Distance Education
Vol. 26, No. 3, November 2005, pp. 367–383

A Guide to Analyzing Message–Response Sequences and Group Interaction Patterns in Computer-mediated Communication

Allan Jeong*
Florida State University, USA

This paper proposes a set of methods and a framework for evaluating, modeling, and predicting group interactions in computer-mediated communication. The method of sequential analysis is described along with specific software tools and techniques to facilitate the analysis of message–response sequences. In addition, the Dialogic Theory and its assumptions are presented to establish a theoretical framework and guide to using sequential analysis in computer-mediated communication research. Step-by-step instructions are presented to illustrate how sequential analysis can be used to measure the way latent variables (e.g., message function, response latency, communication style) and exogenous variables (e.g., gender, discourse rules, context) affect how likely a message is to elicit a response, the types of responses elicited by the message, and whether or not the elicited sequence of responses (e.g., claim → challenge → explain) mirror the processes that support group decision-making, problem-solving, and learning.

Introduction

Current research in computer-mediated communication (CMC) is in need of alternative theories, methods, and software tools to achieve a deeper and more thorough understanding of CMC and its effects on group interaction, group performance, and learning (Garrison, 2000; Koschmann, 1999; Mandl & Renkl, 1992). One approach is to examine group processes by studying the sequential nature of messages and responses exchanged between students to determine how particular processes, and the variables that affect the processes, help or inhibit groups from achieving the desired outcomes (Jeong, 2003a; Koschmann, 1999). As a result, a process-oriented

*Florida State University, Stone Building 305E, Tallahassee, FL 32317, USA. Email: jeong@coe.fsu.edu

ISSN 0158-7919 (print); 1475-0198 (online)/05/030367–17

DOI 10.1080/01587910500291470

approach to studying CMC enables researchers to develop computational models to explain and predict patterns in group interaction based on specific characteristics of the message and the conditions surrounding the exchange of messages.

At this time, content analysis is one of the current methods used in CMC research. Its primary purpose is to identify message categories and measure the frequency of messages observed in each category (Rourke, Anderson, Garrison, & Archer, 2001). This approach generates results that are mainly descriptive rather than prescriptive in nature, reporting for example the frequencies of arguments, challenges, and explanations observed in a discussion. However, message frequencies provide little information to explain or predict how participants respond to given types of messages (e.g., argument → challenge versus argument → simple agreement), how response patterns are influenced by latent variables (e.g., message function, content, communication style, response latency) and exogenous variables (e.g., gender, personality traits, discussion protocols, type of task), and how particular response patterns help to improve group performance to achieve desired outcomes. Therefore, new approaches are needed to examine to what extent messages elicit responses based on what is said in conjunction with when, how, who, and why messages are presented, and whether or not the elicited responses help produce sequences of speech acts that support critical discourse (e.g., claim → challenge → explain) and group performance in decision-making, problem-solving, and learning.

What follows is a detailed description of the tools, techniques, and the seven steps to using sequential analysis to study group interaction in CMC, based on Bakeman and Gottman (1997) and the previous studies of event sequences in CMC (Jeong, 2003a, b, c, 2004a, 2005b; Jeong & Joung, in press). In general, this method has also been used in studies on interpersonal communication conducted over the past 30 years, which include studies on the conversational patterns between married couples, children at play, mother–infant play (Bakeman & Gottman, 1997, pp. 184–193; Gottman, 1979), and studies on human–computer interaction (Olson, Herbsleb, & Rueter, 1994). This method has been claimed to be the "missing factor" in research on the effects of computer-mediated environments and computer-based instruction (England, 1985; King & Roblyer, 1984). The following discussion begins with a proposed set of theoretical assumptions to establish the foundation for the proposed metrics for measuring group interaction, specific methods and software tools to support sequential analysis, and research designs for investigating the effects of latent and exogenous variables on group interaction patterns.

Theoretical Framework

The dialogic theory (Bakhtin, 1981) provides a theoretical framework for reconceptualizing and operationalizing group interaction in collaborative learning (Koschmann, 1999). In this theory, language is viewed as part of a social context in which all possible meanings of a word interact, possibly *conflict*, and affect future meanings. As a result, meaning does not reside in any one utterance (or message).

Instead, meaning emerges from examining the relationship between multiple utterances (e.g., a message and replies to the message). Through the process of examining the interrelationships and conflicts that emerge from a social exchange, meaning is renegotiated and reconstructed through extended social interaction. Conflicts that emerge from the interactions are what drive further inquiry, reflection, and articulation of individual viewpoints and underlying assumptions.

Support for this theory can be drawn from extensive research on collaborative learning showing that conflict and the consideration of both sides of an issue is needed to drive inquiry, reflection, articulation of individual viewpoints, and underlying assumptions, and to achieve deeper understanding (Johnson & Johnson, 1992; Wiley & Voss, 1999). The need to explain, justify, or understand is felt and acted upon only when conflicts or errors are brought to attention (Baker, 1999). This process not only plays a key role in increasing students' understanding, but also in improving group decision-making (Lemus, Seibold, Flanagin, & Metzger, 2004).

As a result, the two main assumptions are that conflict is produced not by ideas presented in one message alone, such as an argument or claim, but by the juxtaposition of opposing ideas presented in a message and responses to the message; and that conflicts produced in exchanges help to trigger subsequent responses that can serve to verify (e.g., argument → challenge → evidence) and justify (e.g., argument → challenge → explain) stated arguments and claims. These assumptions imply that we should be focusing on analyzing the frequency of specific message–response pairs (e.g., argument → challenge, challenge → explain) and not the frequency of messages alone (e.g., arguments, challenges, explanations).

Step 1: Choose a metric for measuring and comparing group interaction patterns

A number of possible metrics can be used to analyze and identify patterns in message–response sequences. The two metrics that are perhaps the most meaningful are transitional probabilities—which determine, for example, what percentage of the observed responses to arguments (ARG) are challenges (BUT) versus supporting evidence (EVI) versus explanations (EXPL)—and mean response scores—the mean number of specific responses elicited per message category, such as the mean number of challenges, supporting evidence, or explanations elicited per stated argument.

Transitional probabilities are computed by tallying the frequency of a particular response posted in reply to a particular message type and by reporting the results in a frequency matrix, as illustrated in Table 1. The observed frequencies are converted into relative frequencies to determine the transitional probabilities for each response type for each message category (see Table 2). To determine whether or not the transitional probabilities of each response to each message category are significantly higher or lower than expected, and to determine whether a pattern exists in the way participants respond to messages in a particular category, Z scores are computed and reported in a Z-score matrix (see Table 3). As opposed to using the independence chi-square statistic, this Z-score statistic proposed by Bakeman and Gottman (1997, pp. 108–111) takes into account not only the observed total number of responses to

Table 1. Frequency matrix of responses to messages across message categories

	ARG	BUT	EVID	EXPL	Replies	No replies	Givens	% Targets	% Givens
ARG	3	**101**	**73**	**<u>16</u>**	193	35	112	0.25	0.30
BUT	3	**<u>82</u>**	88	91	264	24	149	0.35	0.40
EVID	0	64	50	48	162	22	35	0.21	0.09
EXPL	0	51	**<u>22</u>**	**71**	144	55	74	0.19	0.20
	14	307	233	229	763	136	370		

The number of challenges posted in reply to arguments (in bold) was *higher* (n = 101) than expected. The number of challenges posted in reply to challenges (underlined) was significantly *lower* (n = 82) than expected.

a particular message category, but also the marginal totals of each response type observed across all message types.

The transitional probabilities presented in Table 2 are represented in a state diagram (shown in Figure 1) that provides a Gestalt view of the group processes and a means to visually identify response patterns and predict event sequences that are most likely to occur. For example, the diagram can be used to determine or predict

Table 2. Transitional probability matrix

	ARG	BUT	EVID	EXPL	Replies	No replies	Givens	Reply rate
ARG	0.02	**0.52**	**0.38**	**<u>0.08</u>**	193	35	112	0.69
BUT	0.01	**<u>0.31</u>**	0.33	0.34	264	24	149	0.84
EVID	0.00	0.40	0.31	0.30	162	22	35	0.37
EXPL	0.00	0.35	**<u>0.15</u>**	**0.49**	144	55	74	0.26
	14	307	233	229	763	136	370	0.52

The proportion of replies to ARG that were BUT (52%) was significantly *higher* than expected. The proportion of replies to BUT (31%) was significantly *lower* than expected.

Table 3. *Z*-score matrix

	ARG	BUT	EVID	EXPL
ARG	–0.34	**3.96**	**2.54**	**<u>–7.62</u>**
BUT	–1.05	**<u>–3.76</u>**	1.22	1.95
EVID	–1.96	–0.21	0.10	–0.12
EXPL	–1.82	–1.31	**<u>–4.41</u>**	**5.61**

Z scores < –2.32 reveal probabilities (bolded and underlined) that were significantly *lower* than expected. *Z* scores > 2.32 reveal probabilities (bolded) that were significantly *higher* than expected.

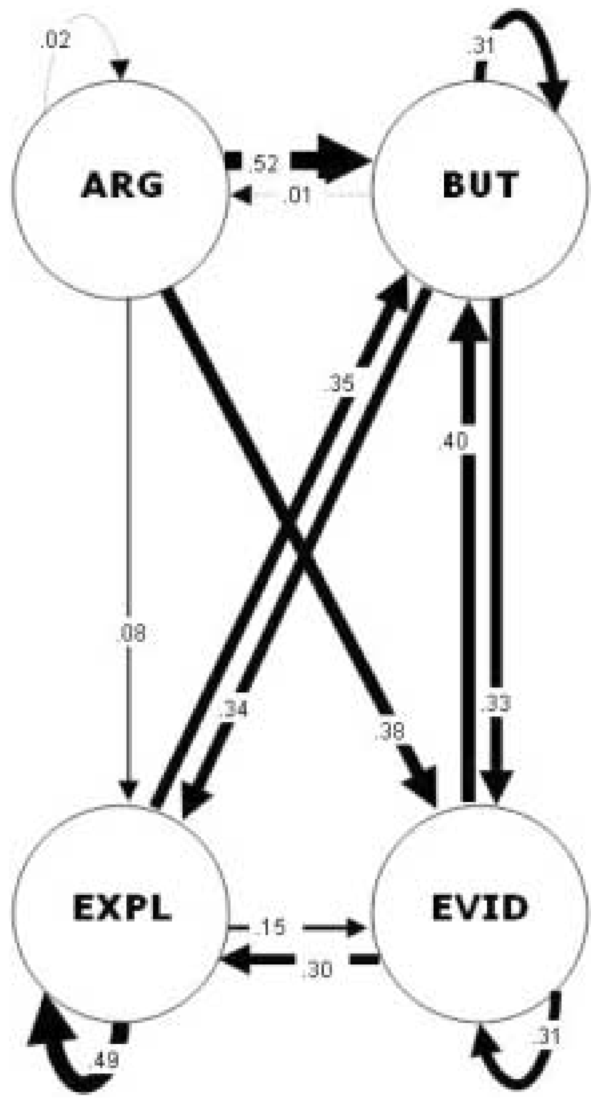

Figure 1. Transitional state diagram

how often arguments will elicit challenges versus counter-arguments, and in turn to predict how often challenges will elicit explanations versus counter-challenges to determine, overall, how likely the observed patterns of interaction will lead to constructive dialog (e.g., argument → challenge → explanation) versus non-productive dialog (e.g., argument → opposing argument).

The second metric, the mean number of specific responses elicited per message category or mean response scores, determines how many times a given type of message is able to elicit a particular type of response. This metric *describes* the overall level of performance by measuring, for example, the mean number of challenges elicited per argument and the mean number of explanations elicited per challenge, which is similar to measuring the percentage of arguments left unchallenged and the percentage of challenges left unresolved. As a result, this particular metric can be used to determine at what level participants are critically analyzing arguments (e.g., argument → challenge → explain), or to what extent participants engage in processes (e.g., argument → counter-argument, argument → no response) that block critical discourse. By using mean scores, statistical methods like *t* tests and analyses of variance can be used to test for differences in response patterns between experimental conditions, and effect sizes can be computed to determine to what extent the observed differences are meaningful differences.

Between these two metrics—transitional probabilities and mean response scores—transitional probabilities can be used to explain observed differences in mean response scores. For example, one group might exhibit a tendency to respond to arguments with more challenges than with supporting evidence, whereas another group might exhibit an opposite tendency to respond to arguments with more supporting evidence but fewer challenges. If a significant difference is found in the mean number of challenges elicited per argument between groups, the differences in interaction patterns would suggest that the second group posted fewer challenges in response to arguments because more time and resources were allocated by the group to developing evidence to support arguments leaving less time and resources to challenge arguments. As a result, both metrics can be used at the same time, with one metric used as the main dependent variable and the other used for post-hoc analysis. However, transitional probabilities are best used as the main dependent variable when conducting an exploratory study, whereas mean response scores are best for conducting experimental studies.

Step 2: Specify a priori tests for specific message–response pairs

When using either of the metrics already described, the specific message–response pairs (or event pairs) examined in a study should be defined a priori because the total number of possible event pairs grows exponentially with the addition of each message category to the coding scheme. For example, a coding scheme consisting of four categories (e.g., argument, challenge, explain, evidence) produces a 4 × 4 matrix resulting in 16 possible event pairs (e.g., argument → challenge, challenge → argument, challenge → explain, explain → challenge, etc.). Testing all 16 event pairs for differences in mean response scores would be too large a number of contrasts to adequately control for Type I error (finding significant differences when the differences are actually the result of random chance alone). Power can be increased by testing only a select number of event pairs—particularly those that are believed to support group performance (e.g., argument → challenge, challenge → explain). To identify the most important sequences to examine in your study, review existing literature and research that present specific models for completing specific tasks. The other alternative is to closely examine social exchanges while groups perform a particular task (Mandl & Renkl, 1992) and identify the subordinate skills and skill sequences needed to successfully complete the task by using the techniques for analyzing intellectual skills (Dick, Carey, & Carey, 2005 pp. 38–56).

Step 3: Collect discussions and messages parsed and classified by speech act

The next step in sequential analysis is to parse the discussion transcripts into discrete units of analysis. Each unit must be classified by function (or speech act) based on an established coding scheme using the same procedures for conducting quantitative content analysis (Rourke et al., 2001). However, the process of parsing and coding is fraught with a number of methodological challenges where the reliability, validity, and feasibility of parsing and coding messages pose significant problems. Messages often address multiple topics or functions, making the process of parsing each message into discrete segments extremely difficult to achieve with high interrater reliability. As a result, researchers have debated the merits of parsing and categorizing messages by sentence, paragraph, message, unit of meaning, and speech act. The problem with interrater reliability is then compounded when one attempts to map the links between units presented within a message with units presented within responses to the message (Gunawardena, Lowe & Anderson, 1997; Newman, Johnson, Cochrane, & Webb, 1996).

One technique for resolving this problem is to instruct participants to classify, label, and post messages to address one, and only one, function at a time (e.g., argument, evidence, challenge, explanation). See example instructions in Figure 2 for structuring online group debates. By using this approach, each message is associated with one, and only one, speech act. As a result, the process of parsing messages into discrete units of analysis and the challenges associated with this process are essentially minimized, if not eliminated. Also eliminated are the challenges associated with the process of mapping the links between speech acts observed in messages and responses to messages. The additional advantage of using this approach is that larger data sets can be more easily produced in order to generate a sufficient number of event pairs within the probability matrix to test transitional probabilities and mean response scores.

Message labeling has been implemented in a number of computer-supported collaborative argumentation systems to scaffold argumentation and problem-solving (Carr & Anderson, 2001; Cho & Jonassen, 2002; McAlister, 2003; Sloffer, Dueber, & Duffy, 1999; Veerman, Andriessen, & Kanselaar, 1999) and to enable participants to see the overall structure and organization of their arguments (see Figure 3). However, message labeling in itself can affect group interactions and the validity of the findings. At this time, the effects of message labeling have not yet been fully investigated and initial findings are still inconclusive (Beers, Boshuizen, & Kirschner, 2004; Jeong & Joung, in press; Strijbos, Martens, Jochems & Kirschner, 2004). Nevertheless, message labeling seems be a practical, although not perfect, solution to address the problems that have prevented previous researchers from examining event sequences in CMC. Regardless, the interactions and findings produced with this type of approach will be useful for improving the design and

Symbol	Description of symbol
+	Identifies a message posted by a student assigned to the team supporting the given claim/statement
-	Identifies a message posted by a student assigned to the team opposing the given claim/statement
ARG#	Identifies a message that presents one and only one argument or reason for using or not using chats (instead of threaded discussion forums). Number each posted argument by counting the number of arguments already presented by your team. Sub-arguments need not be numbered. ARG = "argument".
EXPL	Identifies a reply/message that provides additional support, explanation, clarification, elaboration of an argument or challenge.
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).
EVID	Identifies a reply/message that provides proof or evidence to establish the validity of an argument or challenge.

Figure 2. Example instructions on how to label messages during the online debates

implementation of computer-supported collaborative argumentation, where approaches like message labeling are used to structure and facilitate discourse.

Step 4: Download messages with message threads intact

Once the group discussions are completed and all the posted messages have been labeled by the participants, the message threads must be downloaded and prepared for analysis. At this time, little (if any) software is available for downloading discussions from systems in current use. Among the systems that support downloading, messages are directed into flat files where the explicit links between multi-threaded messages are not recorded and therefore do not remain intact. Even with existing qualitative content analysis tools, such as Atlas-ti™ and NUDIST™, and tools like General Sequential Querier (GSEQ™) for performing sequential analysis (Bakeman & Quera, 1995), the multi-threaded nature of discussions are difficult to retain and analyze. However, the computer program ForumManager (Jeong, 2004c) is under development and has been used in recent studies to harvest messages from Blackboard™, a course management system (see Figure 4) into Microsoft Excel™. Once in Excel™, the message headers and full texts are

SUPPORT statement because...	Instructor	Sat Oct 2, 2004 11:18 am
+ARG1 MediaIsButAMereVehicle	Student name	Mon Oct 4, 2004 8:47 pm
-EVID MediaIsButAMereVehicle	Student name	Tue Oct 5, 2004 7:09 pm
+BUT RelativityTheoryOldToo	Student name	Tue Oct 5, 2004 9:43 pm
-BUT RelativityTheoryOldToo	Student name	Sat Oct 9, 2004 10:12 am
-BUT Whataboutemotions?	Student name	Tue Oct 5, 2004 9:53 pm
+EVID DistEdEffectiveAsF2F	Student name	Tue Oct 5, 2004 10:40 pm
-BUT Mediaamerevehicle	Student name	Wed Oct 6, 2004 8:19 pm
+EVID MooreConcurs	Student name	Wed Oct 6, 2004 10:07 pm
+EXPLMediaSelectionComesAfterInstructionalStrategy	Student name	Sun Oct 10, 2004 12:35 am
-BUT WellChosenEffective	Student name	Sun Oct 10, 2004 4:31 pm
+BUT SupportingResearch	Student name	Sun Oct 10, 2004 5:37 pm
-BUT Mediaismorethanamerevehicle	Student name	Fri Oct 8, 2004 5:30 pm
+BUT SupportingEvidence?	Student name	Sat Oct 9, 2004 8:51 am
-BUT LearningNotSimplyAPassiveResponseToDeliveryMethod	Student name	Mon Oct 11, 2004 9:54 am
+ARG2 Standards for teaching	Student name	Wed Oct 6, 2004 1:48 pm
+BUT Clarification?	Student name	Sun Oct 10, 2004 5:39 pm
+ARG3 MediaUnrelatedtoLearningObjectives	Student name	Wed Oct 6, 2004 3:12 pm
-BUT MediaUnrelatedtoLearningObjectives	Student name	Wed Oct 6, 2004 8:26 pm
+BUT MediaSelection	Student name	Thu Oct 7, 2004 9:20 am
-BUT MediaSelection	Student name	Sun Oct 10, 2004 11:21 am
+EVID MethodNotMedia	Student name	Wed Oct 6, 2004 11:04 pm
-BUT MediaUnrelatedtoLearningObjectives	Student name	Sat Oct 9, 2004 10:59 am
+EXPL Media'sContribution?	Student name	Sat Oct 9, 2004 9:10 pm
-BUT Media'sContribution?	Student name	Sun Oct 10, 2004 3:42 pm

Figure 3. Example of online debate with labeled messages in a Blackboard™ forum

archived and the message threads are structurally maintained to enable the user to read and analyze message threads.

Step 5: Prepare data for analysis according to variables under investigation

To prepare the data for sequential analysis, the Discussion Analysis Tool (DAT) has been developed (Jeong, 2005a) and used to parse out the students' labels from message headers (see column 3 in Figure 5) so that the codes are recorded into column 1 in an Excel™ worksheet. Note that the message labels in Figure 5 identify the message category and the debate team that posted the message (s = supporting team, o = opposing team). Once extracted, the codes must be checked for interrater reliability against the Cohen Kappa coefficient (Rourke et al., 2001, p. 6). Next, the code sequences must be extracted and explicitly mapped using a numerical system based on the thread level of each message (see column 2 in Figure 5).

At this point, the codes in column 1 can be manipulated to examine group interaction patterns from a number of different perspectives depending on the variables under investigation. The present data in column 1 of Figure 5 produce the transitional probability matrix in Figure 6. This matrix can be used to compare performances between the two debate teams. Such a comparison might be meaningful, for example, if the members of the supporting team are all male and the members of the opposing team are all female. Comparing the transitional probabilities in the

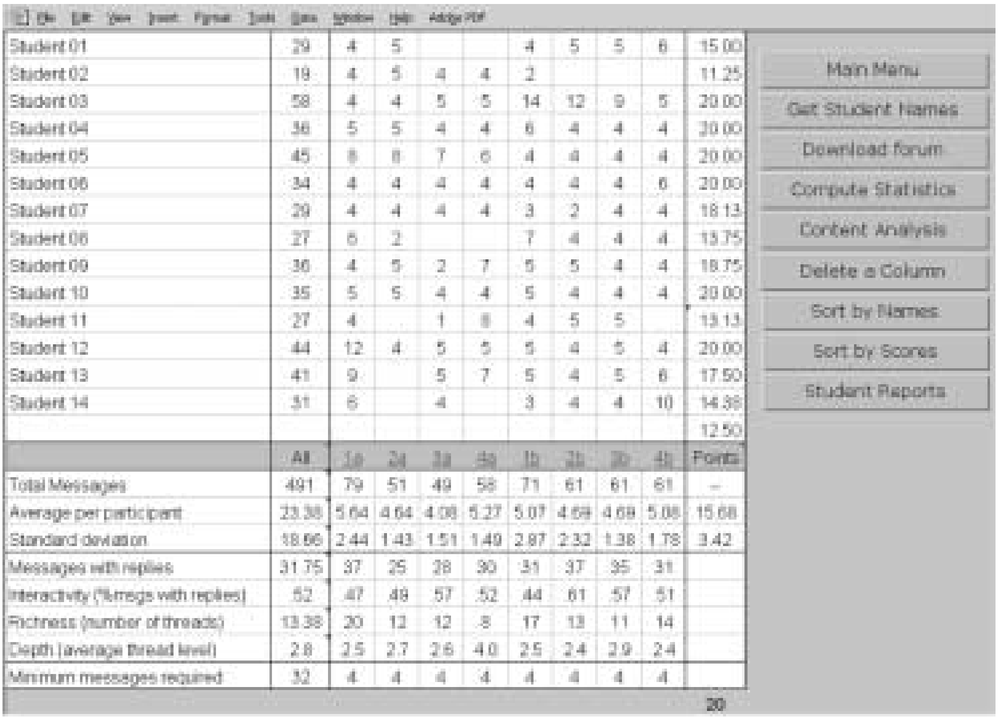

Figure 4. Screenshot of ForumManager™ for downloading discussion threads

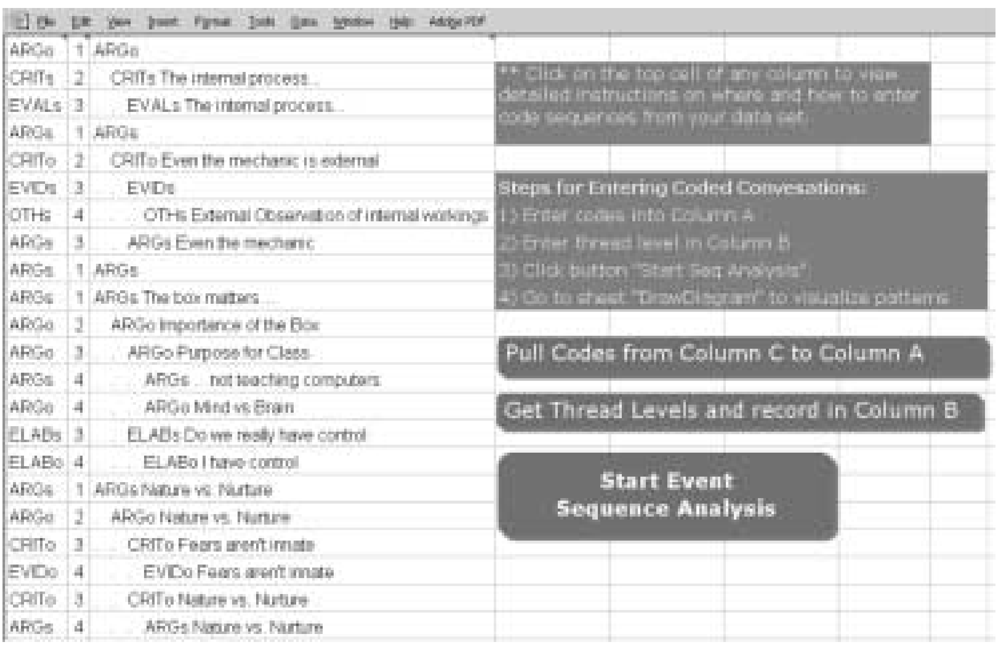

Figure 5. Screen shot of DAT for processing and analyzing message sequences

	+ARG	+BUT	+EVI	+EXP	-ARG	-BUT	-EVI	-EXP	Replies	No Replies	Givens	Reply Rate
+ARG	.01	.35	.20	.10	.00	.20	.06	.07	231	27	143	81%
+BUT	.00	.43	.11	.05	.00	.34	.03	.04	130	133	239	44%
+EVI	.00	.38	.09	.18	.00	.31	.04	.00	45	44	81	46%
+EXP	.00	.48	.10	.00	.00	.34	.00	.07	29	27	50	46%
-ARG	.02	.46	.07	.06	.00	.35	.04	.00	54	8	32	75%
-BUT	.00	.43	.04	.06	.00	.31	.05	.11	94	91	174	48%
-EVI	.00	.25	.25	.00	.00	.37	.12	.00	8	22	30	27%
-EXP	.00	.20	.05	.10	.00	.45	.05	.15	20	20	37	46%
n =	3	238	79	50	0	174	30	37	611	372	786	68%

Figure 6. Transitional probability matrix of event sequences produced by DAT

upper-right quadrant of the probability matrix would reveal how females on the opposing team responded to the males on the supporting team, and the lower-left quadrant would reveal how the males on the supporting team responded to the females on the opposing team. Each quadrant can be converted into a state diagram to visually compare, identify differences, and model interaction patterns produced between genders.

Take another example where members of both teams are mixed in gender, and the goal is to determine whether differences exist in the way males respond to males versus females and the way females respond to females versus males. To examine this question, the team tags in the present codes ("s" and "o") shown in column 1 of Figure 5 can be stripped out and replaced with the gender of the participant that posted the message (e.g., ARGm, CRITf, EVALm, etc.) by using the "find and replace" function in Excel™ to strip the tags, substituting each student name with the corresponding gender tag, and combining the code and gender tag with the "&" function. The new codes can then be sequentially analyzed by DAT to test for differences in interaction patterns between genders or any other individual traits (e.g., extroversion, cognitive style) using this same procedure. The procedure can also be used to analyze the effects of latent variables such as the use of supportive language (e.g., I agree, thank you, inviting replies, ask questions), like the results of a recent study shown in Figure 7 (Jeong, 2005b), and qualifiers when making a statements (e.g., maybe, I think) by replacing the team tags in column 1 with tags to indicate whether or not the message contained supportive language (e.g., CRITs versus CRIT), or qualifiers (e.g., ARGq versus ARG). This approach in particular, examines the combined effects of "how" messages are conveyed and the function of messages on the way participants respond to messages.

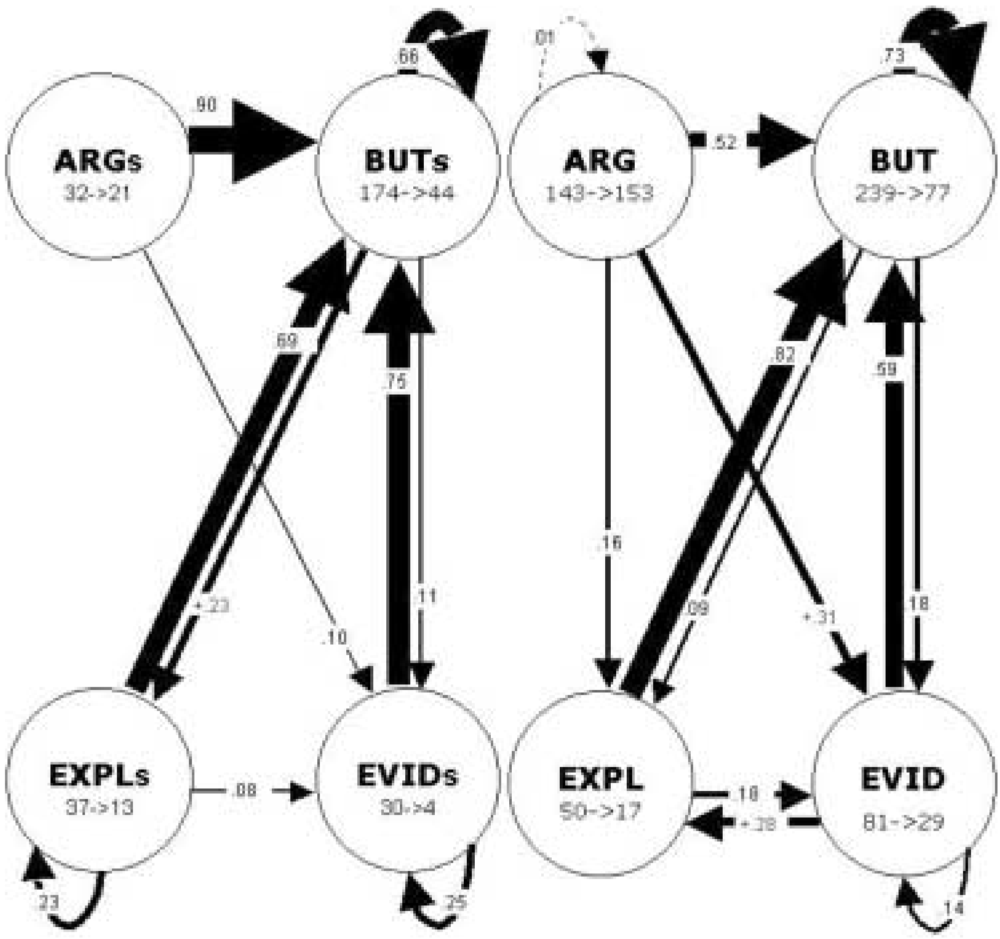

Figure 7. Results of a study comparing interactions produced by messages presented with versus without supportive style of communication. For example: The 32 arguments that were presented using a supportive style of communication elicited 21 total responses, where 90% of these responses were challenges. Probabilities presented with "+" indicate those that were significantly higher than the expected probability with Z scores > 2.32 at $p < .01$.

To test for differences in interaction patterns produced by all-male debates versus all-female debates (or mostly male versus mostly female debates), one can conduct an experimental study where the discussions generated by each group are *separately* collected and analyzed. Therefore, imagine that the probability matrix in Figure 6 was produced by an all-male group. To examine the interaction patterns between the males, the team tags in column 1 are removed to produce a 4 × 4 probability matrix (instead of an 8 × 8 matrix) using only the codes ARG, BUT, EVI, and EXP (without tags). The same procedure is used to separately analyze the response patterns in the all-female group discussion. Then compare the resulting Z scores and state diagrams between the all-male and all-female groups to see, for example, whether significant differences exist in response patterns, and whether or not the gender composition of discussion groups affects, for example, the mean number of challenges posted in response to arguments and the mean number of explanations posted in respond to challenges. This experimental design can also be used to test the effects of other exogenous variables—contextual variables that cannot be directly observed within the messages—choice of debate rules, constraints on response sequences (Jeong, 2003b), choice of message labels (Jeong & Joung, in press), assigning students to teams, and asynchronous versus real-time discussions.

Step 6: Compute transitional probabilities, Z scores and state diagrams

Once the codes have been prepared for analysis, DAT combs through each message thread to compute the frequency, transitional probability, and *Z* score for each message–response pair. DAT also computes the frequency distributions for the observed responses and messages, the number of messages that did not elicit a response, and the overall response rate. At this time, the frequency of event pairs for up to six categories can then be selected to produce state diagrams such as those presented in Figures 1 and 7. In addition, DAT supports the analysis of mean response scores by outputting the necessary numerical data for computing and testing mean response scores in statistical analysis programs like SPSS™ and Systat™ to conduct *t* tests, analyses of variance, regression analysis, multi-dimensional scaling, and other tests that might prove useful in gaining further insights into group interaction patterns and the effects of latent and exogenous variables.

The alternative to DAT is the GSEQ™ developed by Bakeman and Quera (1995). GSEQ™ performs a wide range of statistical functions that analyzes event sequences, timed-event sequences, interval sequences, and cross-classified events. What separates DAT from GSEQ™ is that DAT analyzes multi-threaded or multi-branching sequences of events (often observed in online threaded discussions, as illustrated in Figure 3), extracts message labels (if available) from message headers in discussion transcripts, makes the formulas and functions used to compute probabilities and *Z* scores more transparent within MS Excel™, provides immediate access to MS Excel™'s tools and functions for data preparation and analysis, identifies the location of each event pair tallied in frequency matrices, generates transitional state diagrams, and produces diagrams using arrows with varying densities to help discriminate response patterns.

Step 7: Interpret the transitional probabilities for interaction patterns

Arriving at a meaningful interpretation of interaction patterns revealed from the sequential analysis is often a difficult process due to the large number of statistics associated with each possible event pair and the inherently complex nature of group interaction. However, these difficulties can be largely avoided by focusing the analysis on only those event sequences that exemplify the processes believed to improve group performance and specified in your a priori hypotheses. The final word of caution is that when a particular pattern of interaction is revealed in a Z-score matrix (where the transitional probability of an event sequence is significantly higher or lower than the expected frequency), check to see that the finding is supported by sufficient cell frequencies in the frequency matrix for the given message–response pair, and the findings are not biased by coding errors in the message labels.

Implications for Instruction and Research

These tools and methods provide a road map to studying and modeling group interaction and the effects of specific variables on group interaction in CMC. This approach to studying online interaction will produce the research needed to guide instructional designers in developing collaborative learning activities that focus not on optimizing the sequencing of instructional content, but on optimizing the sequencing of speech acts to maximize group performance. Specifically, the methods and tools for supporting sequential analysis in CMC research can help produce the much-needed empirical research that designers need for improving online learning environments. In terms of the long-range implications, the proposed methods will provide a starting point for building computational models to explain, predict, and perhaps simulate group discussions in computer-mediated environments. Computational models of group processes combined with the use of techniques such as message labeling may serve as the mechanism for building intelligent discourse environments and simulators, and using them as learning objects as they dynamically model, catalog, and strategically sequence speech acts and content acquired from messages accumulated over time to facilitate, optimize, and/or simulate group discussions.

More detailed discussions of the limitations of the methods described are presented in the cited references. Nevertheless, some of the main limitations identified in previous studies provide additional insights on how best to conduct future research using this approach. The following are some of the following recommendations: Examine multiple discussion groups to prevent the idiosyncrasies of any one particular group from exerting too large an influence on the results; examine the interrelationship between multiple variables and their relative impact using multiple regression; examine the links between interaction patterns and group performance; expand the analysis to measure the frequency of three-event sequences to determine whether some event pairs are more effective in eliciting desired responses than other event pairs; identify sequences that distinguish experts from novices using multidimensional scaling; and test and validate process models across different types of tasks using new message categories and labels to facilitate discussions and to identify new patterns of interaction that support group performance.

In conclusion, these methods and tools can be used to model interaction patterns in any social exchange, including exchanges between instructors and students, coaches and athletes, counselors and patients, and humans and computers. Participants in face-to-face discussion can be asked to state their function during individual turns to facilitate the analysis of interaction patterns observed in face-to-face communications. This would lay the groundwork to studying the differences between face-to-face versus computer-mediated discussions in terms of interaction patterns produced by the presence versus absence of non-verbal behaviors, and how the differences in patterns contribute to group performance. Finally, these methods can also be used to model sequential patterns in cognitive operations performed by the individual while performing individual tasks. The hope is that these methods and tools will one day enable more researchers to apply sequential analysis to study and improve human learning and performance.

Notes on Contributor

Dr Allan Jeong is an assistant professor in the Instructional Systems program at Florida State University, teaching courses on distance education and conducting research on group interaction in computer-mediated communication.

References

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

Bakeman, R., & Quera, V. (1995). *Analyzing interaction: Sequential analysis with SDIS and GSEQ.* New York: Cambridge University Press.

Baker, M. (1999). Argumentation and constructive interaction. In P. Courier & J. E. B. Andriessen (Eds.), *Foundations of argumentative text processing* (pp. 179–202). Amsterdam: Amsterdam University Press.

Bakhtin, M. (1981). *The dialogic imagination.* (Ed., M. Holquist). Austin, TX: University of Texas Press.

Beers, P. J., Boshuizen, E., & Kirschner, P. (2004, April). *Computer support for knowledge construction in collaborative learning environments.* Paper presented at the Annual American Educational Research Association Conference, San Diego, CA.

Carr, C., & Anderson, A. (2001, March). *Computer-supported collaborative argumentation: Supporting problem-based learning in legal education.* Paper presented at the Annual Computer Support for Collaborative Learning (CSCL) 2001 Conference. Retrieved October 30, 2003, from http://www.mmi.unimaas.nl/euro-cscl/Papers/25.pdf

Cho, K., & Jonassen, D. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development, 50*(3), 5–22. Retrieved March 3, 2004, from http://tiger.coe.missouri.edu/~jonassen/Argumentation.pdf

Dick, W., Carey, L., & Carey, J. (2005). *The systematic design of instruction* (6th ed.). Boston: Allyn & Bacon.

England, E. (1985). Interactional analysis: The missing factor in computer-aided learning design and evaluation. *Educational Technology, 25*(9), 24–28.

Garrison, R. (2000). Theoretical challenges for distance education in the 21st century: A shift from structural to transactional issues. *International Review of Research in Open and Distance Learning, 1*(1), 1–17.

Gottman, J. M., (1979). *Marital interactions: Experimental investigations.* New York: Academic Press.

Gunawardena, C., Lowe, C., & Anderson, T. (1997). Analysis of global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17,* 397–431.

Jeong, A. (2003a). 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. (2003b, October). *The effects of message-reply and time-based structures on group interactions and critical thinking in asynchronous online discussions.* Paper presented at the Annual Association of Educational Communication and Technology Conference, Anaheim, CA.

Jeong, A. (2003c, October). *Gender interactions in online debates: Look who's arguing with whom.* Paper presented at a meeting of the American Educational Research Association, Chicago, IL.

Jeong, A. (2004a). The combined effects of response time and message content on group interactions in computer-supported collaborative argumentation. *Journal of Distance Education, 19*(1), 36–53.

Jeong, A. (2004c). *ForumManager.* Retrieved July 6, 2005, from http://garnet.fsu.edu/~ajeong

Jeong, A. (2005a). *Discussion analysis tool (DAT).* Retrieved April 18, 2005, from http://garnet.fsu.edu/~ajeong/DAT

Jeong, A. (2005b). *The effects of supportive styles of communication on group interaction patterns and argumentation in online discussion.* Proceedings of the Association of Educational Communication and Technology Conference 2005, Chicago, IL [CD-ROM].

Jeong, A., & Joung, S. (in press). The effects of constraint-based argumentation on interaction patterns and argumentation in online threaded discussions. *Computers and Education.*

Johnson, D., & Johnson, R. (1992). *Creative controversy: Intellectual challenge in the classroom.* Edina, MN: Interaction Book Company.

King, F., & Roblyer, M. (1984). Alternative designs for evaluating computer-based instruction. *Journal of Instructional Development, 7*(3), 23–29.

Koschmann, T. (1999). Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In C. M. Hoadley & J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) Conference 1999,* Stanford University, Palo Alto, CA (pp. 308–313). Mahwah, NJ: Lawrence Erlbaum.

Lemus, D., Seibold, D., Flanagin, A., & Metzger, M. (2004). Argument and decision-making in computer-mediated groups. *Journal of Communication, 54*(2), 302–320.

Mandl, H., & Renkl, A. (1992). A plea for "more local" theories of cooperative learning. *Learning and Instruction, 2,* 281–285.

McAlister, S. (2003). Assessing good argumentation. Retrieved April 10, 2004, from http://iet.open.ac.uk/pp/s.r.mcalister/personal/AssessingGEA.htm

Newman, D., Johnson, C., Cochrane, C., & Webb, B. (1996). An experiment in group learning technology: Evaluating critical thinking in face-to-face and computer-supported seminars. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, 4*(1), 57–74.

Olson, G., Herbsleb, J., & Rueter, H. (1994). Characterizing the sequential structure of interactive behaviors through statistical and grammatical techniques. *Human–Computer Interaction, 9*(3/4), 427–472.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education, 12,* 8–22. Retrieved July 09, 2005, from http://aied.inf.ed.ac.uk/members01/archive/vol_12/rourke/full.html

Sloffer, S., Dueber, B., & Duffy, T. (1999). *Using asynchronous conferencing to promote critical thinking: Two implementations in higher education.* Retrieved October 30, 2003, from http://crlt.indiana.edu/publications/crlt99-8.pdf

Strijbos, J. W., Martens, R. L., Jochems, W., M. G., & Kirschner, P. A. (2004, April). *The effect of functional roles on perceived group efficiency and communication during computer-supported collaborative learning.* Paper presented at the American Educational Research Association, San Diego, CA.

Veerman, A., Andriessen, J., & Kanselaar, G. (1999). Collaborative learning through computer-mediated argumentation. In C. M. Hoadley & J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) Conference 1999,* Stanford University, Palo Alto, CA (pp. 640–650). Mahwah, NJ: Lawrence Erlbaum.

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.