# Will Structuring the Collaboration of Students Improve Their Argumentation?

Oliver Scheuer[1], Bruce M. McLaren[1, 2], Maralee Harrell[2], Armin Weinberger[1]

[1] Saarland University, Saarbrücken, Germany
[2] Carnegie Mellon University, Pittsburgh, PA U.S.A.
{o.scheuer@mx.uni-saarland.de, bmclaren@cs.cmu.edu, mharrell@andrew.cmu.edu, a.weinberger@mx.uni-saarland.de}

**Abstract.** Learning to argue in a computer-mediated and structured fashion is investigated in this research. A study was conducted to compare dyads that were scripted in their computer-mediated collaboration with dyads that were not scripted. A process analysis of the chats of the dyads showed that the scripted experimental 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 non-scripted control group.

**Keywords:** computer-supported collaborative learning, argumentation

## 1 Introduction

Researchers have been increasingly more interested in studying how to use technology to help students learn argumentation skills [1]. For instance, computer-based argument mapping tools have been developed and used to teach argument analysis skills in philosophy classes [2]. 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 argumentation process [3, 4]. Finally, the AIED community has investigated how to analyze arguments and how to provide feedback to support learning or argumentation in domains such as science [5], the law [6], and ethics [7].

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 take a reasoned position. Our main research question is: Will structured student collaboration lead to higher quality argumentation? This work follows from others who have investigated similar questions of the effect of scripts [4] and structuring [8] on the learning of argumentation.

## 2 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 [4]. Students often lack sufficient argumentation skills to engage in collaborative argumentation and productively resolve conflicts [9]. On a macro level, process losses due to inefficient task coordination often outweigh the advantages of combining forces [10]. On a micro level, students often avoid taking a critical stance towards peers' contributions, but typically aim for quick consensus [11]. 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 teamwork usually requires a combination of individual and collaborative activities [12]. Individual preparation gives students time to make up their own minds about a controversial issue without social pressure. Furthermore, such preparation appears to support collaborating students in engaging in fruitful social interaction.

**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 [13]. 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 [14].

**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* [15]. 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 [16]. We encourage students to take a critical position by identifying and discussing possible weaknesses in their partner's contributions. (3) *Productive collaboration*. 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:

**H1:** 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 [11].

Our second hypothesis (**H2**) is that a higher quality of collaboration will lead to more learning. This paper focuses on H1 and a process analysis to evaluate it; an investigation of H2 is deferred to future work.

## 3 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. Three sessions (with required attendance) were conducted between Nov 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 control condition; students in a late recitation class constituted the experimental condition.

### 3.1 Population

The control group consisted of 28 students (39% female); the experimental group consisted of 26 students (54% female). The experimental condition had a higher percentage of freshmen and sophomores (85%), while a majority of the control group students were juniors and seniors (71%). However, the groups were similar in terms of midterm course grades (85% of the experimental students and 75% of the control scored A or B), so we assume homogenous abilities across the groups. Since not all students attended all sessions, the final process analysis is based on 8 control dyads (16 students) and 11 experimental dyads (22 students).

### 3.2 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.

Brown [17] argues for large reductions of greenhouse gas emission levels by developed countries. He outlines moral and legal obligations for (in particular) the U.S. 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 lead to global warming.").

Lomborg [18] 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: "The U.S. 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.

### 3.3 Procedure

Fig. 1 depicts the experimental procedure. The data was collected on Nov 19th and Dec 3rd, 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 control group worked collaboratively and in a self-organized manner on both days (unscripted collaboration). On Nov 19th 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 Dec 3rd 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 Nov. 19th session. We expected livelier discussion when students use their own interpretations rather than skimming through the source texts again.

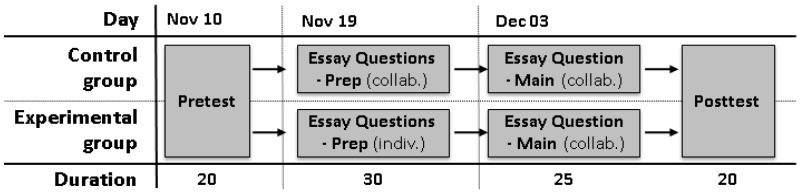


**Fig. 1:** Experimental procedure

The experimental group differed from the control group in several respects. On Nov 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 control group (yet individually), students decided on the argument they preferred (Q3). On Dec 03 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 control group (i.e., select the more compelling argument and justifying this decision).

### 3.4 Analysis Approach

In this paper we report on the results of analyzing student argumentation and collaboration during the intervention, as it relates to H1[1]. To determine the general level of student engagement we analyzed the *quantity of participation* (total # of

[1] In an initial partial grading of the pre and posttests by two independent evaluators, we did not achieve sufficient inter-rater reliability. Thus, an analysis of learning – and an investigation of our second hypothesis (H2) – is deferred for later work.

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* [19] to code chat protocols. 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: (1) "Argumentation" (statements used to increase / decrease the believability of a thesis) and (2) "Broaden & Deepen" (arguing and elaborating on arguments, e.g., rebutting an argument or discussing or concepts central to an argument), as well as a third category of our own design (3) "Text Talk" to code messages that elaborate content but not in an argumentative fashion. We used a code-and-count approach to aggregate data per discussion.

We used *ANOVA*s to determine whether differences between groups are significant and Cohen's *d* to determine effect sizes. In order to fairly compare the control and experimental group interactions, we analyzed the collaboration of the control group that took place on both Nov 19 and Dec 03 but analyzed the collaboration of the experimental group that took place only on Dec 03. Recall that the experimental group did not collaborate on Nov 19 (i.e., they worked individually that day, see Fig. 1), but the control group did. We didn't want to penalize the control group by comparing the interactions that occurred only on Dec 03, since one could argue that relevant discussion/collaboration in the control group already took place on Nov 19 and thus may not have reoccurred again on Dec 03, resulting in an advantage for the experimental group. Thus, we compared the Dec 03 experimental protocols with the combined control protocols of the Nov 19 and Dec 03 sessions. Note that this is a much stricter control than comparing only the Dec 03 sessions across groups, since experimental students thus had less than half as much time to collaborate.

## 4 Results and Discussion

Regarding quantity of participation, experimental students 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:** Comparison of conditions based on aggregated Rainbow codes

| Code | Control | | Experimental | | Comparison | | | |
|---|---|---|---|---|---|---|---|---|
| | M | SD | M | SD | Diff | F | p | d |
| **Text Talk** | 1.88 | 2.64 | 2.09 | 2.59 | 0.21 | 0.03 | 0.86 | 0.08 |
| **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 |

Table 1 summarizes the results with respect to the three codes discussed above. Note, first of all, that the experimental 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 “Text Talk” and “Argumentation” took place in the two groups. Yet, the experimental group required less than half the time for the same amount of this elaborative activity. Although not shown in Table 1, we also compared the chat activity across groups that occurred solely on Dec 03 (a weaker control, as discussed earlier) and found a small effect (that was not significant) in which the experimental group produced more “Text Talk” and “Argumentation.”

We also did a more qualitative analysis of the chats, which showed a pattern of quick consensus building [11] by the control group, i.e., students seemingly agreed not because they were convinced but (more likely) to quickly complete the task. An example of this is shown in the control 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 is very representative of what we observed of the control group; if anything, this dyad made more points than most of the control dyads (e.g., line 9)

**Table 2:** Chat segment from control group (Nov 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 experimental chats clearly showed more discussion of the contrasting arguments. An excerpt from a representative experimental 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 experimental 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 who favored Brown 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).”

In summary, both from the perspective of the quantitative and qualitative analyses, the structured intervention appeared to successfully promote collaborative argumentation. The experimental 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 control group. Thus, hypothesis H1 was generally confirmed.

**Table 3:** Chat segment from experimental group (Dec 03)

| # | 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 esle... |
| 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 Nov 12, but a technical problem occurred for some in the control group and thus this session was dropped from the experiment. Thus, the experimental 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.

## 5 Conclusion

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 experimental 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 [20].

**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.

## References


1. Scheuer, O., Loll, F., Pinkwart, N., McLaren, B. M.: Computer-Supported Argumentation: A Review of the State of the Art. ijCSCL 5(1), 43–102 (2010)
2. Harrell, M.: Using Argument Diagramming Software in the Classroom. Teaching Philosophy 28, 163-177 (2005)
3. Andriessen, J.: Arguing to Learn. In: Sawyer, K. (ed.) Handbook of the Learning Sciences, pp. 443–459, Cambridge University Press (2006)
4. Weinberger, A., Stegmann, K., Fischer, F.: Learning to Argue Online: Scripted Groups Surpass Individuals (Unscripted Groups do not). Comp. in Hum. Beh., 26, 506-515 (2010)
5. Suthers, D. D., Connelly, J., Lesgold, A., Paolucci, M., Toth, E. E., Toth, J., et al.: Representational and Advisory Guidance for Students Learning Scientific inquiry. In: Forbus, K. D., Feltovich, P. J. (eds.) Smart Machines in Education: The Coming Revolution in Educational Technology, pp. 7–35, AAAI/MIT, Menlo Park (2001)
6. Pinkwart, N., Ashley, K. D., Lynch, C., Aleven, V.: Evaluating an Intelligent Tutoring System for Making Legal Arguments with Hypotheticals. International Journal of Artificial Intelligence in Education 19(4), 401–424 (2009)
7. McLaren, B.M., Scheuer, O., Mikšátko, J.: Supporting Collaborative Learning and e-Discussions using Artificial Intelligence Techniques. International Journal of Artificial Intelligence in Education (IJAIED) 20(1), 1–46 (2010)
8. Schwarz, B. B., Glassner, A.: The Role of Floor Control and of Ontology in Argumentative Activities with Discussion-Based Tools. International Journal of Computer-Supported Collaborative Learning (ijCSCL), 2(4), 449–478 (2007)
9. Kuhn, D.: The Skills of Argument. New York: Cambridge University Press (1991)
10. Strijbos, J. W., Martens, R. L., Jochems, W. M. G., Broers, N. J.: 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, 195-229 (2004)
11. Weinberger, A., Fischer, F.: A Framework to Analyze Argumentative Knowledge Construction in Computer-Supported Collaborative Learning. Computers & Education 46(1), 71–95 (2006)
12. Rummel, N. Spada, H.: Learning to Collaborate: An Instructional Approach to Promoting Collaborative Problem Solving in Computer-Mediated Settings. Journal of the Learning Sciences 14(2), 201–241 (2005)
13. Doise, W., Mugny, W. The Social Development of the Intellect. Oxford: Pergamon (2010)
14. Nussbaum, E. M.: Collaborative Discourse, Argumentation, and Learning: Preface and Literature Review. Contemporary Educational Psychology 33(3), 345–359 (2008)
15. Van Eemeren, F. H., Grootendorst, R.: A Systematic Theory of Argumentation: The Pragma-Dialectical Approach, Cambridge University Press, Cambridge (2004)
16. Postmes, T., Spears, R., Cihangir, S.: Quality of Decision Making and Group Norms. Journal of Personality and Social Psychology 80, 918–930 (2001)
17. Brown, D. A.: 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 (2002)
18. Lomborg, B.: Conclusion: Making Our Top Priorities Cool. In: Cool It: The Sceptical Environmentalist's Guide to Global Warming. Alfred A. Knopf, New York (2007)
19. Baker, M., Andriessen, J., Lund, K., van Amelsvoort, M., Quignard, M.: Rainbow: A Framework for Analyzing Computer-Mediated Pedagogical Debates. ijCSCL 2(2–3), 247–272 (2007)
20. Asterhan, C.S.C., Schwarz, B.B.: Online Moderation of Synchronous e-Argumentation. ijCSCL 5(3), 259–282 (2010)