Distance Education,
Vol. 27, No. 3, November 2006, pp. 371–390

Effects of Pre-structuring Discussion Threads on Group Interaction and Group Performance in Computer-supported Collaborative Argumentation

C. Darren Brooks* and Allan Jeong
Florida State University, USA

This study examined the effects of pre-structuring discussion threads on group performance in computer-supported collaborative argumentation where students labeled their messages as arguments, challenges, supporting evidence, and explanations on a threaded discussion board. In the pre-structured group students were required to post supporting and opposing arguments to designated and separate discussion threads. In the control group no discussion threads were designated for posting supporting and opposing arguments. The mean number of challenges elicited per argument was 64% greater with pre-structured threads (*ES* = +.47). No significant differences were found in the mean number of counter-challenges, supporting evidence, and explanations posted in reply to challenges. The findings suggest that prestructured discussion threads can increase the frequency of argument→challenge exchanges needed to initiate critical discourse.

Introduction

Computer-mediated communication (CMC) and online threaded discussion boards have been widely used to support learner–learner interaction and facilitate collaborative argumentation to promote critical thinking (Collins & Collins, 1996; Jeong & Joung, in press; Pilkington & Walker, 2003; Ravits, 1997; Ward & Tiessen, 1997). Collaborative argumentation is an inquiry-based learning strategy used to increase critical thinking skills in online environments (Derry, Levin, Osana, Jones, & Peterson,

*Corresponding author. Educational Psychology and Learning Systems, Instructional Systems Program, Stone Building 305E, Florida State University, Tallahassee, FL 32306-4453, USA. Email: dbrooks615@comcast.net

ISSN 0158-7919 (print); 1475-0198 (online)/06/030371–20

DOI 10.1080/01587910600940448

2000). The nature of the cognitive tasks required in argumentation makes it a form of reasoning essential to formulating ideas and beliefs (Cho & Jonassen, 2002). By using CMC to support collaborative argumentation, students practice argumentation skills using online text-based communication tools to gain deeper subject understanding and create new meaning by testing personal knowledge and uncertainties against those of their peers (Baker, 1999; McAlister, 2001; Reiser & Dempsey, 2002). With CMC, online discussions can be conducted asynchronously with discussion threads (Hewitt, 2003)—hierarchically organized sequences of alternating messages and responses—enabling students to post contributions to multiple and concurrent conversational threads without being constrained by time and the processes of turn-taking often used to manage face-to-face discussions. As a result, students have more time to carefully evaluate and respond to other students' contributions to produce more in-depth discussions.

Despite the affordances of using asynchronous threaded discussions, studies still find that students rarely respond to one another's points, often repeat points already made by other students (Koschmann, 2003; Veerman, 2003), and often produce discussions that lack coherence and depth (Herring, 1999). Ironically, the absence of turn-taking may be what is contributing to these behaviors because the function of turn-taking is to: (a) allow students to express a point with the group's undivided attention; (b) manage the turn-taking transitions so that discussions stay on topic; (c) manage the transitions so that changes in discussion topics follow a logical flow in ways that help the group achieve its objectives. With no turn-taking procedures in place, repeated points in a discussion can split a discussion into separate threads. Each thread then competes for the attention of group members and reduces the limited pool of cognitive resources that can be used to advance each discussion. Without turn-taking, students are also more apt to create sequences of topical threads that do not follow a coherent line of inquiry. Further contributing to this problem is that students tend to post replies to a message regardless of the message's relevance to the main point of discussion (Hewitt, 2003), and will sometimes post "orphaned" replies to messages by erroneously posting a reply in a new thread (Wiley, 2004). These behaviors altogether inhibit students' ability to advance a discussion.

To address these problematic behaviors, computer-supported collaborative argumentation (CSCA) systems have been developed to facilitate critical discourse. In many of these systems constraints or collaborative scripts are imposed by the system to control the types of messages and responses (e.g. claims, challenges, supporting evidence, and explanations) students can post to a discussion in order to keep groups on task (Cho & Jonassen, 2002; Jeong & Joung, in press; McAlister, 2001; Veerman, Andriessen, & Kanselaar, 1999; Weinberger, Fischer, & Mandl, 2004). When using this approach messages are tagged with a label to allow students to clearly identify the function of each message to help students logically connect and track postings in a threaded discussion. In addition, CSCA systems can also allow instructors to impose constraints on message–response sequences (Jonassen & Remidez, 2002) by limiting the types of messages students can post in response to particular types of messages (e.g. argument→critique→counter-critique or argument→critique→evidence). These methods have been found to increase the generation of coherent arguments and the number of problem-solving actions during collaborative discussions (Cho & Jonassen, 2002). When constraints were placed on message–response sequences, Jeong (2003) found no significant increase in the number of challenges posted in reply to stated arguments, but found significant increases in the frequency of elaborative messages posted in reply to challenges. As a result, these findings suggest that these methods can promote higher levels of critical discourse.

However, the methods described above are used to manage messages exchanged among students within discussion threads, not to manage discussion threads within discussion forums. Given the problematic behaviors identified above, there is good reason to believe that how one manages discussion threads within a discussion forum will affect how students exchange messages within discussion threads and affect the quality and depth of an online discussion. Tools for managing discussion threads can be found in course management systems like Blackboard™ and WebCT™, where instructors (but not students) can create separate discussion forums to host different topics of discussion. However, only Blackboard™ enables the instructor to structure a discussion forum with pre-configured discussion threads that can be designated to discuss sub-topics within the forum. In Blackboard™ the instructor also has the option to prevent students from creating new discussion threads within a forum. When this option is used, the instructor must create the discussion threads in advance to enable students to post responses to the preconfigured threads and to the forum.

When these tools for managing discussion threads are used in combination with CSCA (i.e. message constraints and message labels) a discussion forum in Blackboard™ can be structured (see Figure 1) to host, for example, a team debate by: (a) pre-configuring the forum with one discussion thread on which to post all arguments in support of a position and another but separate discussion thread on which to post all arguments in opposition to a position; (b) require students to post arguments by replying directly to the top message in the thread and tagging the argument with a number to identify its location and order relative to all previously posted arguments; (c) removing the students' ability to create new discussion threads. These three procedures can be used together to create "pre-structured threads" to address some of the problematic behaviors noted above.

As a result, using "pre-structured threads" in CSCA presents a number of potential advantages: (a) each designated thread (with clearly stated discussion topics presented within the message title) helps to remind students of the main points of discussion and can, therefore, help to reduce the number of responses posted in reply to "interesting messages" that are not relevant to the main discussion; (b) all opposing and supporting arguments are presented separately and in consecutive order (not interspersed among one another in a series of disjointed threads) and, thus, previously posted arguments can be quickly located to help students avoid posting repeated arguments and "cross-arguing," i.e. when a counter-argument is posted in reply to an opposing argument without directly addressing the merits of the opposing argument (Jeong, 2003); (c) disallowing students from creating new threads can help reduce the frequency of "orphaned responses," help reduce instances where discussions are split into separate and disjointed threads, and thus help the group in harnessing their limited cognitive resources to thoroughly develop, analyze, and/or contest each presented argument.

Although the rationale for using pre-structured discussion threads is supported in some ways by findings from research on face-to-face collaborative learning that showed that well-designed instructional settings support relevant discourse circulation and interaction patterns that facilitate collaborative work (Johnson & Johnson, 1994), no reported studies at this time have tested the effects of pre-structured threads using the procedures described above. To fully understand the impact of pre-structured threads on group performance in CSCA the quality and depth of argumentation must be measured in terms of the sequential nature of strategic moves (e.g. arguments, challenges, supporting evidence, and explanations) used to advance and contest arguments (Jeong, 2005a). For example, studies are needed to examine how many times arguments are challenged (argument→challenge), challenges are counter-challenged (challenge→counter-challenge), how many times challenges are countered with supporting evidence to verify arguments (challenge→evidence), and how many times challenges are countered with explanations to justify arguments (challenge→explain). These types of questions must be addressed in order to understand the strategic value of using pre-structured discussion threads to improve group performance in CSCA, and ultimately in computer-supported collaborative work, decision-making, and problem-solving.

Theoretical Assumptions

The underlying assumptions used to ground the research questions and methods used in this study were based on assumptions of the dialogic theory of language (Bakhtin, 1981). The main assumption in dialogic theory is that social meaning is renegotiated and reconstructed as a result of conflict brought about through social interactions. Accordingly, conflict is the primary catalyst that drives critical inquiry and discourse. The second assumption is that conflict is produced by the juxtaposition of interlocking pairs of statements, not by an individual statement alone. In other words, new meaning occurs through the exchange of opposing ideas within a social context. Some current research in CMC supports these assumptions that "conflict and consideration of both sides of an issue" (Baker, 1999; Jeong, 2003, 2005a; Johnson & Johnson, 1992; Wiley & Voss, 1999) produce critical inquiry and deeper understanding.

These assumptions imply that a student's performance in CSCA should not be analyzed in terms of messages examined in isolation from the responses elicited by the messages. Examining message frequencies alone does not measure the extent to which presented arguments are challenged and the extent to which challenges trigger critical reflection. As a result, the theoretical framework used in this study demands a process-oriented and micro-analytical approach that examines the sequencing of messages and responses exchanged between participants in terms of their functional roles (or speech acts) during collaborative argumentation and examines how particular sequences model the processes of critical inquiry. The types of exchanges of most interest in this study were those that produced conflict because conflict is assumed to be the primary catalyst that drives critical inquiry and discourse.

Purpose of Study

The purpose of the study was to compare the effects of using versus not using prestructured discussion threads on how often students responded to messages in ways that support higher levels of critical discourse. Based on the assumptions of dialogic theory (Bakhtin, 1981), this study addressed four main questions regarding specific response patterns that were determined to be key indicators of critical discourse:

1. Challenging arguments. Do prestructured threads increase the number of challenges elicited per stated argument?
2. Counter-challenging. Do prestructured threads increase the number of counter-challenges elicited per challenge?
3. Explanations. Do prestructured threads increase the number of explanations elicited per challenge?
4. Supporting evidence. Do prestructured threads increase the number of responses presenting supporting evidence elicited per challenge?

Method

Participants

The participants in this study were 30 graduate students from a major south-eastern university in the USA, with ages ranging between 25 and 50 years. The pre-structured group comprised 16 students, 10 female and 6 male, enrolled in the fall 2004 term. The control group comprised 14 students, 10 female and 4 male, enrolled in the fall 2003 term. Each group was registered on a 16 weeks online graduate course titled "Introduction to distance education." The same course requirements and activities were administered to both groups.

Debate Procedures

Prior to each debate students were assigned to one of two teams to either support or oppose a given position. The teams were balanced by gender and by level of individual participation based on the number of messages posted in previous group discussions earlier in the course. Students in both groups conducted the debates using threaded discussion forums in the Blackboard™ course management system. Students were required to post at least four messages to each debate in order to receive participation points. Class participation in all the weekly group discussions over the course of the semester could contribute as much as 25% to of the course grade. The purpose of each debate was to critically examine topical issues, concepts, and principles covered in the course. For example, students debated the following claims: "The Dick and Carey ISD model is the best model for designing distance learning activities"; "The role of the instructor should change when teaching at a distance"; "Media makes very little or no significant contributions to learning"; "Print is the preferred medium for delivering a course study guide."

Students in the control group participated in six debates during the fall 2003 term, while students in the pre-structured group participated in a total of five debates during the fall 2004 term. Only the first four debates from each group were examined in this study, with all four debates performed on the same topics and in the same sequence. In the pre-structured group students were required to post supporting arguments and opposing arguments within a designated discussion thread within each debate forum (see Figure 1). To post a new argument each student was required to open the message prompt located at the top of the thread (which was titled "Opposing arguments" or "Supporting arguments"), post a reply to the prompt with the stated argument, and enter into the subject heading of the reply the label "ARG" and argument number (based on the number of arguments already posted to the thread).

In the control group students were not provided with pre-configured discussion threads on which to post supporting and opposing arguments (see Figure 2). Instead, students in the control group clicked on the "Add new thread" button to create a new discussion thread on which to post each (supporting or opposing) argument. Students in the control group were not instructed nor required to sequentially cluster or number all the arguments presented by members of their own teams. As a result, opposing and supporting arguments were often presented in random and alternating sequences.

Procedures Used to Label Messages

In each debate students were instructed to review and refer to a given set of response categories, based loosely around Toulmin's (1958) model of argumentation, when posting messages to the debates (see Figure 3). Each student was required to classify each posted message by category by inserting the corresponding label into the subject heading of each posted message. Students were also required to identify team membership by adding a "+" or "–" (or "s" and "o" for the control group) to each message label (e.g. +ARG, –ARG). The message labels were then followed by a message title.

In the control group postings were restricted to six message categories (argument, elaboration, issue, evidence, evaluate, and suggest) to facilitate critical discourse in the fall 2003 term. In the pre-structured conditions students used the label BUT (instead of ISSUE) to tag messages that were challenges, and used EXPL (instead of ELAB) to tag messages that provided additional support, elaboration, and/or explanations. These superficial changes in message labels were believed to exert very little or no confounding effects in this study. At the same time the message categories "evaluate" and "suggest" (used in the control group, but not in the pre-structured group)

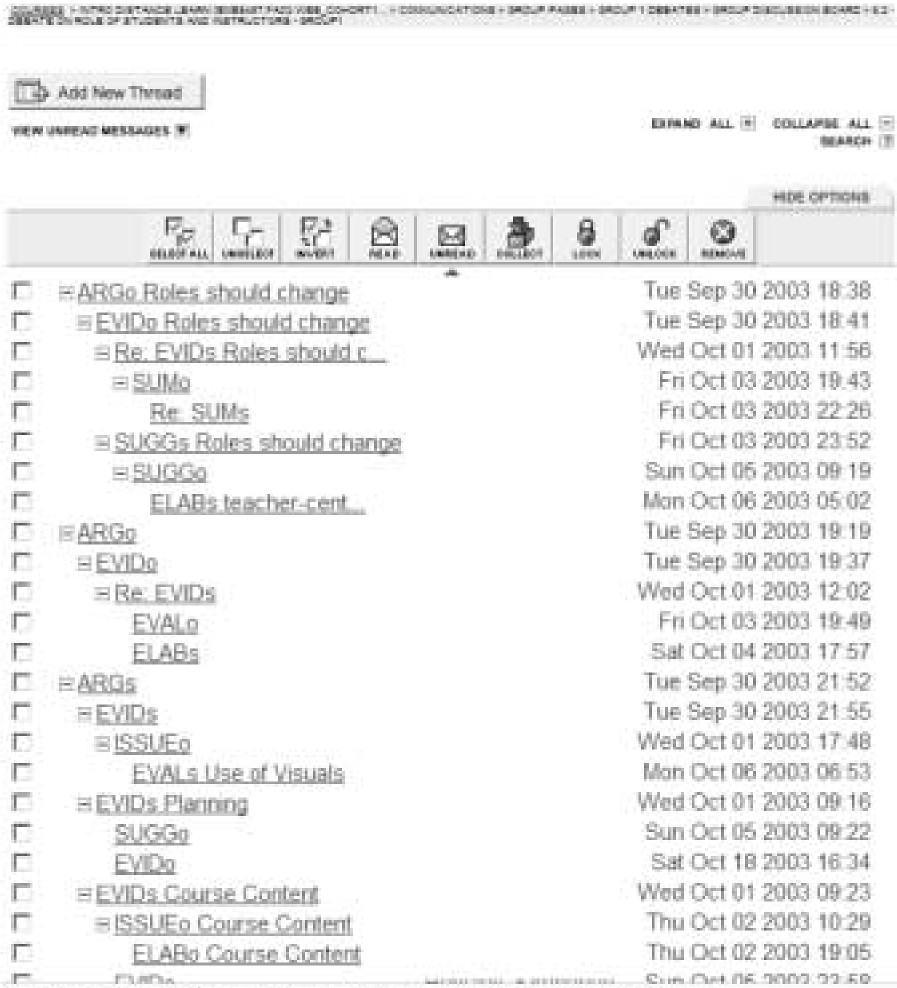
Note: To ensure subject confidentiality student names were blocked out.

The discussion thread for posting all supporting arguments is located in the first message in the forum. The thread for posting all opposing arguments is positioned below the thread for supporting arguments, but is out of view in the above illustration.

Figure 1. Excerpt from a pre-structured discussion forum with designated threads, numbered arguments, and without access to the button for adding new threads

were reassigned and subsumed under the "explain" and "but" categories, respectively, because of the course instructor's desire to simplify the debate procedures used in the fall 2004 term.

In order to compare performance between the groups the 11 EVAL responses observed in the control group were recoded as EXPL messages, given the similarities between the two categories. Three SUGG messages posted in reply to students on the opposing team were recoded as BUT and two SUGG messages posted in reply to students on the same team were recoded as EXPL. Given that the number of EVAL and SUGG messages were relatively small, reassigning these messages to



Note: To ensure subject confidentiality student names were blocked out.

Figure 2. Excerpt from a discussion forum without pre-structured threads, without numbered arguments, and with access to the button for adding new threads

EXPL and BUT (which were also similar in function) was not believed to adversely affect the accuracy nor the validity of the findings reported in this study.

Data Set

A computer program called ForumManager (Jeong, 2005b) was used to download the student-labeled messages from the Blackboard™ discussion forums into Microsoft Excel™ in its original hierarchical format in order to preserve information used to determine which responses were posted in reply to which messages. Another computer program, the Discussion Analysis Tool or DAT (Jeong, 2005a,

Symbol	Description of symbol
+ [1]	Identifies a message posted by a student assigned to the team supporting the debate claim
- [1]	Identifies a message posted by a student assigned to the team opposing the debate claim
ARG#	Identifies a message that presents one and only one argument or reason (supporting or opposing) for the debate claim. Number each posted argument by counting the number of arguments already presented by your team. Sub-arguments need not be numbered BC = because . Example posting: -ARG *2 ProducesDeeperDiscussions*
BUT	Identifies a reply/message that questions or challenges the merits, logic, relevancy, validity, accuracy, or plausibility of a presented argument (ARG) or challenge (BUT). Example posting: *+BUT ProducesCognitiveOverload*
EXPL	Identifies a reply/message that provides additional support, explanation, clarification, elaboration for an argument, or challenge. Example posting: *-EXPL CanParticipateInMultipleThreads*
EVID	Identifies a reply/message that provides proof or evidence to establish the validity of an argument or challenge. Example posting: *-EVID Threads50%LongerOnAverage*
EVAL [2]	Identifies a reply/message that evaluates, reviews, compares or otherwise weighs presented arguments in favor of an argument
SUGG [2]	Identifies a reply/message that makes suggestions, requests or asks questions to strengthen a posting

[1] In the control group, students used s and o to tag messages posted by the supporting and opposing teams, respectively (instead of + and -).

[2] Due to the course instructor s desire to simplify the message categories, the EVAL and SUGG categories were reassigned and subsumed under the EXPL and BUT categories. As a result, the 11 EVAL responses observed in the control group were recoded as EXPL given their similarities in function. Three SUGG messages posted in reply to students on the *opposing* team were recoded as BUT, and two SUGG messages posted in reply to students on the same team were recoded as EXPL.

Figure 3. Students' instructions on how to label messages during the online debates

2005c), was used to automatically extract the codes (or labels) from the subject headings of each message to tag each message as an argument (ARG), evidence (EVID), challenge (BUT), or explanation (EXPL). To check for possible errors in the labels that were assigned by students to each message, the experimenter coded all the messages from the second and third debates in both groups. The experimenter's codes were then compared with the students' codes. Cohen's κ test of reliability of the coding for the control group showed perfect agreement (k = 1.0) for both the second and third debates. The pre-structured group showed perfect agreement (k = 1.0) for the second debate and very good agreement (k =.973) for the third debate. In this study the experimenter's codes (not the students' codes) were used for the data analysis.

The frequency table in Figure 4 shows that in the control group a total of 230 messages were observed, with 46 ARG, 91 BUT, 57 EXPL, and 36 EVID. The table also shows the proportion of messages (for each message category) that received at least one or more replies: 31% of arguments, 36% of challenges, 16% of explanations, and 17% of presented evidence received one or more replies. The frequency table in Figure 5 shows that in the pre-structured group a total of 323 messages were observed, with 85 ARG, 165 BUT, 34 EXPL, and 39 EVID. Figure 5 also shows the proportion of messages (for each message category) that received at least one or more replies: 53% of arguments, 32% of challenges, 6% of explanations, and 8% of presented evidence received one or more replies.

Data Analysis

A 2 × 4 univariate analysis of variance was conducted to examine the effects of using versus not using pre-structured threads across each of the four types of exchanges on the mean response scores (e.g. the mean number of challenges posted in reply to each stated argument) using an experiment-wise rate of error of p <.05. Independent sample t-tests were also used to conduct post-hoc tests to locate where significant differences existed in the mean response scores within each type of exchange. In other words, post-hoc t-tests were conducted (each at p <.05) to compare the mean number of challenges elicited per argument, the mean number of counter-challenges elicited per challenge, the mean number of elaborative messages elicited per challenge, and the mean number of responses presenting supporting evidence elicited per challenge between the two groups. The DAT software was used to tally, for example, the number of challenges (messages labeled BUT) posted in reply to each argument (ARG) within each group to produce the raw scores needed to conduct the tests described above (see details in Jeong, 2005a).

Results

Effects of Using Prestructured Threads

No significant differences were found in the mean number of responses per message posted across all four types of exchanges (argument→challenge, challenge→counter-challenge, challenge→explanation, challenge→evidence) [F(1, 879) =.161, p >.05]. Table 1 shows that the mean number of responses per message was 0.26 (SD = 0.58, N = 580) in the pre-structured group and 0.25 (SD = 0.543, N = 307) (ES = +0.010). However, a significant interaction was found between groups and exchange type [F(3, 879) =.000, p <.05]. This suggests that the effects of using pre-structured threads depended on the type of exchange. To determine which of the four types of exchanges were most affected by the use of pre-structured threads, post-hoc t-tests were conducted on the mean number of responses observed within each exchange type between groups.

Challenges posted in reply to each argument. A significant difference was found in the mean number of challenges posted in reply to each argument between the groups [t(153) = 2.34, p =.020]. The mean number of challenges per argument in the prestructured group was 0.96 (SD = 0.89, N = 85) compared with 0.59 (SD =

Table 1. Statistics on mean number of replies posted to each discussion thread for each posted argument

Group	Exchange	Mean	STD	*N*
Structure	ARG-BUT	0.965[a]	0.892	85
	BUT-BUT	0.339	0.568	165
	BUT-EXPL	0.073	0.283	165
	BUT-EVID	0.055	0.253	165
	Total	0.259	0.578	580
Control	ARG-BUT	0.587	0.717	46
	BUT-BUT	0.333	0.623	87
	BUT-EXPL	0.161	0.428	87
	BUT-EVID	0.092	0.328	87
	Total	0.254	0.543	307
Total	ARG-BUT	0.832	0.852	131
	BUT-BUT	0.337	0.586	252
	BUT-EXPL	0.103	0.342	252
	BUT-EVID	0.032	0.197	252
	Total	0.257	0.566	887

For example, the mean number of BUT messages posted in reply to ARG in the pre-structured group was 0.965.

[a]The mean number of responses per message in the argument→challenge exchanges in the pre-structured group was significantly higher than in the control group [$t(129) = 2.47$, $p = .015$, ES = +0.47].

0.717, $N = 46$) in the control group. As a result, pre-structured threads increased the frequency of challenges per argument by 64%, with a moderate and positive effect size of +0.47.

Counter-challenges posted in reply to each challenge. No significant difference was found in the mean number of counter-challenges posted in reply to each challenge between the groups [post-hoc $t(250) = 0.08$, $p = .94$]. The mean number of counter-challenges per challenge in the pre-structured group was 0.34 (SD = 0.57, $N = 165$) and 0.33 (SD = 0.62, $N = 165$) in the control group, with a trivial effect size of +0.01. As a result, pre-structured threads did not significantly increase (2%) the number of counter-challenges per challenge.

Elaborations posted in reply to each challenge. No significant difference was found in the mean number of elaborative messages posted in reply to each challenge between the groups [post-hoc $t(250) = -1.96$, $p = .051$]. The mean number of explanations per challenge in the pre-structured group was 0.07 (SD = 0.28, $N = 165$) compared with 0.16 (SD = 0.43, $N = 87$) in the control group, with a moderate and negative effect size of –0.40. Although the observed difference did approach statistical significance ($p < .05$), pre-structured threads had no significant effect on the mean number of explanations posted per challenge.

Supporting evidence posted in reply to each challenge. No significant difference was found in the mean number of messages with supporting evidence posted in reply to each challenge between the groups [post-hoc $t(250) = -1.00$, $p = .032$]. The mean number of messages with supporting evidence per challenge in the pre-structured group was 0.05 (SD = 0.25, $N = 165$) compared with 0.09 (SD = 0.33, $N = 87$) in the control group, with a very small effect size of –0.13. As a result, pre-structured threads had no effect on the number of supporting evidence posted per challenge.

Exploring the Effects of Response Patterns on Differences in Mean Response Scores

Exploratory analysis was conducted to identify potential differences in response patterns between the two groups, response patterns that might help to explain the observed differences in the mean number of challenges posted per argument between the groups. Analyzing response patterns can help to determine, for example, whether or not the observed differences in the mean number of challenges per argument can be attributed to students' inclinations to respond more often to arguments with supporting evidence instead of responding to arguments with challenges.

To identify patterns in students' responses, sequential analysis (Bakeman & Gottman, 1997) was conducted using the DAT software to: (a) count the frequency of specific responses to each type of message; (b) convert each response frequency into a transitional probability to determine, for example, what percentage of responses to arguments were challenges versus supporting evidence; (c) test the observed transitional probabilities with Z-scores to determine which probabilities were significantly higher than the expected probabilities to determine dominant patterns in students' responses; (d) translate the transitional probabilities into state diagrams to produce a visual means of summarizing and identifying patterns in message–response sequences. In Figures 4 and 5 are the observed frequencies, transitional probabilities, and Z-scores for all possible message–response exchanges for the control and pre-structured groups, respectively. These results are summarized in the transitional state diagrams shown in Figure 6, with the transitional probabilities that occurred at higher than expected probabilities presented as darker arrows. The transitional diagrams are visual representations of the interaction patterns observed in each condition, illustrating the types and frequencies of responses that were elicited by each type of message.

To identify response patterns that might explain the observed differences in the mean number of challenges posted per argument the response distributions in the first rows of the frequency matrix in Figures 4 and 5 were tested against one another using a 2×4 χ^2 test of independence (two groups × four possible types of responses to arguments). No significant differences were found in the distribution of responses posted in reply to arguments based on the results of a χ^2 test of independence [χ^2 (3, $N = 183$) = 4.50, $p = .212$]. Using the same approach, no significant differences were found in the distribution of responses posted in reply to challenges [χ^2 (3, $N = 142$) = 7.47, $p = .058$], although the observed differences did approach statistical

Frequency matrix

	ARG	BUT	EXPL	EVID	Replies	No Replies	Givens	% replies	% givens
ARG	1	27	10	18	56	18	46	.31	.20
BUT	2	35	19	9	65	41	91	.36	.40
EXPL	0	15	12	1	28	38	57	.16	.25
EVID	0	14	9	8	31	16	36	.17	.16
	3	91	50	36	180	113	230		

Transitional probabilities matrix

	ARG	BUT	EXPL	EVID	Replies	No Replies	Givens	Reply Rate
ARG	.02	.48	.18	.32	56	18	46	.61
BUT	.03	.54	.29	.14	65	41	91	.55
EXPL	.00	.54	.43	.04	28	38	57	.33
EVID	.00	.45	.29	.26	31	16	36	.56
	3	91	50	36	180	113	230	.13

Z-score matrix

	ARG	BUT	EXPL	EVID
ARG	0.08	-0.42	-2.00	2.74
BUT	1.11	0.66	0.33	-1.55
EXPL	-0.75	0.35	1.94	-2.37
EVID	-0.80	-0.66	0.17	0.89

Figure 4. Frequency, transitional probability, and *Z*-score matrices produced by DAT for message–response sequences observed in the control group

Frequency matrix

	ARG	BUT	EXPL	EVID	Replies	No Replies	Givens	% replies	% givens
ARG	1	82	17	27	127	17	85	.53	.26
BUT	0	56	12	9	77	104	165	.32	.51
EXPL	0	13	2	0	15	20	34	.06	.11
EVID	0	14	3	3	20	26	39	.08	.12
	1	165	34	39	239	167	323		

Transitional probabilities matrix

	ARG	BUT	EXPL	EVID	Replies	No Replies	Givens	Reply Rate
ARG	.01	.65	.13	.21	127	17	85	.80
BUT	.00	.73	.16	.12	77	104	165	.37
EXPL	.00	.87	.13	.00	15	20	34	.41
EVID	.00	.70	.15	.15	20	26	39	.33
	1	165	34	39	239	167	323	.12

Z-score matrix

	ARG	BUT	EXPL	EVID
ARG	0.94	-1.59	-0.40	2.20
BUT	-0.69	0.85	0.41	-1.34
EXPL	-0.26	1.53	-0.10	-1.77
EVID	-0.30	0.10	0.10	-0.17

Figure 5. Frequency, transitional probability, and *Z*-score matrices produced by DAT for message–responses sequences observed in the group with pre-structured threads

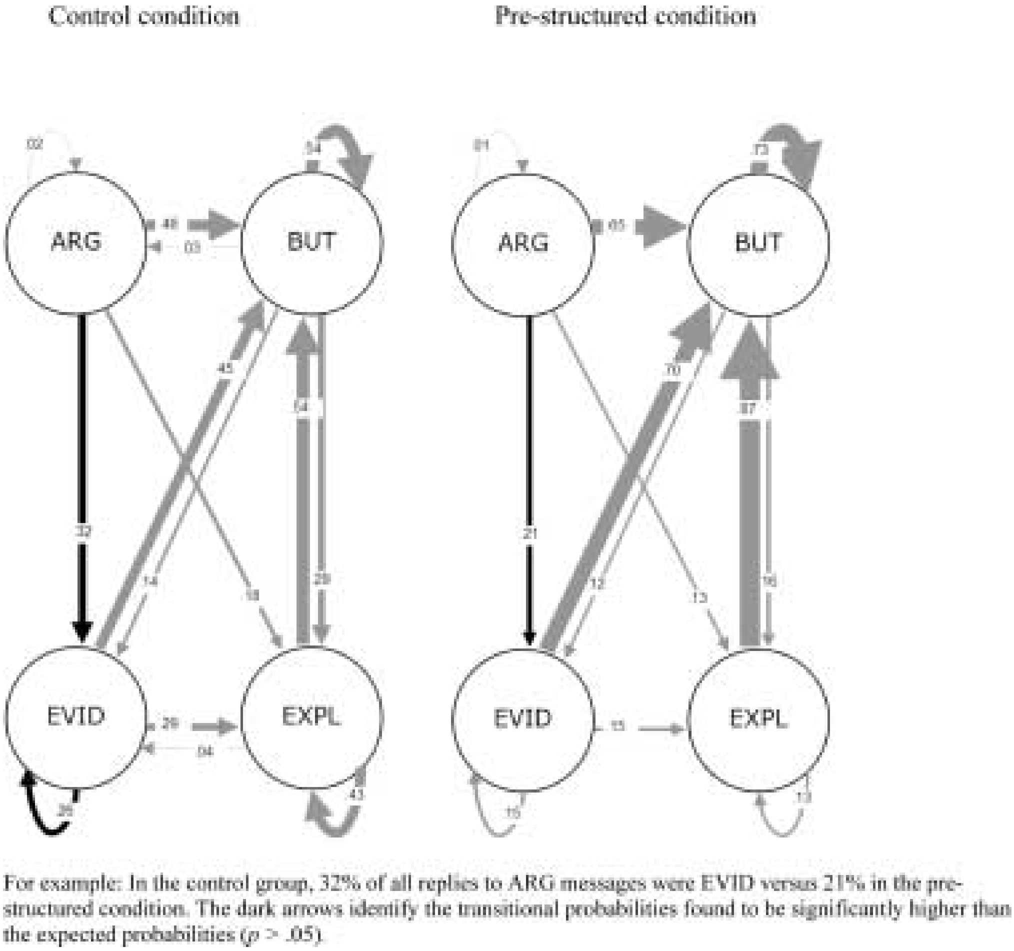


Figure 6. State diagrams of message–response sequences observed in the control condition and pre-structured condition produced by the DAT software

significance. Overall, the results of these tests show that pre-structured threads did not produce any significant differences in the way students were inclined to respond to arguments and challenges.

However, a closer analysis revealed that the percentage of responses to arguments that were supporting evidence was significantly higher than expected in both groups. Although this pattern of responses to arguments was observed in both groups (32 versus 21% in the control and pre-structured groups, respectively), students in the pre-structured group were 11% less likely than students in the control group to respond to arguments with supporting evidence. As a result, students in the pre-structured group appeared to have shifted their attention away from responding to arguments with supporting evidence, and in turn shifted more attention to responding to arguments with challenges. This response pattern helps to explain why students in the pre-structured group posted a higher percentage of challenges in response to arguments (65%) than students in the control group (48%).

Discussion

The purpose of this study was to examine how pre-structured discussion threads in CSCA affected the way students exchanged responses in ways that foster and demonstrate critical discourse while participating in online team debates. Requiring students to post supporting and opposing arguments only to designated and separated discussion threads was found to have no effect on the number of total responses posted per message across all four types of exchanges examined in this study (argument→challenge, challenge→counter-challenge, challenge→elaborate, and challenge→evidence). However, this study found that the effects of pre-structured threads depended on the type of exchange in which the responses were posted. Students who used pre-structured threads posted significantly more challenges per argument (with moderate effect size) than students who did not use pre-structured threads, while no significant differences were found in the number of counter-challenges posted per challenge, explanations posted per challenge, and supporting evidence posted per challenge.

As predicted, this study found that more challenges were posted per argument with pre-structured threads. A possible explanation for this finding is based in theory on the assumption that pre-structured threads reduce the frequency of repeated arguments, responses to irrelevant messages, and cross-arguing (as described earlier). Another explanation for this finding was that the students who used pre-structured threads were more inclined than the students who did not use pre-structured threads to respond to arguments with challenges and in turn were less inclined to respond to arguments with supporting evidence. However, finding no significant differences in the frequency of responses to challenges was not expected, because one of the underlying assumptions in this study was that more argument→challenge exchanges would produce more conflict and that the conflict would trigger more rebuttals to challenges.

One factor that might explain why no differences were found in the number of responses to challenges is because the use of pre-structured threads appears to have encouraged students to post twice the number of arguments than the control group. In other words, the numbering of arguments (or simply the clustering of supporting and opposing arguments) may have allowed students to easily count and compare the number of arguments posted by each team and induced the teams to compete with one another based on the sheer number of arguments presented. As a result, students in the pre-structured group may have channeled more effort to posting arguments (and posting challenges to opposing arguments), leaving less time and effort for posting responses to challenges. The response patterns found in the exploratory analysis (channeling attention to posting challenges instead of supporting evidence in response to arguments) provides some preliminary evidence to support this explanation. Another factor that might have contributed to this unexpected finding is the possibility that the argument→challenge exchanges did not produce sufficient conflict to trigger further inquiry compared with the level of conflict produced by challenge→counter-challenge exchanges. Jeong (2003) found preliminary evidence to show that challenge→–counter-challenge exchanges triggered longer strings of subsequent responses than argument→challenge exchanges alone.

Instructional Implications

Overall, the findings in this study, although not conclusive, support the use of pre-structured threads to facilitate critical discourse in CSCA. The findings in this study show that pre-structuring discussion threads can significantly increase the number of times students challenge each presented argument. Increasing the number of argument→challenge exchanges is the first and most important step to generating the cognitive conflict needed to trigger further inquiry.

The findings in this study suggest that in order to increase the number of responses to challenges and move a discussion forward extra caution should be taken when numbering or clustering arguments under designated threads because its use could lead students to generate a larger number of arguments. A large number of arguments can adversely affect the amount of time and effort that is available for thoroughly responding to each and every message that challenges an argument and hinder the processes of advancing the discussion and the processes of verifying (argument–challenge–evidence) and justifying (argument–challenge–elaborate) arguments. As a result, placing constraints on the total number of arguments that can be presented by each team and/or increasing the minimum number of required postings (in order to increase the amount of time and effort needed to generate deeper threads) are potential solutions to helping students produce more response challenges in order to trigger further inquiry.

Given these findings and given that the procedures described in this study can be implemented for the most part on any threaded discussion board, pre-structuring threads may prove to be an effective and feasible method of facilitating critical discourse and critical thinking in online group discussions.

Implications for Future Research

Once again, the findings in this study are not conclusive, given several limitations in its design and scope. Future studies will need to: re-examine the effects of pre-structured threads by placing some type of control over the number of arguments examined in each debate and the minimum number of required postings per student; examine a larger sample and multiple groups within each condition to ensure that idiosyncratic events (e.g. personal flaming, an unusual mix of students with personality traits that are not conducive to argumentation) that occur within a group do not inadvertently bias the main findings; conduct a close examination of the message content to test the theory that pre-structured threads reduce the frequency of repeated arguments and the frequency of responses to messages that are not relevant to the main points of discussion (work in progress); examine the effects across a wider range of exchanges (e.g. challenge→concede, challenge→–derogatory remark) that mark both constructive and non-constructive interactions; conduct further testing using identical message categories across all experimental conditions; use smaller debate groups so that the participants in both experimental conditions belong to the same course at any given time; compare students' perceptions and reactions to the online debates to gain further insights and evidence to support the reported advantages and disadvantages of using pre-structured threads.

In conclusion, this study has provided a preliminary look at the effects of pre-structuring discussion threads on group interaction patterns and group performance in CSCA. The methods and tools described in this study will hopefully serve as a framework for future investigations into the socio-cognitive processes that are supported and inhibited in CSCA. The types of message–response sequences observed in this study and the methods used to measure differences or changes in message–response sequences provides a common metric for future researchers and instructional designers to develop, refine, and assess this and similar strategies for managing group communication and maximizing group performance in computer-mediated environments.

Notes on Contributors

C. Darren Brooks is a doctoral student of instructional systems in the Department of Educational Psychology and Learning Systems at Florida State University.

Allan Jeong is a Professor of Instructional Systems in the Department of Educational Psychology and Learning Systems at Florida State University.

References

Bakeman, R., & Gottman, J. (1997). *Observing interaction: An introduction to sequential analysis.* Cambridge, UK: Cambridge University Press.

Baker, M. (1999). Argumentation and constructive interaction. In P. Courier & J. E. B. Andriessen (Eds.), *Foundations of argumentative text processing* (pp. 179–2002). Amsterdam, The Netherlands: University of Amsterdam Press.

Bakhtin, M. (1981). Discourse in the novel. In M. Holquist (Ed.), *The dialogical imagination: Four essays by M. Bakhtin* (pp. 259–422). Austin, TX: University of Texas Press.

Cho, K., & Jonassen, D. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research & Development, 50*(3), 5–22.

Collins, C., & Collins, S. (1996). The Internet as a tool, call of the north. *Proceedings of the 17th Annual National Educational Computing Conference.* Minneapolis, MN, 97–101. (ERIC Document Reproduction Service No. ED398883).

Derry, S. J., Levin, J. R., Osana, H. P., Jones, M. S., & Peterson, M. (2000). Section on teaching, learning, and human development—fostering students' statistical and scientific

thinking: Lessons learned from innovative college courses. *American Educational Research Journal, 37*(3), 747–775.
Herring, S. (1999). Interactional coherence in CMC. *Journal of Computer Mediated Communication, 4*(4). Retrieved August 26, 2005, from http://jcmc.indiana.edu/vol4/issue4/herring.html.
Hewitt, J. (2003). *Toward an understanding of how threads die in asynchronous computer conferences.* Paper presented at the 2003 Annual Meeting of the American Educational Research Association, Chicago, IL.
Jeong, A. (2003). The sequential analysis of group interaction and critical thinking in online threaded discussions. *The American Journal of Distance Education, 17*(1), 25–43.
Jeong, A. (2005a). A guide to analyzing message–response sequences and group interaction patterns in computer-mediated communication. *Distance Education, 26*(3), 367–383.
Jeong, A. (2005b). *Forum manager.* Retrieved March 21, 2006, from http://bbproject.tripod.com/ForumManager.
Jeong, A. (2005c). *Discussion analysis tool.* Retrieved May 10, 2005, from http://garnet.fsu.edu/~ajeong/DAT.
Jeong, A., & Joung, S. (In press). The effects of response constraints and message labels on interaction patterns and argumentation in online discussions. *Computers and Education.*
Johnson, D., & Johnson, R. (1992). *Creative controversy: Intellectual challenge in the classroom.* Edina, MN: Interaction Book Co.
Johnson, D. W., & Johnson, R. T. (1994). *Cooperative learning in the classroom.* Alexandria, VA: Association for Supervision and Curriculum Development.
Jonassen, D., & Remidez (2002). *Mapping alternative discourse structures onto computer conference.* Paper presented at the 2002 Computer Support for Collaborative Learning Conference: Foundations for a CSCL Community, Boulder, CO. Retrieved August 13, 2005, from http://newmedia.colorado.edu/cscl/12.html.
Koschmann, T. (2003). CSCL, argumentation, and Deweyan inquiry: Argumentation is learning. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 261–269). Boston, MA: Kluwer Academic.
McAlister, S. R. (2001). *Argumentation and a design for learning,* CALRG Report no. 197. Retrieved August 13, 2005, from http://iet.open.ac.uk/pp/s.r.mcalister/personal/197.pdf.
Pilkington, R., & Walker, A. (2003). Using CMC to develop argumentation skills in children with a literacy deficit. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 144–175). Boston, MA: Kluwer Academic.
Ravits, J. (1997). An ISD model for building online communities: Furthering the dialogue. In O. Abel, N. Maushak, & K. Wright (Eds.) *Proceedings of the 1997 National Convention of the Association for Educational Communications and Technology,* 297–307.
Reiser, R. A., & Dempsey, J. V. (2002). *Trends and issues in instructional design and technology.* Upper Saddle River, NJ: Pearson Education.
Toulmin, S. (1958). *The uses of argument.* Cambridge, UK: Cambridge University Press.
Veerman, A. L. (2003). Constructive discussion through electronic dialogue. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 117–143). Boston, MA: Kluwer Academic.
Veerman, A. L., Andriessen, J. E. B., & Kansellaar, G. (1999). *Collaborative learning through computer-mediated argumentation.* Paper presented at the 1999 Conference on Computer Supported Collaborative Learning, San Francisco, CA.
Ward, S. C., & Tiessen, E. L. (1997). Adding educational value to the Web: Active learning with Alivepages. *Educational Technology, 37*(5), 22–30.
Weinberger, A., Fischer, F., & Mandl, H. (2004). *Knowledge convergence in computer-mediated learning environments: Effects of collaboration scripts.* Paper presented at the American Educational Research Association Conference. San Diego, CA.

Wiley, D. (2004). *Measuring interactivity in online group discussions.* Paper presented at the 2004 American Educational Research Association Conference. Chicago, IL.

Wiley, J., & Voss, J. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91,* 301–311.