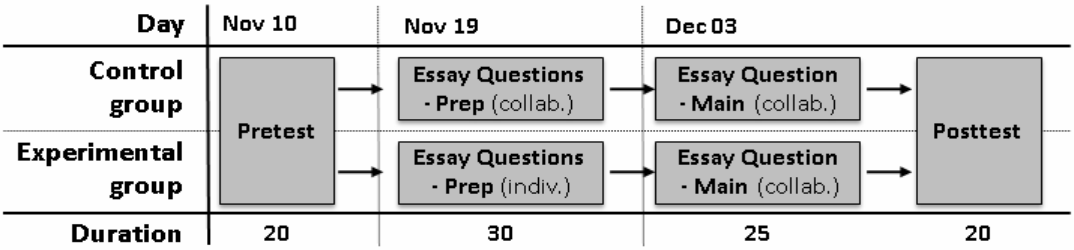

Fig. 1: Experimental procedure

The treatment group differed from the comparison group in several respects. On November 19 they worked individually (**E1**). To increase the chances of creating different preferences we used two slightly different versions of the essay questions, one emphasizing the Brown perspective (Q1: reproduce Brown's argument; Q2: rebut Lomborg's argument) and the other emphasizing the Lomborg perspective (Q1: reproduce Lomborg's argument; Q2: rebut Brown's argument). Analogous to the comparison group (yet individually), students decided on the argument they preferred (Q3). On December 3 students who preferred Lomborg were paired up with ones who preferred Brown (**E2**). Collaboration was scripted in this session through a set of instructions that included prompts for all aspects of **E3**. The task itself was identical to that of the comparison group (i.e., select the more compelling argument and justifying this decision).

2.4 Analysis Approach

In this paper we report on the results of analyzing student argumentation and collaboration during the intervention, as it relates to **H**. Our unit of analysis is the dyad rather than the individual student; thus we avoid the problem of statistical dependencies between collaborating students, a problem that can lead to an alpha-error inflation [6].

To determine the general level of student engagement we analyzed the quantity of participation (total # of contributions and words per chat) and the heterogeneity of participation (percentage deviation from a 50/50 distribution of words between the students of each dyad.

To obtain a more detailed picture of students' argumentation, we used some of the elements of the Rainbow framework [4] to code chat protocols, and a code-and-count approach to aggregate the coded data. Rainbow has been validated and used to analyze

interactive knowledge elaboration in CSCL environments (both argumentative and not) and distinguishes seven categories of collaboration. We employed the two Rainbow categories focused on collaborative argumentation ("Argumentation" and "Broaden & Deepen") as well as a third category of our own design ("Other Elaboration"):

1. *Argumentation* moves are used to increase / decrease the believability of a thesis (e.g., supporting the Brown position: "I think brown addresses pretty tangible issues, such as legal responsibility and limited resources").
2. *Broaden & Deepen* moves are used to argue and elaborate on arguments. For instance, students might rebut an argument, discuss concepts central to an argument, or interrelate arguments (e.g., juxtaposing arguments from Brown and Lomborg: "i feel like lomborg relies more on the moral obligations of the us for future generations but brown points out the tangible factors of today (legal and limited resources)").
3. *Other Elaboration* moves are used to elaborate content, yet not as part of or in reference to an argument (e.g., jointly recollecting what the two texts were about: "Brown did talk about proportional contributions, right?").

We used ANOVAs to determine whether differences between groups are significant and Cohen's d to determine effect sizes. In order to fairly compare the comparison and treatment group interactions, we analyzed and compared the collaboration of the comparison group that took place on both November 19 and December 3 with the collaboration of the treatment group that took place only on December 3. Recall that the treatment group did not collaborate on November 19 (i.e., they worked individually that day, see Fig. 1), but the comparison group did. We did not want to penalize the comparison group by comparing the interactions that occurred only on December 3, since one could argue that relevant discussion/collaboration in the comparison group already took place on November 19 and thus may not have reoccurred again on December 3, resulting in an advantage for the treatment group.

We also analyzed the chat traces qualitatively to determine how students achieved consensus. In [24] five social modes of argumentative knowledge construction have been distinguished, three of which referring to modes of consensus building. *Quick consensus building* describes a behavior of uncritically accepting the contributions of others without further discussions. This behavior is driven by the goal to continue the discourse rather than by deep convictions. *Integration-oriented consensus building* involves the mutual transformation of positions, based on reasoned argument, in order to arrive at a joint position. *Conflict-oriented consensus building* involves critically reviewing, challenging and defending claims and arguments. Our goal is to promote discussions in which integration- and conflict-oriented forms of consensus building dominate and quick consensus building is minimized.

3. Results

Regarding quantity of participation, treatment group dyads used significantly more words, with a large effect size, $F(1, 17)=4.96$, $p=0.04$, $d=1.03$. Regarding heterogeneity of participation, there was no significant difference between groups $F(1, 17)=0.00$, $p=1.00$, $d=0.00$; the conditions were balanced, with the more active student of each dyad producing (on average) approximately 57% of all words.

Table 1 summarizes the results with respect to the three codes discussed above. Note, first of all, that the treatment group dyads produced more than 5 times as many instances of "Broaden & Deepen" (4.27 vs. 0.75 messages), a significant and large effect. On the other hand, notice that approximately the same amount of "Other Elaboration" and

"Argumentation" took place in the two groups. Yet, the treatment group required less than half the time for the same amount of this argumentative and elaborative activity. Although not shown in Table 1, we also compared the chat activity across groups that occurred solely on December 3 (a weaker control, as discussed earlier) and found a small effect (that was not significant) in which the treatment group produced more "Other Elaboration" and "Argumentation."

Table 1: Evaluation of conditions based on aggregated Rainbow codes

Code	Comparison		Treatment		Difference			
	M	SD	M	SD	Diff	F	p	D
Argumentation	2.13	2.17	2.09	1.45	-0.04	0.00	1.00	0.00
Broaden & Deepen	0.75	1.04	4.27	3.74	3.52	6.62	0.02*	1.20
Other Elaboration	1.88	2.64	2.09	2.59	0.21	0.03	0.86	0.08

We also did a more qualitative analysis of the chats, which showed a pattern of quick consensus building [24] by the comparison group, i.e., students seemingly agreed not because they were convinced but (more likely) to quickly complete the task. An example of this can be found in the comparison group chat segment shown in Table 2. After student 1 states that Lomborg was more convincing (lines 1 and 2), student 2 agrees (line 3), brings Brown into play (line 4) yet without following up or elaborating. Instead of discussing the different positions, the dyad quickly agrees on "Lomborg" (lines 5 - 8). A possible argument against Brown is brought forward only after the decision had already been made (line 9). The chat of Table 2 illustrates a general trend we observed in which the comparison group dyads made far less critical points in their chat. In fact, if anything, this dyad made more points than most of the other comparison group dyads (e.g., line 9).

Table 2: Chat segment from comparison group (November 19)

#	Stud.	Contribution
1	S1	what do you think for the last question?
2		i think lomborg was more convincing
3	S2	yeah, I think Lomborg had some good points
4		but Brown has some too
5	S1	ya so we can say somewhat convincing
6		should we say brown or Lomborg
7	S2	I think Lomborg
8	S1	ok good we agree
9	S2	because Brown doesn't actually say what will happen even if the US takes responsibility and takes the initiative

Conversely, the treatment group chats clearly showed more discussion of the contrasting arguments. An excerpt from a representative treatment group chat protocol is shown in Table 3. Here, student 1 brings forward an argument in favor of Lomborg (line 1), which is questioned by student 2 (line 2). In response, student 1 points to a possible misunderstanding of student 2 and clarifies his point (lines 3 and 4). Student 2 concedes his mistake (lines 5 and 7), while student 1 provides further clarification on the Lomborg position (line 6). Despite the fact that treatment group students entered the collaborative phase with opposite conclusions, most dyads demonstrated critical objectivity in discussing. For instance, one student stated "I was for Lomborg, but I can argue Browns." In another instance, a student indicated that the decision could be driven by pragmatic considerations rather than deep convictions: "I feel like it would be easier to argue (for Lomborg's practical viewpoint)."

Table 3: Chat segment from treatment group (December 3)

#	Stud.	Contribution
1	S1	ok, so basically I think Lomborg's argument is better b/c his solution also covers what you were saying about equity (R & D should improve that...
2	S2	How would putting a tax on CO2 make more funds available for R & D? Or are you saying the tax would be an incentive for the US and other rich countries to do R&D?
3	S1	He didn't support putting tax...
4		He was saying that if we impose tax, that would decrease CO2 but there would be another cost
5	S2	Ah
6	S1	I don't know where he was gonna get the money from but he kind of just said we can use resources that we use for reducing CO2 for something else...
7	S2	I guess I misread the article. It seemed to me that Lomborg wanted a tax in addition to other methods to help

One limitation of our finding is that we originally conducted a day of computer-based argument mapping of the Brown text on November 12, but a technical problem occurred for some in the comparison group and thus this session was dropped from the experiment. Therefore, the treatment group theoretically had more exposure to the Brown argument text on this day. This confounding factor should be considered minor, however, since both conditions were asked to read the texts in advance.

4. Conclusion and Outlook

In summary, both from the perspective of the quantitative and qualitative analyses, the structured intervention appeared to successfully promote collaborative argumentation. The treatment group used significantly more words, engaged in significantly more broadening and deepening of the discussion, and appeared (in a qualitative sense) to engage in more critical and objective argumentation than the comparison group. Thus, hypothesis **H** was generally confirmed.

Our study results show that structured collaboration (i.e., scripts) can promote argumentative content elaboration and critical discussion norms. On the other hand, despite overall significant effects, the treatment group dyads sometimes appeared to collaborate in a suboptimal way, preferring a least-effort solution. In our next study we will put more emphasis on promoting collaborative elaboration. We also plan to provide dynamic feedback, using AI techniques, to challenge students and provoke a more critical discourse. For instance, there is some empirical evidence that shows a “devil’s advocate” approach can stimulate students to reason more critically [2].

Acknowledgements

The German Research Foundation (DFG) provided support for this research. We would also like to thank Donald Brown and Maria Daskolia for their helpful suggestions and comments on our global climate ethics materials. Finally, we would like to thank Dimitra Tsovaltzi for helping with the data analysis.

References

[1] Andriessen, J. (2006). Arguing to Learn. In: K. Sawyer (Ed), *Handbook of the Learning Sciences* (pp. 443–459). Cambridge University Press.
[2] Asterhan, C.S.C., Schwarz, B.B. (2010). Online Moderation of Synchronous e-Argumentation. *IJCSCL 5*(3), 259–282.
[3] Baker, M. (2003). Computer-mediated argumentative interactions for the co-elaboration of scientific notions. In J. Andriessen, M. Baker, & D. D. Suthers (Eds), *Arguing to learn: confronting cognitions in computer-supported collaborative learning environments* (pp. 1–25). Kluwer, Dordrecht.
[4] Baker, M., Andriessen, J., Lund, K., van Amelsvoort, M., Quignard, M. (2007). Rainbow: A Framework for Analyzing Computer-Mediated Pedagogical Debates. *IJCSCL 2*(2–3), 247–272.
[5] Brown, D. A. (2002). U.S. Obligation to Act Even if the Developing World Does Not. In: *American Heat: Ethical Problems with the United States' Response to Global Warming*. Rowman & Littlefield, New York.
[6] Cress, U. (2008). The need for considering multilevel analysis in CSCL research - An appeal for the use of more advanced statistical methods. *IJCSCL 3*(1), 69–84.
[7] Doise, W., Mugny, W. (1984). *The Social Development of the Intellect*. Oxford: Pergamon.
[8] Harrell, M. (2005). Using Argument Diagramming Software in the Classroom. *Teaching Philosophy 28*(2), 163–177.
[9] Jermann, P., Dillenbourg, P. (2003). Elaborating new Arguments through a CSCL Script. In: J. Andriessen, M. Baker, & D. D. Suthers (Eds), *Arguing to Learn: Confronting Cognitions in Computer-Supported Collaborative Learning Environments* (pp. 205–226). Kluwer, Dordrecht.
[10] Kuhn, D. (1991). *The Skills of Argument*. New York: Cambridge University Press.
[11] Lomborg, B. (2007). Conclusion: Making Our Top Priorities Cool. In: *Cool It: The Skeptical Environmentalist's Guide to Global Warming*. Alfred A. Knopf, New York.
[12] McLaren, B.M., Scheuer, O., Mikšátko, J. (2010). Supporting Collaborative Learning and e-Discussions using Artificial Intelligence Techniques. *IJAIED 20*(1), 1–46.
[13] Nussbaum, E. M. (2008). Collaborative Discourse, Argumentation, and Learning: Preface and Literature Review. *Contemporary Educational Psychology 33*(3), 345–359.
[14] Pinkwart, N., Ashley, K. D., Lynch, C., Aleven, V. (2009). Evaluating an Intelligent Tutoring System for Making Legal Arguments with Hypotheticals. *IJAIED 19*(4), 401–424.
[15] Postmes, T., Spears, R., Cihangir, S. (2001). Quality of Decision Making and Group Norms. *Journal of Personality and Social Psychology 80*(6), 918–930.
[16] Rummel, N. Spada, H. (2005). Learning to Collaborate: An Instructional Approach to Promoting Collaborative Problem Solving in Computer-Mediated Settings. *JLS 14*(2), 201–241.
[17] Scheuer, O., Loll, F., Pinkwart, N., McLaren, B. M. (2010). Computer-Supported Argumentation: A Review of the State of the Art. *IJCSCL 5*(1), 43–102.
[18] Scheuer, O.; McLaren, B. M.; Loll, F. & Pinkwart, N. (to appear 2011). Automated Analysis and Feedback Techniques to Support and Teach Argumentation: A Survey. In N. Pinkwart, & B. M. McLaren (Eds), *Educational Technologies for Teaching Argumentation Skills*. Bentham Science Publishers.
[19] Schwarz, B. B., Glassner, A. (2007). The Role of Floor Control and of Ontology in Argumentative Activities with Discussion-Based Tools. *IJCSCL 2*(4), 449–478.
[20] Strijbos, J. W., Martens, R. L., Jochems, W. M. G., Broers, N. J. (2004). The Effect of Functional Roles on Group Efficiency: Using Multilevel Modeling and Content Analysis to Investigate Computer-Supported Collaboration in Small Groups. *Small Group Research 35*(2), 195–229.
[21] Suthers, D. D., Connelly, J., Lesgold, A., Paolucci, M., Toth, E. E., Toth, J., et al. (2001). Representational and Advisory Guidance for Students Learning Scientific inquiry. In K. D. Forbus, & P. J. Feltovich (Eds), *Smart Machines in Education: The Coming Revolution in Educational Technology* (pp. 7–35). AAAI/MIT, Menlo Park.
[22] Van Eemeren, F. H., Grootendorst, R. (2004). *A Systematic Theory of Argumentation: The Pragma-Dialectical Approach*. Cambridge University Press, Cambridge.
[23] Weinberger, A., Ertl, B., Fischer, F., Mandl, H. (2005). Epistemic and social scripts in computer–supported collaborative learning. *Instructional Science 33*(1), 1–30.
[24] Weinberger, A., Fischer, F. (2006). A Framework to Analyze Argumentative Knowledge Construction in Computer-Supported Collaborative Learning. *Computers & Education 46*(1), 71–95.
[25] Weinberger, A., Stegmann, K., Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction 17*(4), 416–426.
[26] Weinberger, A., Stegmann, K., Fischer, F. (2010). Learning to Argue Online: Scripted Groups Surpass Individuals (Unscripted Groups do not). *Computers in Human Behavior 26*(4), 506–515.