THE AMERICAN JOURNAL OF DISTANCE EDUCATION, *17*(1), 25–43

The Sequential Analysis of Group Interaction and Critical Thinking in Online Threaded Discussions

Allan C. Jeong
Florida State University

The purpose of this study was to examine group interaction and critical thinking in online threaded discussions. The Discussion Analysis Tool (DAT) was used to identify patterns in interactions and determine which interactions promoted critical thinking. With DAT, discussion transcripts were coded across twelve critical thinking events, and transitional probabilities between events were computed using the method of sequential analysis (Bakeman and Quera 1995). By computing the transitional probabilities, DAT generated useful quantitative descriptions of interaction patterns and critical thinking events that followed. The findings show that interactions involving conflicting viewpoints promoted more discussion and critical thinking, and that evaluation of arguments was more likely to occur as conclusions were being drawn—not as arguments were being presented. Tools such as DAT will be useful for empirically testing interactions and structures that enhance online discussions, providing the basis for more systematic testing of instructional interventions and computer-conferencing technologies.

Computer-mediated communication (CMC) is widely used to support group discussions in distance learning (Harasim 1993; Berge 1997) by applying the principles of constructivism (Muilenburg and Berge 2000), which focus on social interaction and the learning process (Driscoll 2000). Despite its popularity as an instructional tool, few theories and little empirical research account for the effects of CMC and instructional strategies on student interaction and learning processes in online discussions (Collins

Requests for reprints should be sent to Allan C. Jeong, Florida State University, Department of Educational Research, Open & Distance Learning, Stone Building 305E, Tallahassee, FL 32306–4453. E-mail: jeong@coe.fsu.edu

and Berge 2001). This shortcoming can be attributed to the lack of methods and tools capable of measuring group interactions and processes. Such tools and methods are sorely needed to advance the research in CMC (Hiltz, Johnson, and Turoff 1986; Anderson and Garrison 1995; Fahy, Crawford, and Ally 2001).

Content analysis is one method of analysis that has taken on increased emphasis in CMC research, focusing on the quality of messages in relation to performance in critical thinking and argumentation. However, content analysis cannot be used to examine the relation between threaded messages and how message sequence and group processes affect subsequent discussion and cognitive outcomes. Examining the relation between messages is the key to understanding group processes and interaction in CMC. Some attempts have been made to study patterns in the relation between messages (Levin, Kim, and Riel 1990; Newman, Webb, and Cochrane 1995; Gunawardena, Lowe, and Anderson 1997; Sudweeks and Simoff 1999; Fahy, Crawford, and Ally 2001). Levin, Kim, and Reil attempted to map and analyze message flow. Sudweeks and Simoff applied neural network analysis by assigning numerical values to the strength of interrelations between messages. Gunawardena, Lowe, and Anderson examined transitions between phases of critical thinking to illustrate the social construction of knowledge. These studies fall short in providing a robust methodology for measuring student interactions and examining how specific event sequences affect subsequent discussion and cognitive outcomes.

England (1985) and King and Roblyer (1984) proposed the use of sequential analysis (Bakeman and Quera 1995; Bakeman and Gottman 1997) to study computer-assisted instruction. In this study, sequential analysis was used to study group interactions in CMC, specifically to study the relations between messages and to determine which interactions support critical thinking (see Table 1). Sequential analysis is particularly appropriate for studying student interactions in threaded discussions. In threaded discussions, messages are hierarchically organized into threads or main headings and subsequent responses are displayed in subheadings. This structural organization allows students to review and selectively respond to messages presented earlier in the discussion. Every message and response is threaded or linked, forming a unit of interaction for sequential analysis.

Table 1. Event Categories in Coding System

Position (+/–)	A statement explicitly citing the individual's position. Assign valences for pro (+) or con (–) on the position debated by the group. If valence is unclear, omit from code.
Agreement	A statement of agreement.
Disagreement	A statement of disagreement.
Argument	A statement containing information to develop or support a position. Arguments can be in the form of (1) predicted consequences, implications, or problems through hypothetical or analogical reasoning (e.g., "they would, he could"); (2) proposed solutions to resolve or mitigate problems challenging the position; (3) personal beliefs, principles, or set of assumptions; (4) factual information describing events, objects, and circumstances.
Experience	A description of (1) events, objects, or circumstances drawn from personal experiences, actions, and observations; (2) emotional reactions/feelings to an issue.
Literature	Information drawn or cited from literature and reports, including TV and radio.
Data	Information or observations drawn from formal data *collected by participants* (not cited from literature).
Hypothetical action	A statement describing a *personal preferred* course of action used to evaluate the extent and validity of predicted consequences and implications. Look for "I would, I will."
Evaluation	A statement that *judges* the accuracy, likelihood, validity, logic, relative importance or value of an argument or claim by (1) making explicit judgment with words like "good," "true," "not likely"; (2) raising alternative viewpoints from which to make judgments.
Summary	A statement to review or summarize points raised in discussion.
Negotiation	A statement that relates to the negotiation of (1) meanings or definition of terms; (2) terms or conditions of an agreement/compromise on solutions to a problem or positions on an issue. Look for statements that declare a position on the issue with given terms, limits, or conditions.
Comments	A statement addressing issues related to the discussion *process* rather than the discussion *content.* These statements address (1) group procedures and participant interactions (if not already included in any codes previously listed); (2) personal train of thought or flow of written text; (3) the acknowledgment of member contributions.

Purpose of Study

The purpose of this study was to conduct a preliminary test of a software tool designed to perform complex and extensive computations required in event sequence analysis—a tool that will enable researchers to examine, measure, and empirically test student interactions and critical thinking in threaded discussions. The Discussion Analysis Tool (DAT) was developed to compute transitional probabilities between critical thinking events, providing the basis for measuring and describing the relation between threaded messages and student interaction. The computed probabilities generated by DAT were converted into graphical illustrations to provide a bird's-eye view of student interactions and critical thinking processes. Using this methodology, this study addressed the following questions:

1. What types of *two-event sequences* or interactions are most likely to occur in threaded discussions? For example, how likely is a position statement to elicit a response? How likely is a position statement to elicit a response with supporting arguments versus opposing arguments? Which of these possible types of event sequences and transitional probabilities are statistically significant and reveal significant patterns in student interactions? How do the observed patterns challenge or support existing models of critical thinking?
2. What events are likely to follow each type of interaction? For example, how many responses and what types of critical thinking can be expected to follow a "position statement → agreement" interaction versus a "position → disagreement" interaction in subsequent discussions? Which interactions draw the most responses, student participation, and critical thinking? What implications do these findings have on prescribed models of critical thinking and strategies for structuring discussions?

Theoretical Context

The theory of dialogism (Bakhtin 1981) was used to frame the research questions and methods addressed in this study. In this theory, language is viewed as part of a larger whole or social context in which all possible meanings of a word interact, possibly conflict, and affect future meanings. Meaning and critical thinking is produced by the relation between one utterance and another and is affected, renegotiated, and reconstructed as a result of conflict in social interactions. This conflict energizes and drives inquiry, reflection, and articulation of individual viewpoints and underlying assumptions. The theory's emphasis is on "depth and mutual understanding" of alternative conceptions through "dialogic reasoning," struggle, and voice (Tella and Mononen-Aaltonen 1998). Social interaction is essential to producing conflict and the social construction of new knowledge and meaning.

Based on the assumptions of the theory, this study focused not on the individual message in the analysis (as with content analysis) but on the relation and transitions between messages in threaded discussions. *Interaction* was operationally defined as two-event sequences composed of a given message and target message (or responding message). When pairs of threaded messages are followed by subsequent responses to advance a discussion thread, all subsequent messages within the thread are labeled as *lag events.*

Methodology

Participants

The participants were thirty-four master's of business administration graduate students (ten women, twenty-four men) enrolled in a face-to-face course in Business Ethics at a major midwestern university in the United States. The course was designed to educate students regarding the legal, moral, and ethical issues in business, to create sensitivity to the consequences of one's decisions, and to train students in critical thinking and moral/ethical analysis.

Instructional Task

Students were assigned to teams to debate an ethical issue for a period of four weeks on a threaded bulletin board. In groups of eight to nine students, Groups 1 and 2 debated the issue of electronic monitoring of employees, Group 3 debated the issue of collecting personal information for targeted advertising, and Group 4 debated the issue of hiring and firing employees based on employee conduct beyond the context of work. These issues were selected based on a survey that identified them to be the most controversial issues among classroom students. Students were assigned to groups by their initial position on an issue to maximize differences in viewpoints, and groups were balanced by gender and previous experience with threaded discussion boards. During the debates, the instructor abstained from intervening and participating in the discussions and sent e-mail only to remind students to meet the required minimum of two messages per week and to include descriptive titles with each posted message.

Procedures

The coding system (Table 1) was constructed from the analysis of discussion transcripts from a pilot test. The unit of analysis was the unit of meaning (Henri 1992), consisting of sentences or phrases. Using a grounded theory approach, twelve event categories were identified: position statements, statements of simple agreement or disagreement, arguments, personal experiences, literature, formal data, personal or hypothetical actions and choices, evaluation or critiquing of arguments, summary, negotiation or conclusions, and process comments.

Of the four debates, one was randomly selected and coded by the experimenter and a second observer to test for interrater reliability. The codings resulted in an 84.6% level of agreement. The estimated Cohen's Kappa Coefficient of reliability was .766—a very high interrater reliability given that a coefficient of .40 to .60 is considered fair, .60 to .75 is good, and over .75 is excellent reliability (Bakeman and Gottman 1997, 66). The coefficient takes into account the expected probabilities of agreement and disagreement according to the number of categories in the coding system. The coefficient was statistically significant with a squared variance of 0.012 and z score of 65.17.

Given the unique demands of sequential analysis and analysis of threaded messages, DAT was developed to perform functions that cannot be readily performed by existing transcript analysis tools such as Nudist and Atlas. DAT was designed to perform various types of sequential analysis on threaded messages. For each analysis, DAT identified and followed the links between messages in discussion threads—particularly links between messages spanning multiple levels of branching subthreads. The ability to operationalize and process the complex network of linked messages directly from the transcripts with DAT was critical to success in sequential analysis of threaded messages. Furthermore, automating these tasks using DAT eliminated manual data entry and the possibility of data entry error. DAT was also used to compute event frequencies, the transitional probabilities between events, and measures of statistical significance. In addition, DAT recorded the locations of each observed sequence in the transcripts to enable the review and retrieval of message texts for qualitative analysis of each observed interaction.

Results

Event Pairs and Transitional Probabilities

To examine the relation between messages, the function of each message was identified and coded according to the codes assigned to the message titles. During the debates, students were instructed to include descriptive titles to identify the main purpose of their responses. A total of 195 messages were coded. In thirteen messages, students forgot to submit a message title. In Table 2, the relative frequencies of events observed in the titles (see column "Given Events") were highly correlated (.94) with the relative frequencies of events observed in message texts. This finding suggests that coding messages based on message titles can provide a fair and practical method for coding messages. This process is similar to the approach used in scaffolded conferencing systems such as CoVis and CaMILE, where students select and assign labels to messages prior to posting them to discussion (Jonassen 1996).

The transitional probabilities between messages and responses are shown in Table 2. In each cell of the matrix is the transitional probability for each possible pairing of given and target events. For example, a message posted in the category "disagreement" was followed 38% of the time by an agreement, 13% of the time by disagreement, and 38% of the time by arguments. These probabilities were based on a total of eight target events (or replies) to the five given disagreements. The overall response ratio (eight targets to five givens) to any given disagreement was 1.60, higher than the average response ratio of .64. These probabilities were based on small cell frequencies, and, as a result, they must be interpreted accordingly.

Transitional State Diagram

The state transitional diagram in Figure 1 provides a visual illustration of the flow of events observed in the threaded discussions. In the diagram, the interactions that were most likely to occur are identified with darker lines, and interactions that were least likely to occur are identified with dotted lines. This visual representation distinguishes the common interactions from the less common interactions and highlights the flow between multiple-event sequences. For example, the diagram shows how position statements were most often followed by arguments and that arguments were then followed by additional arguments but rarely followed evaluation of arguments. For purposes of demonstration, only six of the twelve event categories were in-

Table 2. Relative Frequencies of Observed Events

	Position	Agree	Disagree	Argue	Exp	Lit	Data	Hypoth Act	Evaluate	Summary	Neg	Comment	# Targets	# Observed	RSP Ratio	Given Events (%)
Position	**.33**	.17	.00	.33	.00	.00	.00	.00	.00	.00	.00	.17	6	10	0.60	5.1
Agree	.00	.22	.00	.33	.00	.00	.00	.00	.11	.00	.22	.11	9	22	0.41	11.3
Disagree	.00	**.38**	**.13**	.38	.00	.00	.00	.00	.00	.00	.00	.13	8	5	1.60	2.6
Argue	.01	.16	.03	**.49**	.03	.00	.00	.04	<u>**.04**</u>	.00	.06	<u>**.13**</u>	69	72	0.96	36.9
Exp	.00	.00	.00	.00	**.50**	.00	.00	.00	**.50**	.00	.00	.00	2	5	0.40	2.6
Lit	.00	.00	.00	.00	.00	.00	.00	.00	**1.00**	.00	.00	.00	1	2	0.50	1.0
Data													0	0	0.00	0.0
Hypoth Act	.00	.00	.00	.00	.00	.00	.00	**.75**	.00	.00	.00	.25	4	6	0.67	3.1
Evaluate	.00	.20	.00	.20	.00	.00	.00	.00	**.40**	.00	.20	.00	5	14	0.36	7.2
Summary													0	0	0.00	0.0
Neg	.08	.04	.00	<u>**.20**</u>	.04	.00	.00	.00	**.20**	.00	.16	.28	25	18	1.39	9.2
Comment	.06	<u>**.03**</u>	.03	.30	.00	**.03**	.00	.00	.03	.00	.09	**.42**	33	41	0.80	21.0

Note: Agree = agreement; Disagree = disagreement; Argue = argument; Exp = experience; Lit = literature; Hypoth Act = hypothetical action; Evaluate = evaluation; Neg = negotiation; RSP = # targets/# observed. Values in bold are higher than the expected probability (z-score > 1.65, alpha < .10), and values in underline/bold are lower than the expected probability (z-score < 1.65, alpha < .10).

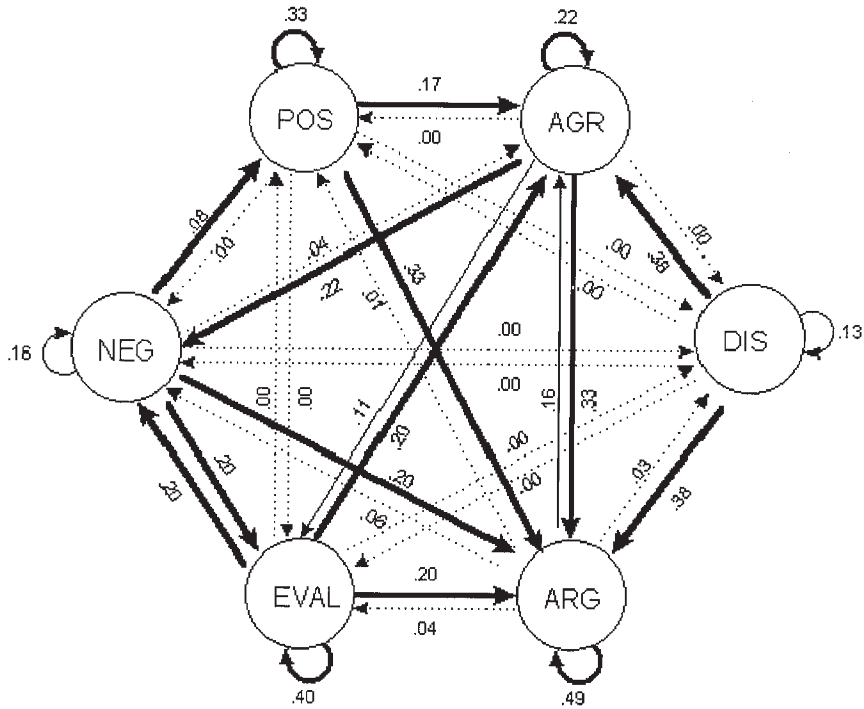

Figure 1. State Transitional Diagram. The circles in the diagram depict different events, and the arrows display the transitional probabilities between events. Pos = Position statement, Agr = Agreement, Dis = Disagreement, Arg = Argument, Neg = Negotiation, Eval = Evaluation.

cluded and were chosen according to observed frequencies and their relevance to interactions associated with conflict and argumentation.

Tests of Significance

Table 3 identifies the interactions that occurred at significantly higher probabilities than other interactions. For each possible event pairing, z scores were computed to reveal transitional probabilities that were significantly higher and lower than the expected probability (Bakeman and Quera 1995, 109). The z scores were computed for each possible event pairing while taking into account the differences in relative and observed frequencies of both given and target events. A z score of ±1.65 at .10 alpha level was used for this exploratory study.

Table 3. ***Z*-scores for Tests of Significance**

	Position	Agree	Disagree	Argue	Exp	Lit	Data	Hypoth Act	Evaluate	Summary	Neg	Comment
Position	**3.56**	0.33	–0.40	–0.13	–0.40	–0.20	0.00	–0.49	–0.77	0.00	–0.77	–0.26
Agree	–0.66	0.93	–0.49	–0.16	–0.49	–0.24	0.00	–0.61	0.27	0.00	1.49	–0.75
Disagree	–0.62	**2.22**	1.88	0.10	–0.46	–0.23	0.00	–0.57	–0.89	0.00	–0.89	–0.60
Argue	–1.55	1.20	0.30	**3.08**	0.30	–0.86	0.00	0.37	**–1.68**	0.00	–1.11	**–2.14**
Exp	–0.30	–0.53	–0.23	–1.06	**4.36**	–0.11	0.00	–0.28	**2.09**	0.00	–0.44	–0.73
Lit	–0.21	–0.38	–0.16	–0.75	–0.16	–0.08	0.00	–0.20	**3.26**	0.00	–0.31	–0.52
Data	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Hypoth Act	–0.43	–0.76	–0.32	–1.51	–0.32	–0.16	0.00	**7.65**	–0.62	0.00	–0.62	0.20
Evaluate	–0.48	0.53	–0.36	–0.75	–0.36	–0.18	0.00	–0.45	**2.53**	0.00	0.92	–1.17
Summary	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Neg	0.98	–1.38	–0.87	**–1.79**	0.54	–0.43	0.00	–1.07	**2.20**	0.00	1.42	0.94
Comment	0.55	**–1.82**	0.23	–0.74	–1.02	**1 .98**	0.00	–1.26	–1.29	0.00	0.10	**3.39**

Note: Agree = agreement; Disagree = disagreement; Argue = argument; Exp = experience; Lit = literature; Hypoth Act = hypothetical action; Evaluate = evaluation; Neg = negotiation. Values in bold identify transitional probabilities that were higher than the expected probability (z-score > 1.65, alpha < .10), and values in underline/bold identify transitional probabilities that were lower than the expected probability (z-score < 1.65, alpha < .10).

Despite the small cell frequencies, the z score tests revealed twelve interactions that occurred at probabilities significantly higher than expected and four interactions that occurred at probabilities significantly lower than expected. For example, the .49 probability that arguments were followed by another argument was significantly higher than expected (z score = 3.08, $p < .10$). In contrast, the .04 probability that Arguments were followed by an evaluation of the argument was significantly lower than expected (z score = –1.68, $p < .10$). Although larger cell frequencies are needed in order to draw any conclusions, the results of this study demonstrate how sequential analysis can be used to identify significant interaction patterns when larger cell counts are obtained. When cell counts are small, this method of analysis can be useful as a diagnostic tool for monitoring student performance in an instructional context.

Events Following Interactions

DAT was used to compute the number and types of responses following each interaction pair up to the end of a discussion thread. A total of fifty-one unique paired interactions were observed in this study. Table 4 shows the responses to the nineteen most frequent interactions. For each given interaction, the types of critical thinking events and how likely they were to follow the given interaction are represented in transitional probabilities. For example, the interaction Disagree → Disagree was observed at a frequency of 1 and was followed by a total of six threaded responses. Responses that followed this interaction were likely to be 33% arguments, 33% hypothetical actions, 17% agreements, and 17% process comments. In this case, the average number of messages elicited by the Disagree → Disagree interaction was 6.0 messages. The overall average number of messages subsequent to any given interaction was 1.2 messages. Of the fifty-one interactions, eighteen (or 35%) received no subsequent responses. More data will be needed in future experiments to obtain sufficient cell frequencies to test for statistical significance.

Discussion

The methods and software tools developed in this study generated data to measure and identify significant patterns in student interactions. With the theoretical framework of this study, interaction patterns of most interest were those associated with conflict and argumentation. For example, this study found that responses to position statements were most likely supporting or

Table 4. Response Types and Probabilities Following Interactions

Event Pairs	Position	Agree	Disagree	Argue	Exp	Lit	Data	Hypoth Act	Evaluate	Summary	Neg	Comment	Lag Events	Frequency	Avg. # Replies
DD	.00	.17	.00	.33	.00	.00	.00	.33	.00	.00	.00	.17	6	1	6.0
PC	.00	.20	.00	.40	.00	.00	.00	.00	.00	.00	.00	.40	5	1	5.0
CN	.00	.00	.08	.85	.00	.00	.00	.00	.00	.00	.00	.08	13	3	4.3
PP	.00	.25	.00	.38	.00	.00	.00	.00	.00	.00	.00	.38	8	2	4.0
RC	.00	.09	.03	.59	.00	.00	.00	.00	.06	.00	.06	.16	32	9	3.6
CD	.00	.00	.00	.33	.00	.00	.00	.00	.00	.00	.33	.33	3	1	3.0
EN	.00	.33	.00	.00	.00	.00	.00	.00	.67	.00	.00	.00	3	1	3.0
RR	.00	.08	.01	.62	.00	.00	.00	.05	.11	.00	.03	.09	74	32	2.3
AC	.00	.00	.00	1.00	.00	.00	.00	.00	.00	.00	.00	.00	2	1	2.0
CE	.00	.50	.00	.00	.00	.00	.00	.00	.00	.00	.50	.00	2	1	2.0
CR	.00	.20	.00	.35	.00	.00	.00	.00	.10	.00	.10	.25	20	10	2.0
DC	.00	.00	.00	.50	.00	.00	.00	.00	.00	.00	.50	.00	2	1	2.0
DR	.00	.00	.00	.33	.00	.00	.00	.33	.00	.00	.17	.17	6	3	2.0
NR	.00	.00	.10	.80	.00	.00	.00	.00	.00	.00	.00	.10	10	5	2.0
RD	.00	.00	.00	.75	.00	.00	.00	.00	.00	.00	.25	.00	4	2	2.0
RN	.00	.00	.00	.29	.14	.00	.00	.00	.00	.00	.00	.57	7	4	1.8
CC	.00	.14	.00	.50	.00	.00	.00	.00	.00	.00	.05	.32	22	14	1.6
NC	.10	.10	.10	.20	.00	.00	.00	.00	.00	.00	.20	.30	10	7	1.4
NE	.00	.17	.00	.00	.00	.00	.00	.17	.50	.00	.17	.00	6	5	1.2

Note: P = position statement; A = agreement; D = disagreement; R = argument; E = experience; L = literature; T = data; H = hypothetical action; S = summary; N = negotiation, C = process comments; Agree = agreement; Disagree = disagreement; Argue = argument; Exp = experience; Lit = literature; Hypoth Act = hypothetical action; Evaluate = evaluation; Neg = negotiation. The table displays the types and number of responses to 19 of the 51 event pairs observed in the study. The average number of subsequent responses per interaction was 1.19.

opposing arguments and that arguments were likely to generate additional arguments in subsequent responses. This study also found that disagreements were rarely posted in response to position statements and arguments, whereas agreements were ten times more likely to be posted in response.

Although statements of disagreement were rare, review of the transcripts indicated that disagreement occurred when arguments and counterarguments were exchanged. In this study, the coding scheme did not include event categories to distinguish supporting arguments from counterarguments because they were often difficult to discriminate. One solution to this problem is to assign students to debate teams and instruct students to label their messages by function and by team membership. An argument posted by one team (Team *s*) and a threaded response with an argument from the opposing team (Team *o*) would reveal such an interaction based on the message labels alone (e.g., ARG*s* → ARG*o*). Another possible solution is to subdivide the argument category to distinguish the role of challenging and critiquing of arguments from the role of stating claims and assertions. With these changes, the interactions involving disagreement, arguments, and counterarguments could be examined in more detail.

This study also found that students rarely responded to arguments with evaluation of the argument's accuracy, validity, and relevancy. The results indicate that evaluation is more likely to occur in the discussions when the group is negotiating a conclusion or consensus. In this case, evaluations were presented after arguments were presented and not during the presentation of arguments. The transitional state diagram suggests that evaluation occurred when discussions flowed from negotiation to evaluation (20%) to arguments (20%) to agreement (16%), and back to negotiation in an iterative process. This pattern suggests that evaluation primarily supports the negotiation phase of discussions.

Implications for Critical Thinking Models

The interactions observed in this study were consistent overall with existing models of critical thinking. Gunawardena, Lowe, and Anderson (1997) formulated a critical thinking model consisting of twenty-one categories hierarchically organized into five phases (Table 5). Many of the event sequences observed in this study were consistent with Gunawardena, Lowe, and Anderson's model. These sequences include the transition from position statement to agreement, agreement to arguments, and position statement to arguments. No transitions from position statement to disagreement were observed in this study, which is consistent with the omission of

Table 5. Gunawardena's Critical Thinking Model

Phase I: Sharing/comparing of information (92.7% of observed events)
- A. A statement of observation or opinion
- B. A statement of agreement from one or more other participants
- C. Corroborating examples provided by one or more participants
- D. Asking and answering questions to clarify details of statements
- E. Definition, description, or identification of a problem

Phase II: The discovery and exploration of dissonance or inconsistency among ideas, concepts, or statements (2.4% of observed events)
- A. Identifying and stating areas of disagreement
- B. Asking and answering questions to clarify the source of extent of disagreement
- C. Restating the participant's position and possibly advancing arguments or considerations in its support by references to the participant's experience, literature, formal data collected, or proposal of relevant metaphor or analogy to illustrate point of view

Phase III: Negotiation of meaning/co-construction of knowledge (1.9% of observed events)
- A. Negotiation or clarification of the meaning of terms
- B. Negotiation of the relative weight to be assigned to types of argument
- C. Identification of areas of agreement or overlap among conflicting concepts
- D. Proposal and negotiation of new statements embodying compromise, co-construction
- E. Proposal of integrating or accommodating metaphors or analogies

Phase IV: Testing and modification of proposed synthesis or co-construction (1.0% of observed events)
- A. Testing the proposed synthesis against "received fact" as shared by the participants and/or their culture
- B. Testing against existing cognitive schema
- C. Testing against personal experience
- D. Testing against formal data collected
- E. Testing against contradictory testimony in the literature

Phase V: Agreement statements/applications of newly constructed meaning (1.9% of observed events)
- A. Summarization of agreements
- B. Applications of new knowledge
- C. Metacognitive statements by the participants illustrating their understanding that their knowledge or ways of thinking (cognitive schema) have changed as a result of the conference interaction

Note: Reprinted from C. Gunawardena, C. Lowe, and T. Anderson, "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 (4): 397–431. Copyright © 1997. Used by permission of Baywood Publishing Company, Inc.

disagreement in Phase IB of Gunawardena, Lowe, and Anderson's model. The findings in this study also support the transition from stating disagreements to advancing arguments, from identification of agreements to negotiating conclusions, and evaluation of arguments during negotiation rather than during the sharing of arguments.

Some of the observed interactions also challenge Gunawardena, Lowe, and Anderson's (1997) model. In this study, disagreement with position statements and arguments took the form of counterarguments. Gunawardena, Lowe, and Anderson's model prescribes a process in which positions and supporting arguments are presented in a phase *before* discussion of differences and disagreements in a following phase. The violation of this process and its impact on subsequent discussion must be examined in more detail by applying new methods for coding supporting versus opposing arguments, as described earlier. Testing the impact of this pattern of *cross-arguing* will provide empirical evidence to support, refute, or modify existing models of the critical thinking process.

The observed transition from evaluation to negotiation is consistent with Gunawardena, Lowe, and Anderson's (1997) model that prescribes the transition from weighing of arguments to offering proposals. However, this study found that the transitions between negotiation and evaluation were bidirectional and just as likely to move in either direction. The transitions were better described as a cyclical process of negotiation, evaluation, and validation—a process not suggested in Gunawardena, Lowe, and Anderson's model or any other models of critical thinking (Henri 1992; Garrison 1992). Further analysis of these interaction patterns could identify processes that improve the negotiation process.

Limitations of Study

Although this study was able to find some significant interaction patterns, the high number of event pairs tested for significance made the likelihood of Type I errors unacceptable. As a result, the findings can only be interpreted as exploratory in nature. The application of DAT in future studies may require some or all of the following actions: (1) reduce the testwise alpha; (2) restrict the number of a priori tests to specific event sequences believed to be problematic in any given model of critical thinking; and (3) obtain, code, and analyze larger data sets.

In this study, the data set consisted of only 208 messages. As a result, the observed frequencies in this study were small due to the relatively large number of possible event pairings. Each addition of an event category to the coding system exponentially increases the number of possible event sequences, potentially resulting in yet smaller cell frequencies. As a result, data sets larger than the one used in this study will be necessary to continue the examination of interactions using sequential analysis.

Some of the challenges of collecting and coding large data sets can potentially be resolved by requiring students to label and categorize the messages they post to discussions. This procedure has been applied in discourse systems such as Knowledge Integration Environment (Bell and Linn 1997), Collaboratory Notebook (O'Neill and Gomez 1994), and Computer-Supported Intentional Learning Environments (Scardamalia and Bereiter 1994) to scaffold domain-specific conversations and problem solving. Jonassen and Remidez (2002) proposed a system that enables instructors to define event categories and restrict event sequences. This coding method eliminates the time-intensive task of coding transcripts and will potentially enable real-time analysis of group discussions.

Conclusion

The software tool DAT and methods developed in this study were successfully used to produce some of the first empirical evidence to support critical thinking processes and sequences prescribed in current models of critical thinking. The continued use of DAT will be instrumental in measuring and testing additional forms of interactions and their impact on student discussions. Most of all, DAT will be useful for evaluating the effectiveness of different instructional interventions, discourse structures, and communication technologies on group interactions and learning outcomes.

In this study, minimal interventions were presented in order to obtain baseline measures of event frequencies and event sequences. Over time, different interventions and structures can be introduced and varied systematically to measure and compare their effects on student interactions and outcomes. Many different interventions, discourse structures (Jonassen and Remidez 2002) and strategies for moderating discussions (Collins and Berge 2001) have yet to be empirically tested. Some of these interventions include the manipulation of group size, grouping by gender, setting the length of discussion periods, allowing anonymous participation, and providing discussion structures.

This study demonstrates the potential power of sequential analysis and its ability to provide quantitative as well as qualitative descriptions of group interactions in threaded discussions. The author hopes that the results of this study will form the basis of a new framework for advancing empirical research in distributed learning environments and distance education, and that the development of new technologies will support the monitoring, evaluation, and assessment of student performance in online discussions.

References

Anderson, T., and D. R. Garrison. 1995. Transactional issues in distance education: The impact of design in audioteleconferencing. *The American Journal of Distance Education* 9 (2): 27–45.

Bakeman, R., and J. Gottman. 1997. *Observing Interaction: An introduction to sequential analysis.* Cambridge: Cambridge University Press.

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

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

Bell, M. C., and P. Linn. 1997. Scientific arguments as learning artifacts: Designing for learning on the Web. Paper presented at the annual meeting of the American Educational Research Association, April, Chicago, IL.

Berge, Z. 1997. Characteristics of online teaching in post-secondary, formal education. *Educational Technology* 37 (3): 35–47.

Collins, M., and Z. Berge. 2001. Resources for moderators and facilitators of online discussion. Available online at http://www.emoderators.com/moderators.shtml

Driscoll, M. P. 2000. *Psychology of learning for instruction,* 2nd 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.

Fahy, P. J., G. Crawford, and M. Ally. 2001. Patterns of interaction in a computer conference transcript. *International Review of Research in Open and Distance Learning* 2 (1). Available online at http://www.irrodl.org/content/v2il/fahy.html

Garrison, D. 1992. Critical thinking and self-directed learning in adult education: An analysis of responsibility and control issues. *Adult Education Quarterly* 42 (3): 136–148.

Gunawardena, C., C. Lowe, and T. Anderson. 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 (4): 397–431.

Harasim, L. 1993. Collaborating in cyberspace: Using computer conferences as a group learning environment. *Interactive Learning Environments* 3 (2): 119–130.

Henri, F. 1992. Computer conferencing and content analysis. In *Computer supported collaborative learning,* ed. C. O'Malley, 117–136. Heidelberg, Germany: Springer-Verlag.

Hiltz, S. R., K. Johnson, and M. Turoff. 1986. Experiments in group decision making—Communication process and outcome in face-to-face vs. computerized conferences. *Human Communication Research* 13 (2): 225–253.

Jonassen, D. H. 1996. *Computers in the classroom—Mindtools for critical thinking.* Englewood Cliffs, NJ: Prentice Hall.

Jonassen, D., and H. Remidez. 2002. Mapping alternative discourse structures onto computer conferences. In *Computer support for collaborative learning: Foundations for a CSCL community,* ed. Gerry Stahl, 237–244. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Levin, J., H. Kim, and M. Riel. 1990. Analyzing instructional interactions on electronic message networks. In *Online education,* ed. L. Harasim, 185–213. New York: Praeger.

Muilenburg, L., and Z. Berge. 2000. A framework for designing questions for online learning. Available online at http://www.emoderators.com/moderators/muilenburg.html

Newman, D., B. Webb, and C. Cochrane. 1995. A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century* 3 (2): 56–77.

O'Neill, D. K., and L. M. Gomez. 1994. The collaboratory notebook: A distributed knowledge building environment for project learning. *Proceedings of ED MEDIA,* 94. Available online at http://www.covis.nwv.edu/info/papers/pdf/oneill-edmedia-94 Vancouver, BC, Canada.

Scardamalia, M., and C. Bereiter. 1994. Computer support for knowledge-building communities. *Journal of the Learning Sciences* 3: 265–283.

Sudweeks, F., and S. J. Simoff. 1999. Complementary explorative data analysis. In *Doing Internet research: Critical issues and methods for examining the net,* ed. S. Jones, 29–55. Thousand Oaks, CA: Sage.

Tella, S., and M. Mononen-Aaltonen. 1998. *Developing dialogic communication culture in media education: Integrating dialogism and technology.* Media Education Publication 7:58. Available online at http://www.helsinki.fi/-tella/mep7.pdf