Studying the Effectiveness of the Discussion Forum in Online Professional Development Courses

Susan Lowes
Peiyi Lin
Yan Wang
Teachers College/Columbia University

Abstract

As online professional development courses for teachers have grown, the discussion forum has become a locus of considerable research. This study analyzes the discussion forums in four different sessions of a short (4-week) online course for teachers from six schools in three states. This study also compares four methodologies, all of which have a visualization component: an analysis of data from the CMS; network analysis; content analysis; and sequential analysis. In addition, this study describes the insights into the effectiveness of the course design and facilitation that each approach provides, correlates these with participant satisfaction, and argues for using a combination of methods when studying discussion forums in online courses.

Introduction: Context

Online professional development courses for teachers have been growing rapidly over the past few years. The structure of these courses varies enormously—some are fully online, some meet occasionally face-to-face, some last a full semester, some are more limited in time and scope—but almost all use the threaded discussion forum as a central locus of course activity. The discussion forum has thus become the subject of considerable research, both in terms of designing discussion forum activities that support learning and of using the discussion forums to create a community that will support learning (Collison, Elbaum, Haavind, & Tinker, 2000; Barab et al., 2004; Swan, 2003b; Swan, 2004).

The site of this study was a short (4-week) fully online professional development course for teachers in four middle and two high schools that were in the process of being accredited to offer a whole-school reform program that covers grades 6–10 and thus bridges middle and high school.[1] The course was thus part of a larger whole-school reform initiative and was designed to introduce the teachers to the fundamental principles of the program's approach to learning.

Although the course content was specific to the goals of the particular program, the course structure is representative of an increasingly common online professional development model: it was short term; conducted entirely online; facilitated; and combined readings, discussions, and a group project. Like many such courses, it was designed to be highly interactive, with the discussion forum being the site for most of the course's interactivity, the "place" where participants would meet to exchange ideas, discuss assignments, and share work, both with each other and the facilitator. There were

[1] The program takes an approach to learning whose hallmarks are project-based learning, interdisciplinarity, and collaboration within and across grade levels. Additional teachers from three already accredited middle schools and one accredited high school joined the final sessions.

several assignments each week, all of which required posting to the discussion forum, so that overall the participants were expected to post several times a week.

The course content has been offered for many years in face-to-face, one-day professional development workshops and the online course was expected to provide a similar experience, although with the opportunity for more reflection and possibly more depth. The online version was designed and facilitated by two experienced workshop leaders. They had not previously taught online, however, and were therefore required to take a preparatory course (also online) on effective methods for facilitating online learning. This preparatory course emphasized the role of the discussion forum in the course and the importance of facilitator presence in creating effective discussions.

The four sessions of the course that are the focus of this study (we will abbreviate them and call them Sess1, Sess2, Sess3, and Sess4) were offered at different times in Spring 2004 and Fall 2005. They all had the same content and they were taught by one of the two facilitators; each had approximately 20 participants. The participants were experienced classroom teachers (depending on the session, between 75% and 90% had taught for more than 5 years) and taught many different subjects. Almost all of the participants in Sess2, Sess3, and Sess4 reported in a pre-course survey that they were "very" or "somewhat" comfortable using such everyday technologies as email, word processing, and Internet research--Sess1 was an exception, with much higher percentages of "not very" and "not at all" in answers to these questions. However, all the participants were at best only somewhat familiar with the new program or with the content addressed in the course, and very few had ever taken an online course or participated in an online discussion group or bulletin board. They can thus be considered relatively knowledgeable about technology but not about what a successful online experience would look like.

Research Questions

Since the question for the organization offering the course was whether the existing face-to-face experience would translate successfully to the online environment, the question for us as researchers became one of deciding what to look for online that would provide useful insights into the nature and success of the online learning community. We thus began with two sets of questions.

The first set was methodological: What different insights could different methodologies provide into the structure and functioning of these discussion forums, and into the role of the facilitator? Would the results of different methodologies be complementary, repetitive, mutually exclusive, or conflicting? Equally important, are those methodologies that provide the most useful insights necessarily the most time consuming, or are there methods that provide a lot of information without a huge amount of work? And finally, we were particularly interested in methodologies that use visualizations because they can be so useful in communicating with the public.

The second set of questions was more substantive and has already been subject of considerable study—indeed, as we will see, it is these prior studies that provide the framework for two of the methodologies. These questions address the content of the interaction: How does the content of the interaction in the discussion forums affect participant participation? As an extension of this, are there differences in terms of the content, quantity, timing, and nature of the postings in the discussion forums that we can correlate with overall course satisfaction?

Methodologies

We chose four methodologies, ranging from the very simple (and quick) to the more complex (and time consuming), that look at the nature of interactions in online forums: (1) an analysis of data provided by Blackboard, the course management system; (2) a network analysis of the discussion forum interactions, using UCINET for statistical data and NETDRAW for graphical representations; (3) a content analysis of the discussion forum posts, including a visualization that added a time dimension; and (4) a sequential analysis of the content analysis using Allan Jeong's Discussion Analysis Tool (Jeong, 2003).

Data

Because we knew that some of the methodologies would be very time consuming to implement, for all but the Blackboard analysis we confined our data set to the interactions in the discussion forums during the third week of the 4-week course. We chose this week both because it was a week in which the discussion topics were addressed to the entire group (rather than to separate small-group discussions, as in the second week) and also because by this point course participation had "solidified," so that drop-outs and "fade-aways" would not have to be considered in the analysis.[2] By week 3, the number of participants in each session was:

Sess1: 21
Sess2: 24
Sess3: 18
Sess4: 17

The curriculum for week 3 had four discussion topics, or threads, with the initial question posed by the facilitator (more precisely, by the course designer, so that it was the same from session to session, no matter who was facilitating). Since we were interested in interactivity and since the rules of the course required every participant to post at least one response to each topic, we did not consider these initial responses to be true interactions and so excluded them from the data set. This is an important difference between this analysis and those of other researchers, since it dramatically reduced the number of posts that we had to code.

Statistical Analysis of Blackboard Data

Blackboard provides course facilitators with graphical analyses and downloadable "course statistics" in the form of hit counts. The charts are colorful but difficult to read and essentially useless in terms of the data they analyze (see Figure 1):

[2] Drop-outs were those who registered for the course but never began and fade-aways were those who began but faded away over the four weeks. Both are major issues, both for those who are conducting online courses and for those who are trying to analyze them.

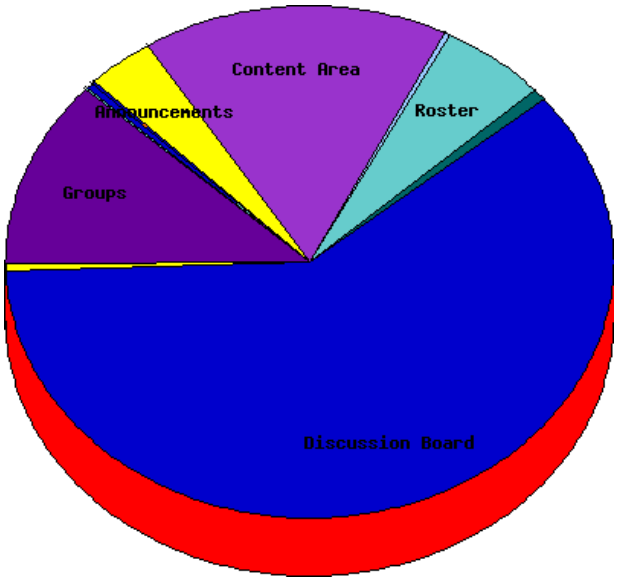


Figure 1. Blackboard Course Statistics

The hit data, however, can be exported and analyzed for each section of the site and for each participant, as well as by time of day and by individual each day. Since it is the easiest data to collect, we thought it was important to include it in our analysis.

Table 1 shows the average number of hits per participant and per facilitator for the discussion forums over the entire four weeks of the course.[3] The first point to note is that the sessions with the larger number of participants have the higher average hits per participant (but not per facilitator). This is probably because the more participants there are, the more posts there are to read and therefore the more times each participant needs to return to the discussion board during the week. The second point to note is that the average number of participant hits per session seems to be inversely correlated with the average number of facilitator hits per session. For instance, Sess1 has the highest average per participant but the lowest average per facilitator, while Sess3 has the highest average per facilitator but the lowest per participant.

Table 1: *Average Number of Discussion Board Hits per Participant and per Facilitator, over Entire Session*

	Participants	**Facilitators**
Sess1	1318	114
Sess2	1277	121
Sess3	1074	180
Sess4	1151	118

[3] To be completely comparable with the other analyses, we should have concentrated on week 3 data only, but although Blackboard lets you narrow to specific dates and to specific sections of the site, it does not let you confine your analysis to specific forums. In addition, since many participants posted late, using the dates would have lost a large number of hits.

If hits were a measure of interaction, we would conclude from this data that Sess1 was the most successful session, followed by Sess2, Sess4, and Sess3.

Hit data counts how many times a section of the site was "hit," or accessed, but says nothing about what the visitor was doing, which could range from passing through quickly, spending time reading posts, actively contributing, or even clicking in an area by mistake. In addition, Blackboard provides no explanations for the counts, other than to warn that "due to the way that statistics are collected, not all totals are consistent"—hardly a reassuring statement. This suggests that the Blackboard data may be seriously misleading. A more sophisticated analysis, one that takes relationships and content into consideration, is necessary.

Social Network Analysis

Social network analysis has been used for some time to analyze patterns of interaction among participants in a group or organization as the result of computer-mediated communication (Garton, Haythornthwaite, & Wellman, 1997; Haythornthwaite, 1998; Haythornthwaite, 2002; Barry Wellman's NetLab). However, most of these studies use surveys and interviews to elicit the quality and content of the network ties. In addition, although there is already considerable research on the facilitator's role in online discussion forums (Paz Dennen, 2005; Pape et al., 2005; Barab et al., 2004; Collison et al., 2000), there is less on the participants' roles and still less on the interaction between the two (Lalli & Feger, 2005). Our goal in considering online discussions to be social networks was to see if, by using the discussion forum itself as the subject of study and by then examining the direction and value (strength) of the relationships among the participants (in other words, who interacted with whom and how often), we could discern differences among the sessions.

To prepare the data, we created spreadsheets that showed which participants responded to each other's and the facilitator's posts and how often they did so. (As noted above, since the initial post was by the facilitator and everyone was required to respond at least once to this post, these initial posts were not considered interaction and were not counted.) Table 2 is a graphical version of one portion of this analysis. The data is cumulative: for example, it shows that the facilitator (T) responded to participant S1 four times during the week and to participant S17 five times.

Table 2: *Partial Matrix, Sender to Receiver*

Receiver

Sender	T	S1	S2	S3	S4	S5	S6	S7	S8	S9	S 10	S 11	S 12	S 13	S 14	S 15	S 16	S 17	S 18
T		4	1	3	**2**	3			1	1	**2**		1	1	3	1		5	4
S1						1	1												
S2				1		1				1								1	1
S3	1																		
S4	**2**					**2**								**2**			1	1	**2**
S5		1	1		1					1		1	1						
S6						1													
S7			1											1					
S8	**2**						1	1						1				1	
S9	1	1			**2**												**2**	1	
S10			1													1			
S11																			
S12					1	1												1	
S13	1	1			1	1											1		3
S14	1																		
S15							1	1										**2**	1
S16					1			1	1					1					
S17	1				1	**2**	1							1	1	**2**	1		**2**
S18	**2**			1	1		1	1						1			1	1	

This spreadsheet was imported into the network analysis software (we used UCINET), which provides three statistical measures:

Density: The number of connections between contributors as a percent of all possible connections. A density of 100% would mean that everyone talked to everyone else at least once—all the squares in the matrix would be filled in.

Network centralization: A measure of the extent to which the network centers on one person or is dispersed among all the participants. A centralization of 100 % would mean that all the participants talked to only one person.

Share: The percentage of each participant's contributions to the total of all contributions.

Reciprocity: This is a measure that we calculated for the number of contributors who got responses from the contributors they posted to.

Table 3 shows the figures for the four sessions, with the highest figures highlighted in red and the lowest figures highlighted in gray:

Table 3: *Network Analysis Data*

	Sess1	**Sess2**	**Sess3**	**Sess4**
Density	18%	23%	23%	35%
Network centralization	15%	26%	17%	40%
Facilitator's contribution (share)	18%	16%	19%	24%
Highest single participant contribution (share)	15%	9%	14%	7%
Second highest participant contribution (share)	6%	9%	10%	6%
Reciprocity	8/22 = 36%	18/25 = 72%	11/19 = 58%	16/18 = 89%

This analysis begins to give us a picture of the differences between the sessions. First, Sess4 and Sess1 appear to be sharply contrasting experiences. Sess4 has the highest network density, meaning that more of the participants interacted with each other than in the other sessions. It also has the highest degree of centralization, in this case because of high facilitator involvement, but this is counterbalanced by a very high degree of reciprocity among participants. Compare this to Sess1, which had the lowest density, lowest network centralization, and lowest level of reciprocity. The low level of network centralization was because of low facilitator involvement and even lower participant involvement overall: in fact, the share data shows that the facilitator and three participants accounted for 46% of all interactions, indicating that there were a few dominant participants and many more who were close to inactive. Second, we can see that Sess2 is more like Sess4, while Sess3 is more like Sess1, with Sess2 and Sess4 appearing to be the more successful sessions, with much higher levels of reciprocity than Sess1 and Sess3.

The network software imports into NETDRAW to produce elegant diagrams that help visualize this kind of data. The network diagrams for Sess1 (Figure 2) and Sess4 (Figure 3) show visually the differences in overall density and also the contrast between the two sessions in terms of the participants' contributions. The size of the circle indicates each participant's relative share while the position of the circle shows centrality and the thickness of the connecting lines shows frequency of posts. In both Sess1 and Sess4, the largest circle in the center (T1) is the facilitator (note, however, that the size of the node is determined by the software and cannot be compared from diagram to diagram). In Sess1, two participants—S1 and S10—dominate, while in Sess4, despite the centralization on the facilitator, the participant interactions are also much more evenly dispersed among the entire group:

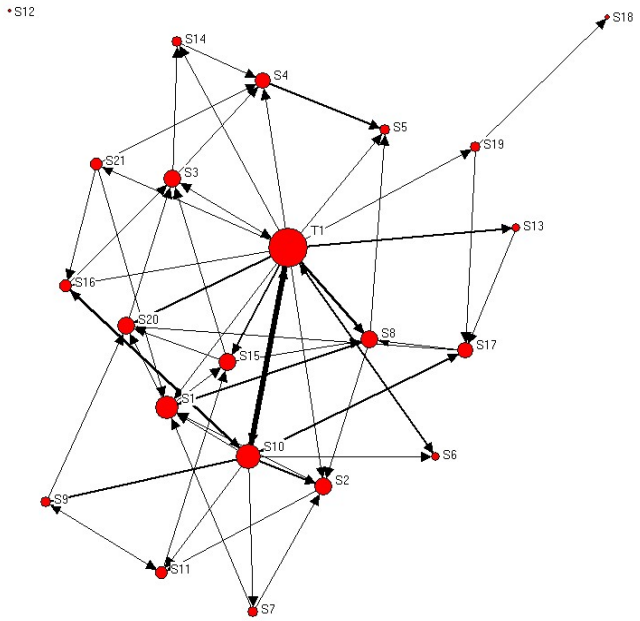


Figure 2. Network diagram for Sess1

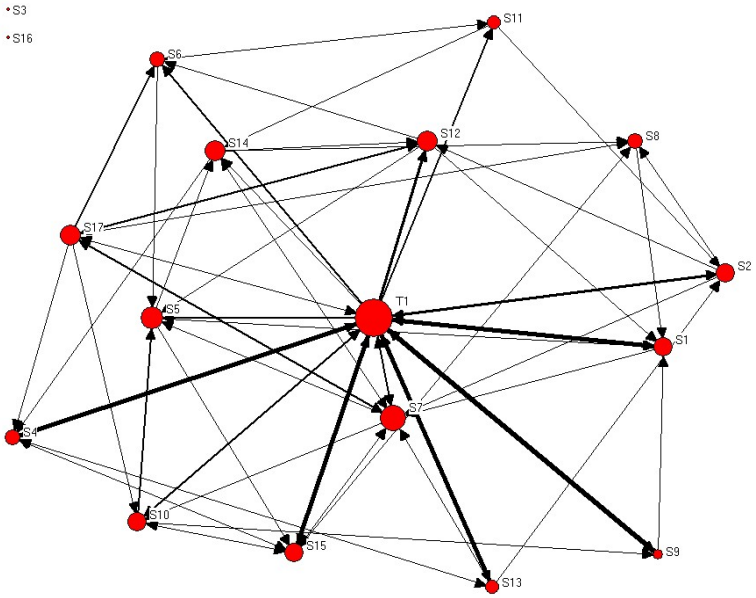


Figure 3. Network diagram for Sess4

The next two diagrams focus attention on the reciprocal relations by highlighting the two-way arrows in bold.[4] They show graphically how Sess4 (Figure 4) had far more reciprocal interactions than Sess1 (Figure 5):

[4] Although NETDRAW produces two-way arrows, they are difficult to distinguish, so they were manually highlighted to produce these diagrams.

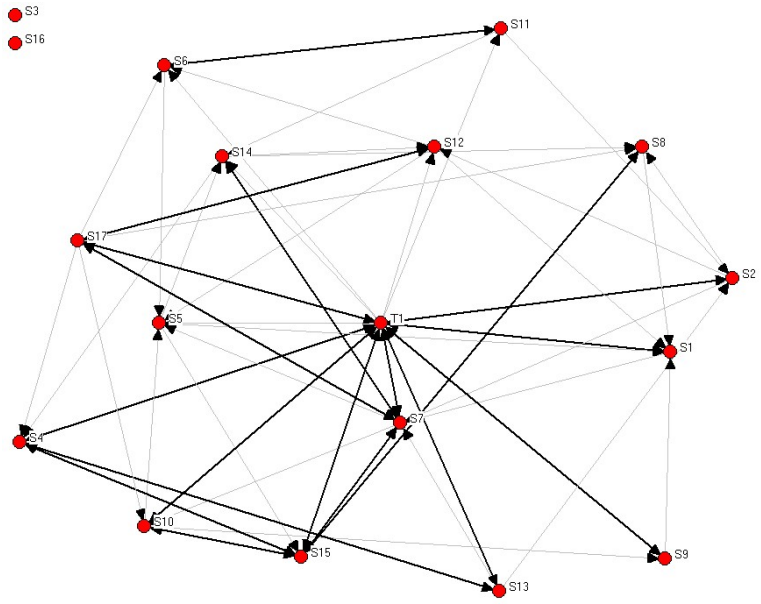


Figure 4. Reciprocity highlighted: Sess4

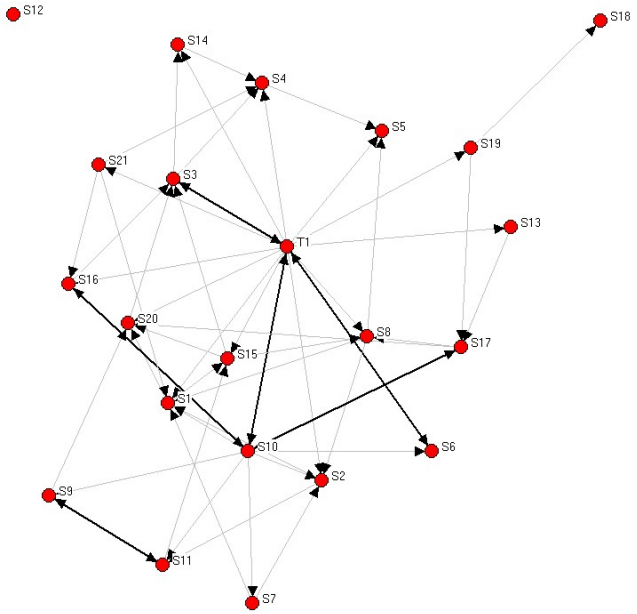


Figure 5. Reciprocity highlighted: Sess1

The network analysis suggests that online courses that are supposed to be the same (or at least similar) may, when implemented, have very different patterns of interaction. It also showed that interactions in the discussion forum are part of a complex whole, so that both participant-participant *and* facilitator-participant interactions need to be taken into account to get a picture of the relationships among the participants.

Content Analysis

While network analysis can help us understand and describe the extent of interaction among participants, it does not tell us anything about the content of this interaction. The next step, therefore, was to look at that content, and then to see if some types of content were correlated with more successful interaction patterns than others.

Content or conversation analysis has been widely used in the research on computer-mediated communication. (Much of this work has been published in the *Journal of Computer-Mediated Communication* ; see, for example, Herring, 2004; also see Mazur, 2004, for a review.). One strand of this research looks at the interactions between the participants in computer-mediated conferences (i.e., discussion forums) in terms of three types of "presence"—social presence, cognitive presence, and teaching presence (Garrison, Anderson, & Archer, 2001; Rourke, Anderson, Garrison, & Archer, 2001). We chose to focus on social presence and teaching presence because previous research has shown that both are associated with high levels of participant satisfaction (Gunawardena & Zittle, 1997; Richardson & Swan, 2003; Swan, 2003a; Swan, 2003b). Participants need to feel the teaching presence of the facilitator and the social presence of other learners—put simply, they need evidence that someone is listening. This requires active participation on the part of both participants and facilitators; without it, they feel abandoned.

Social presence is a factor in any learning environment, but it needs a higher level of assertiveness in online than face-to-face courses: even those who are silent in a face-to-face classroom can indicate presence through unspoken means (nods, smiles, etc.), but this is not the case in an online discussion forum. Teaching presence is important not only because the participants need to feel that the teacher is there for them but because of the role the teacher plays in facilitating the participants' social presence.

As researchers have worked with these concepts, the analytical categories have become increasingly complex (see Table 4). While Garrison, Anderson, and Archer (2001) outlined three categories of social presence, Rourke, Anderson, Garrison, and Archer (2001) relabeled them and added three or four indicators for each, for a total of twelve indicators. Swan (2001) then took the categories set out by Rourke et al. and further refined them, for a total of fifteen indicators (see Appendix 1 for a fuller description of Swan's indicators).

Table 4: *Analytical Categories*

Garrison et al.	**Rourke et al.**	**Swan**
Emotional expression	Affective responses: Expression of emotions Use of humor Self-disclosure	Affective responses: Expression of emotion Humor Self-disclosure Value Paralanguage
Open communication	Interactive responses Continuing a thread Quoting from others' messages Referring explicitly to others' messages Asking questions Complimenting, expressing appreciation Expressing agreement	Interactive responses Continuing a thread Acknowledgment Referring explicitly to others' messages Invitation Appreciation Agreement/disagreement
Group cohesion	Cohesive responses Vocatives Addresses/refers to the group using inclusive pronouns Phatics, salutations	Cohesive responses Vocatives Group reference Greetings and salutations Social sharing, phatics

As Rourke et al. (2001) point out, even with 12 indicators, the data analysis was extremely time-consuming, and in addition there were problems of inter-rater reliability with several of the indicators (i.e., what is humor to one coder may not be humor to another). For our purposes, these highly evolved variables seemed better suited to the fairly extensive discussions in the graduate-level courses that both Rourke et al. and Swan (2001) were studying than to our more limited professional development course, where the participants were full-time teachers fitting the work into busy schedules and where the course design (with three or four discussion threads each week) was less conducive to extended interactions. We had also been struck by how many of the interventions did not seem to contribute materially to the discussion or move it forward—they seemed to be acting as what Hewitt (2005) calls "clunkers." We therefore focused on those variables that we felt were likely to facilitate (or not) interactivity in the forums.

This led us to make rather different distinctions than those made by Rourke et al. (2001) or Swan (2001). Thus, while Swan has agreement/disagreement as one variable, we suspected that there was a huge difference between the two in terms of their role in generating two-way interaction, so we separated them. We also split Swan's Acknowledgment variable to fit what we saw in our forums, which was a difference in the effect of simple acknowledgment compared to acknowledgment that included new information and was therefore likely to contribute to further discussion and

understanding: when the acknowledgment took the form of simple agreement or appreciation, we coded it as Cheerleading/affirming (an example of this is a post that says "I agree with you completely!"), but when it included ideas or concepts that expanded on the previous thread, we coded it as "New information." Our re-coding (Table 5) therefore produced the following much simplified, and much easier to implement, schema:

Table 5: *Coding Schema*

Swan	**Code**	**New category**	**Definition**
Interactive: Agreement Appreciation	1	Cheerleading/affirming	Offering praise and encouragement to others; expressing agreement
Interactive: Acknowledgment	2	New information	Expand on previous posts; introduce new ideas; sharing new information
Interactive: Disagreement Invitation	3	Questioning/challenging	Raise questions that extend previous post, or express disagreement with it

We hypothesized that although Code 1 (cheerleading and/or affirming) was necessary to create social presence, when used alone it would not move the discussion forward. On the other hand, providing new information (Code 2) and questioning/challenging (Code 3) were more likely to do so.[5]

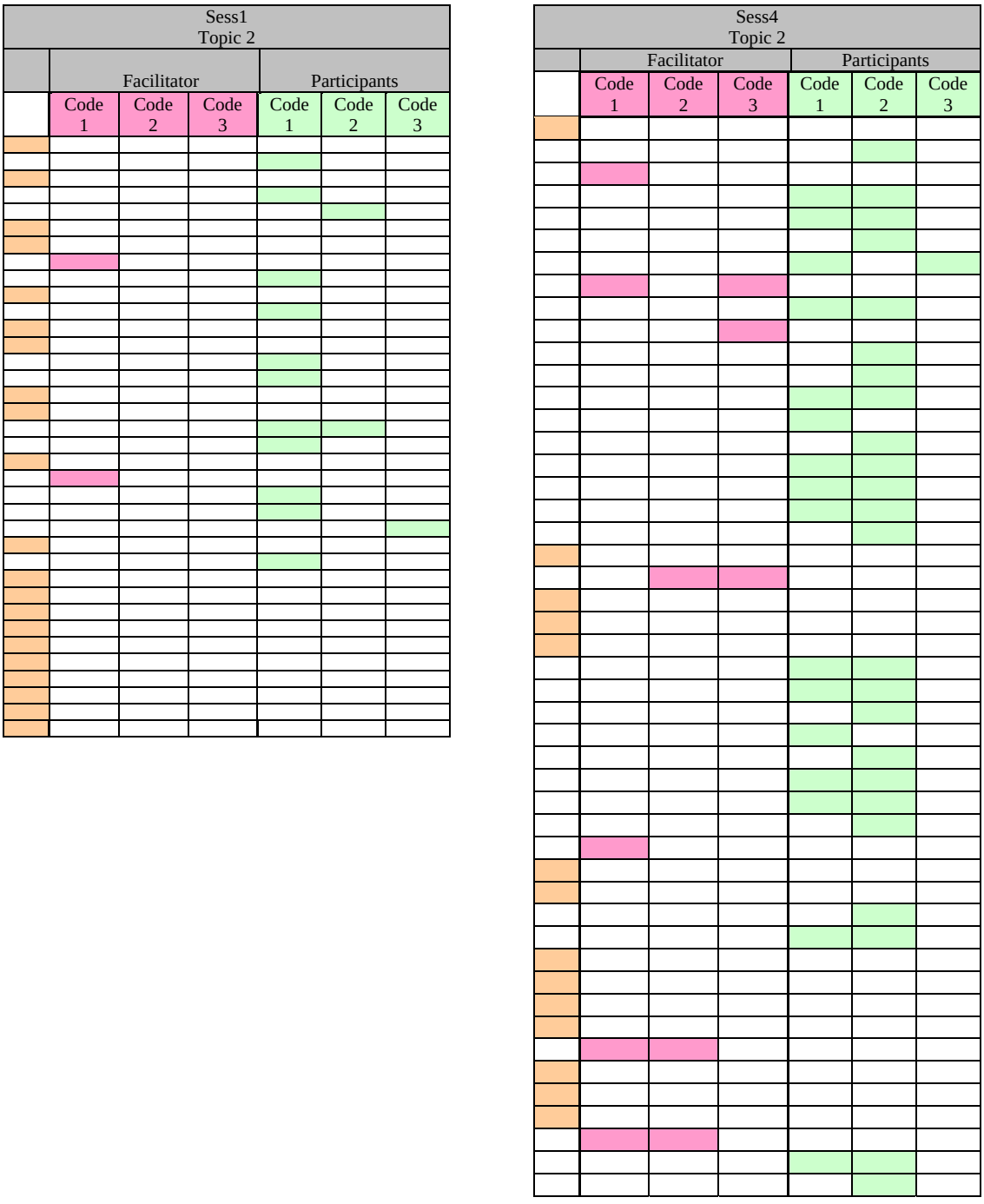
Laying the codes for each post out on a timeline allowed us to see how the conversation evolved. Table 6, which covers one topic during week 3, shows the initial posts (the ones that were required as part of the course structure) in orange. Looking down the left-hand column from orange block to orange block (initial post for each participant), we can see how much further apart the orange blocks are in Sess4 compare to Sess1, indicating much more conversation between initial posts. In fact, by the end of the thread in Sess1, there are no responses at all—the thread has died. When we checked the dates, we could see that this was probably because these posts were made so late that the rest of the group had moved on.[6]

Reading across the rows shows graphically how the responses from both the facilitator and participants in Sess1 were heavily weighted with cheerleading (Code 1), while the participants in Sess4 had more posts that offered new information and the facilitator in that session did more questioning/challenging:

[5] Two researchers coded the data separately, cross-checked their results, and came to an agreement on those entries where they differed. As with the network analysis, in our coding we did not include the first facilitator post (i.e., the start of the thread) or the first response, since it was required and therefore was not considered an indicator of interaction.

[6] See Hewitt 2005 for a discussion of this issue and an alternative hypothesis.

Table 6: *Topic 2, Week 3, Comparison of Code Types*



Content Analysis: Code Frequencies

As the timeline shows, one post could have more than one code: for instance, a post could be simple cheerleading but it could also be cheerleading plus new information. We therefore ended up with a total of seven codes for both facilitators and participants, ranging from single codes to multiple combinations—from simple to more complex responses—as shown in Table 7.

Table 7: *Code Combinations*

Code 1: Cheerleading
Code 2: New information
Code 3: Questioning/challenging
Code 4 (1+2): Cheerleading + new information
Code 5 (1+3): Cheerleading + questioning/challenging
Code 6 (2+3): New information + questioning/challenging
Code 7 (1+2+3): Cheerleading + new information + questioning/challenging

An examination of the percentages of each of these codes for the entire week (Table 8) provides further confirmation of the differences between the sessions, and again shows the contrast between Sess1 and Sess4.[7] When we look at the facilitators, we see that Sess4 had the lowest percentage of Code 1 for the facilitators, as well as the highest percentage of posts that included questioning/challenging (Codes 3, 5, and 7: 39 percent), while almost 50% of the interventions in Sess1 were cheerleading alone and only 12% of posts included questioning/challenging.

Table 8: *Coding by Type of Intervention for Facilitators as Percent of Total Codes*

Code	**Sess1**	**Sess2**	**Sess3**	**Sess4**
1	46%	53%	61%	38%
2	12%	11%	9%	11%
3	8%	11%	9%	14%
4 (1+2)	23%	19%	13%	14%
5 (1+3)	0%	3%	4%	11%
6 (2+3)	8%	3%	4%	14%
7 (1+2+3)	4%	0%	0%	0%
Facilitator total	100%	100%	100%	100%

For the participants, on the other hand, while Sess4 showed an even more striking difference between it and the other sessions in the percentage of cheerleading alone (Code 1), it was also distinguished from the other sessions by the percentage of posts that included new information (Code 2)—a huge 77% (Table 9). In other words, while the facilitator of Sess4 intervened through questions, the participants tended to offer new

[7] Note that the percentages are percentages of total codes not total posts.

information. Contrast this with Sess1, where the facilitator posts most frequently offered new information, but the participants were heavily weighted toward cheerleading:

Table 9: *Coding by Type of Intervention for Participants as Percent of Total Codes*

Code	**Sess1**	**Sess2**	**Sess3**	**Sess4**
1	54%	44%	57%	12%
2	14%	19%	27%	40%
3	13%	10%	2%	4%
4 (1+2)	13%	22%	13%	37%
5 (1+3)	5%	4%	2%	1%
6 (2+3)	2%	0%	0%	6%
7 (1+2+3)	0%	2%	0%	0%
Participant total	100%	100%	100%	100%

This streamlined version of content analysis helped explain the differences between Sess1 and Sess4. The low density and low levels of reciprocity of Sess1 were associated with a facilitator who tended to provide information and participants who tended to simply cheerlead. In Sess4, in contrast, where both density and reciprocity were much higher, there was very little cheerleading: the facilitator was more likely to question/challenge while the participants were more likely to offer new information. Further, although it might be expected that greater interaction among participants would also be associated with lower centralization on the facilitator, in fact this was not the case here: it was the content of the facilitation that was the key, not the extent of the interventions.

Sequential Analysis

The timeline (Table 6) showed graphically that cheerleading alone tended to stop the conversation, but it also seemed to indicate that it was not new information or questioning/challenging, either alone or in combination, that were most likely to lead to further discussion, but cheerleading plus one of these (i.e., Codes 4, 5, or 7). To further analyze the sequential relationship between the different types of postings, we used Allan Jeong's Discussion Analysis Tool (Jeong, 2003). This is a predictive tool that analyzes which types of intervention are more likely to generate additional interventions and what form those subsequent interventions are likely to take. In other words, it allows us to confirm statistically what we thought we could see graphically and in the percentage data. Since Jeong's tool can at present only draw graphs for a maximum of 6 codes and since we had far fewer codes 5, 6, or 7 (all of which involved questioning), we combined these into one. Table 10 shows the re-coding:

Table 10: *Recodes of Coding Schema*

Code 1: Cheerleading	Code 1
Code 2: New information	Code 2
Code 3: Questioning/challenging	Code 3
Code 4 (1+2): Cheerleading + new information	Code 4
Code 5 (1+3): Cheerleading + questioning/challenging	Code 5
Code 6 (2+3): New information + challenging/questioning	Code 5
Code 7 (1+2+3): Cheerleading + new information + challenging/questioning	Code 5

Table 11 shows the reply rates—the chance that a post of a particular type will get a reply—as analyzed by the DAT tool. This confirms the low probably that cheerleading (Code 1) will get a response (21% chance), compared to the much higher probability (51%) that either new information (Code 2) or questioning/challenging (Code 3) will do so:[8]

Table 11: *Response Probability Rates*

Types of posting	**All sessions**
Code 2: New Information	.51
Code 3: Questioning	.51
Code 5: Questioning + other	.47
Code 4: Cheerleading + New info	.43
Code 1: Cheerleading	.21

The DAT tool also provides sequence diagrams that show the likelihood of one type of response leading to another. Figure 6 is the diagram that results when all sessions are combined.[9] Two of the results are statistically significant. When a post that is only cheerleading (Code 1) does get a response (and remember that the likelihood is small), there is a 60% chance that that response will be more cheerleading (Code 1). Similarly, if

[8] Although the results also show that the more complex responses are less likely to get a response than the simple ones, we suspect that this is the result of the codes that we combined into Code 5. We are now looking at other ways to combine our codes (for instance, all codes that include cheerleading plus something else into one new code).

[9] To have some predictive certainty, Jeong's tool also requires a larger population than any one of our sessions provided. We therefore began by looking at Sess4 and Sess2, which had similar profiles in terms of density, reciprocity, facilitator participation, participant interaction, and satisfaction, then added in Sess3 and Sess1. The results were roughly the same in each case, which is what we would expect: if one type of posting is likely to lead to another, this will hold true across all sessions. We are therefore only showing all sessions combined here. The statistical data produced by the tool are included as Appendices 2 and 3.

a post that is questioning/challenging (Code 3) does get a response, there is a 56% chance that that response will be new information (Code 2).

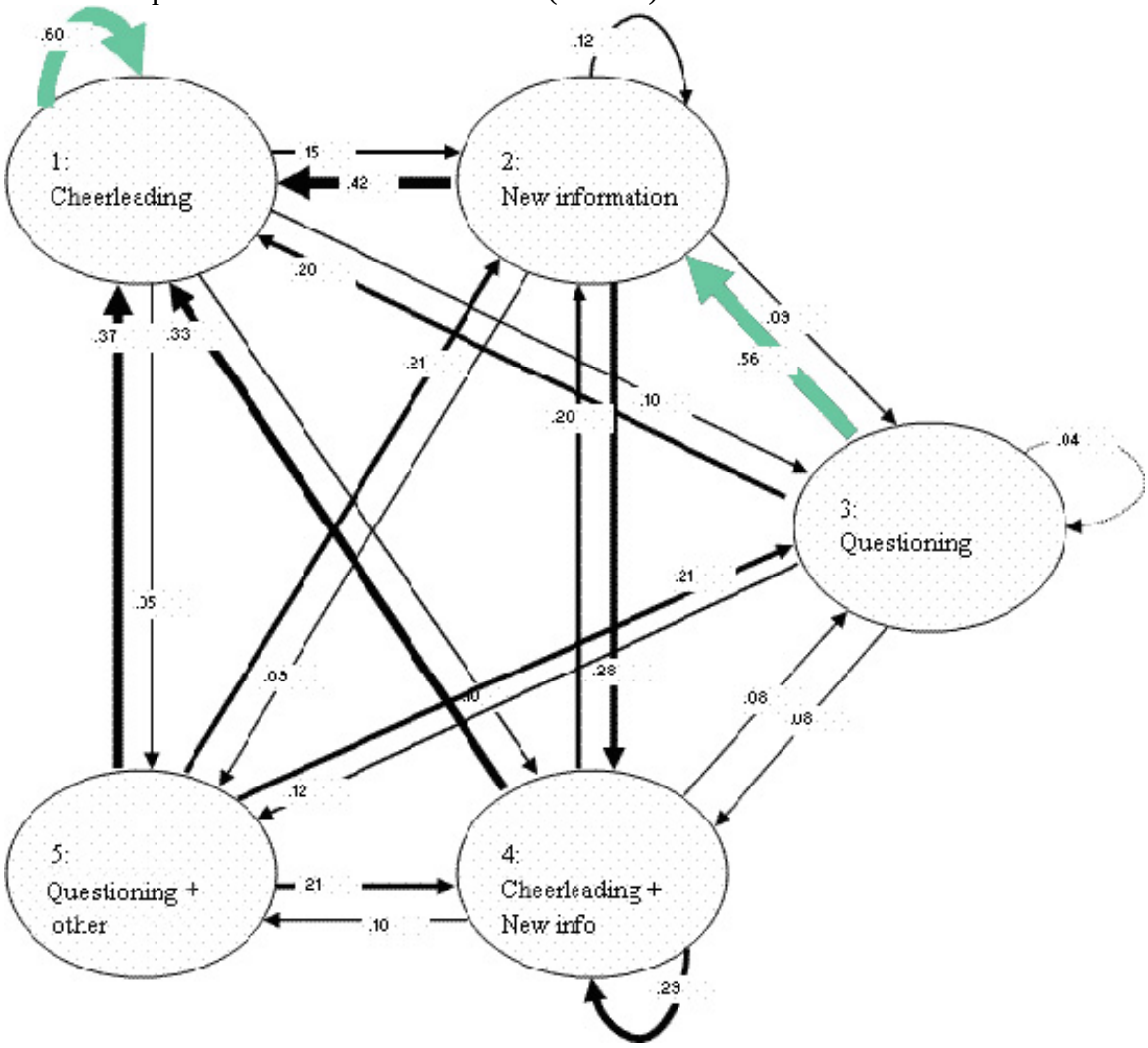


Figure 6. Sequential Analysis

Discussion

We began this research with a set of questions about both methodology and course process. In terms of methodology, we were interested in which methodological approaches would provide the most useful insights; whether the results would be complementary, repetitive, mutually exclusive, or conflicting; and whether those methods that did provide useful insights were necessarily those that took the greatest amount of time to implement. In terms of course process, we also wanted to better understand the relationship between the extent of participant contributions to the discussion forums and the nature or content of those contributions, as well as how (and if) these correlated with overall course satisfaction.

The final course survey asked a number of satisfaction questions. For the purposes of this particular study, we focus on the participants' perceptions of the overall value of the course, as well as their perceptions of the success of those aspects that

involved communication and interaction, including facilitation. (We therefore are not looking at issues surrounding usability of the interface, navigation, etc., which we also asked about.) In order to compare course sessions, we converted the survey's Likert-scale ratings to means. The first three items had a 5-point scale, while the last item had a 3-point scale. The columns are organized from the session with the highest rankings to the session with the lowest rankings, with the top two rankings in each category highlighted.

Table 12: *Course Satisfaction by Session*

	Sess2	**Sess4**	**Sess3**	**Sess1**
Quality of facilitation (range 1 – 5)	4.53	4.20	4.18	3.67
Quality of other participants' contributions (range 1 – 5)	4.37	4.13	4.18	4.05
Enjoyment of discussion forums (range 1 – 3)	2.53	2.60	2.24	2.43

Sess2 and Sess4 had the highest ratings, Sess1 had the lowest, while Sess3 fell in between. The responses to a question that asked participants whether they would take another online course showed the same pattern (see Table 13): far higher percentages in Sess2 and Sess4 answered "Yes, definitely" than for the other two sessions, with Sess1 having the fewest of all:

Table 13: *Willingness to Take another Online Course*

	Sess2	**Sess4**	**Sess3**	**Sess1**
Would you take another online course?	58%	60%	35%	24%

In terms of course satisfaction, then, two of the sessions—Sess2 and Sess4—were ranked higher by the participants than the other two sessions, with Sess1 consistently ranked lowest.

The first data set we looked at, from the Blackboard CMS, was not only inconsistent with these results but was based on an essentially useless measure (hits) that provides no insights into what visitors are actually doing in the discussion forums.

The network analysis, in contrast, provided a wealth of interesting information on all these issues, and was relatively easy to implement. Despite the fact that the four sessions were supposed to mirror each other, the results from the network analysis showed that there were very different patterns of participant-participant and facilitator-participant interaction in each session. If the sessions are ranked from high to low in terms of density, reciprocity, and network centralization, we see that the results are not only highly correlated with each other but are also correlated with the satisfaction ratings (Table 14):

Table 14: *Network Measures from High to Low*

Measure	**Sess4**	**Sess2**	**Sess3**	**Sess1**
Density	35%	23%	23%	18%
Reciprocity	89%	72%	58%	36%
Network centralization	40%	26%	17%	15%
Willingness to take another online course	60%	58%	35%	24%

In addition, the network share analysis provides a window into the dominance (or not) of the facilitator and the dispersion (or not) of the conversation among the participants, which may help explain why some sessions were more highly rated others. For instance, one of the sessions with the highest satisfaction ratings (Sess4) had the highest facilitator involvement, while the one with the lowest satisfaction rating (Sess1) was dominated by one participant (Table 15):

Table 15: *Share Ratings from High to Low*

Measure	**Sess4**	**Sess2**	**Sess3**	**Sess1**
Facilitator share	24%	16%	19%	18%
Highest participant share	7%	9%	14%	15%

Finally, although the graphical view of the networks provided by the software may provide a useful visualization that could help communicate the differences between sessions to a lay audience, we found that the statistical data was not only easier to interpret but allowed us to construct profiles of each session.

Content analysis is a potentially a rewarding methodology that can provide important insights into why a session is successful, but it is highly labor intensive, even using our much simplified coding scheme and even given a course that did not have, overall, a large number of posts. Nevertheless, the results of the content analysis were not only highly correlated with the satisfaction ratings, but also helped explain some of those ratings. Laying the codes out on a timeline showed graphically how the sessions differed in the depth and extent of forum conversations, while the frequency analysis of the types of posts allowed us to compare the differences among sessions in terms of the nature and complexity of the posts. This analysis once again confirmed the differences between the sessions.

None of these distinctions would be important if they did not have consequences in terms of leading to deeper and more engaged interactions. The sequential analysis (which could not be done without the content analysis and is therefore also time

consuming) shows the likelihood that one type of post will lead to another. While on the face of it the results in this case might seem obvious, they would have been obscured if we had used the indicators of social presence that have been classified by other researchers as interactive (i.e., by Rourke et al. 2001 and Swan 2001, among others). In fact, these researchers' interactive posts are only interactive in a passive sense: in other words, they may be interactive in that the person who posts is interacting with a previous poster, but they are not interactive in the sense that they lead to further interaction. The DAT tool helps us look at this post-to-post sequence; it would seem particularly useful for comparing changes after design interventions.

With the exception of the Blackboard data, each of these methodologies—network analysis, content analysis, and sequential analysis—correlated with the participants' satisfaction with the course, yet each provided very different kinds of information on the patterns of interaction during the four sessions. They can thus be considered complementary rather than repetitive, mutually exclusive, or conflicting. In an ideal world, more than one method would be used to evaluate effectiveness, with each being mined for the insights it provides. However, the different methods involved very different time commitments, with network analysis taking far less time than content/sequential analysis. Thus while content analysis has become a locus of much research over the last few years, this comparison suggests that looking at network characteristics may also provide a fruitful way of understanding, comparing, and evaluating discussion forums, not only those used in online professional development courses but in many other settings as well. In addition, the visual impact of the interaction matrix, the network analyses, the content analysis-over-time chart, and the sequential analysis diagrams have proved very useful in discussions with practitioners.

Implications for practice and further research

In addition to having implications for researchers, our analysis has implications for practitioners. While it is highly unlikely that facilitators in online learning environments will have the time to do a full-scale network or content analysis, the analysis shows that if they are to make these experiences as valuable as possible for the participants, they need to go beyond counting thread-depth or looking at lines contributed per participant and look at the shape, content, and sequence of the interaction, using the results as a corrective as the course evolves. For instance, if the facilitator of Sess1 had known what to look for, he/she would have quickly seen that cheerleading was stopping the discussion in its tracks; that facilitator interventions, while possibly helpful, were not pushing the discussion forward; and that one or two people were dominating the interactions. In other words, the analysis shows that it is not only the amount of participation but the content of the participation and most especially the *sequence* of participation that is important for the success of a course.

References

Collison, G., Elbaum, B., Haavind, S., & Tinker, R. (2000). *Facilitating online learning: Effective strategies for moderators*. Madison, WI: Atwood Publishing.

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education, 15*(1), 7–23.

Garrison, R. (2003). Cognitive presence for effective asynchronous online learning: The role of reflective inquiry, self-direction and metacognition. In J. Bourne & J. C. Moore (Eds.), *Elements of quality online education*, Vol. 4 (pp. 47–58). Needham, MA: Sloan Center for Online Education.

Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying online social networks. *Journal of Computer Mediated Communication, 3*(1). Retrieved August 25, 2007, from http://jcmc.indiana.edu/vol3/issue1/garton.html

Haythornthwaite, C. (1998). A social network study of the growth of community among distance learners. *Information Research, 4*(1). Retrieved August 25, 2007, from http://informationr.net/ir/4-1/paper49.html

Haythornthwaite, C. (2002). Building social networks via computer networks: creating and sustaining distributed learning communities. In K.A. Renninger & W. Shumar (Eds.), *Building virtual communities: learning and change in cyberspace* (pp. 159–190). Cambridge: Cambridge University Press.

Herring, S. C. (2004). Computer-mediated discourse analysis: an approach to researching online behavior. In S. Barab, R. Kling, & J. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338–376). Cambridge: Cambridge University Press, 2004. Draft version retrieved August 25, 2007, from http://ella.slis.indiana.edu/~herring/cmda.html

Hewitt, J. (2005) Toward an understanding of how threads die in asynchronous computer conferences. *Journal of the Learning Sciences, 14*(4), 567–589.

Jeong, A. (2003). Sequential analysis of group interaction and critical thinking in online threaded discussions. *American Journal of Distance Education, 17*(1), 25–43. Retrieved August 25, 2007, from http://dev22448-01.sp01.fsu.edu/Research/Publications/SequentialAnalysis_Jeong2003.pdf

Lalli, C. B., & Feger, S. (2005). Gauging and improving interactions in online seminars for mathematics coaches. Providence, RI: Northeast and Islands Regional Educational Laboratory. Retrieved August 25, 2007, from http://www.allinace.borwn.edu/db/ea_catalog.php

Mazur, J. (2004). Conversation analysis for educational technologists: Theoretical and methodological issues for researching the structures, processes and meaning of on-line talk. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1073–1098). New York: McMillan.

Pape, L., Adams, R., & Ribiero, C. (2005). The Virtual High School: Collaboration and online professional development. In Z. Berge & T. Clark (Eds.). *Virtual schools: Planning for success* (pp. 188–132). New York: Teachers College Press.

Paz Dennen, V. (2005). From message posting to learning dialogues: Factors affecting learner participation in asynchronous discussion. *Distance Education, 26*(1), 127 –148.

Rourke, L. Anderson, T., Garrison, R., & Archer, W. (2001). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education, 14*(2). Retrieved August 25, 2007, from http://cade.athabascau.ca/vol14.2/rourke_et_al.html

Swan, K. (2001). Content analysis of course transcripts: Social presence indicators. Retrieved April 1, 2006, from http://www.alnresearch.org/JSP/Resources/instrument_full_text.jsp?index=0 &&article_id=5&

Swan, K. (2003a). Developing social presence in online discussions. In S. Naidu (Ed.), *Learning and teaching with technology: Principles and practices* (pp. 147–164). London and Sterling, VA: Kogan Page,

Swan, K. (2003b). Learning effectiveness: What the research tells us. In J. Bourne & J. C. Moore (Eds.), *Elements of quality online education: Practice and direction* (pp. 13–45). Volume 4, Sloan-C Series. Needham, MA: The Sloan Consortium,

Swan, K. (2004). Relationships between interactions and learning in online environments. Sloan-C Effective Practices, Learning Effectiveness. Retrieved August 25, 2007, from http://www.sloan-c.org/effective/index.asp

Appendix A

Content Analysis of Course Transcripts: Social Presence Indicators

(from Swan 2001)

Category	Indicator	Definition	Examples
affective	expressions of emotion (EM)	Use of descriptive words that indicate feelings such as love, hate, ludicrous, silly, sad	*I just can't stand it . . .* *It is sad that 175 hour requirements have to be mandated*
affective	Value (VL)	expressing personal values, beliefs, & attitudes	*I guess Postman hits a sensitive chord with me. I suspect history will show that what he says about technology is mostly true. We are caught up in in right now, but perhaps for all the wrong reasons.*
affective	paralanguage (PL)	features of written language used outside of formal syntax to convey emotion, such as repetitious punctuation, conspicuous capitalization, emotive spellings, & emoticons	*!!!!!!!!!!!!!!!* *I was reeeeaaaaly tired.* *ANYBODY OUT THERE?* *Thanks :-)*
affective	humor (H)	Use of teasing, cajoling, irony, understatement, sarcasm	*The banana crop in Edmonton (CA) is looking good this year.* *Does he ride a horse without those new-fangled horseshoes?*
affective	self-disclosure (SD)	sharing personal information, expressing vulnerability	*I failed that Classical Literacy test miserably* *Where I work, we . . .* *My daughter spends a lot of time in chat rooms and it worries me*
interactive	acknowledgement (AK)	referring directly to the contents of others' messages, quoting from others' messages	*In your message, you referred to Neil Postman's concept of technopoly . . .*

			I agree with the statements you made about email, "the good and the bad" . . .
interactive	agreement/disagreement (AG)	expressing agreement or disagreement with others' messages	*I couldn't agree with you more.* *I have come to the same conclusion.* *I understand what you are saying but I disagree.*
interactive	appreciation (AP)	offering praise, reinforcement and encouragement to others	*I love your sense of humor.* *What a great answer. I agree.* *You go girl!*
interactive	invitation (IN)	asking questions or otherwise inviting response	*What does everyone think about that show?* *I would love it if somebody else who teaches music could give me some ideas on this.*
cohesive	vocatives (V)	addressing or referring to participants by name	*Jim, ...* *Interesting idea, Mary*
cohesive	greetings & salutations (GS)	greetings, closures	*. . ., Tony* *Hi all!* *That's it for now*
cohesive	group reference (GR)	referring to the group as we, us, or our	-
cohesive	social sharing, phatics (SS)	sharing information not related to the course content, yet with the purpose of enhancing communication	*How about those Yankees!* *When we were in Miami we stayed at . . .* *Am I the only one getting really tired of all this snow?*
cohesive	personal advice (PA)	offering specific advice to classmates	*Also paying attention to the task bar on your desktop can help*
cohesive	course reflection (RF)	reflection on the course as it is progressing	*Now I find myself not only turning it on and using it but taking online classes.*

Appendix B

Network Analysis Statistics

Sess1
Density (matrix average) = 0.1818
Standard deviation = 0.6079
Network Centralization = 15.31%
Heterogeneity = 8.42%. Normalized = 4.06%

Participant	Degree	NrmDegree	Share
1	27.000	18.367	0.188
[Facilitator]	21.000	14.286	0.146
11	9.000	6.122	0.063
2	9.000	6.122	0.063
9	7.000	4.762	0.049
3	7.000	4.762	0.049
16	7.000	4.762	0.049
17	7.000	4.762	0.049
21	6.000	4.082	0.042
18	6.000	4.082	0.042
5	6.000	4.082	0.042
4	4.000	2.721	0.028
12	4.000	2.721	0.028
6	4.000	2.721	0.028
22	4.000	2.721	0.028
10	3.000	2.041	0.021
8	3.000	2.041	0.021
7	3.000	2.041	0.021
14	3.000	2.041	0.021
15	3.000	2.041	0.021
20	1.000	0.680	0.007
19	0.000	0.000	0.000
13			

Sess2
Density (matrix average) = 0.2338
Standard deviation = 0.5954
Network Centralization = 26.40%
Heterogeneity = 6.73%. Normalized = 3.00%

Participant	Degree	NrmDegree	Share
1	40.000	32.000	0.161
[Facilitator]	23.000	18.400	0.093
19	21.000	16.800	0.085
20	16.000	12.800	0.065
7	15.000	12.000	0.060
6	14.000	11.200	0.056
15	11.000	8.800	0.044
25	10.000	8.000	0.040
3	10.000	8.000	0.040
10	9.000	7.200	0.036
11	8.000	6.400	0.032
18	8.000	6.400	0.032
22	8.000	6.400	0.032
4	7.000	5.600	0.028
8	7.000	5.600	0.028
17	6.000	4.800	0.024
9	6.000	4.800	0.024
26	6.000	4.800	0.024
12	5.000	4.000	0.020
5	5.000	4.000	0.020
14	4.000	3.200	0.016
24	4.000	3.200	0.016
16	2.000	1.600	0.008
21	2.000	1.600	0.008
13	1.000	0.800	0.004
2	0.000	0.000	0.000
23			

Sess3
Density (matrix average) = 0.2310
Standard deviation = 0.6856
Network Centralization = 17.16%
Heterogeneity = 9.74%. Normalized = 4.73%

	Degree	NrmDegree	Share
1	39.000	45.882	0.244
[Facilitator]	11.000	12.941	0.069
8	10.000	11.765	0.063
18	10.000	11.765	0.063
6	10.000	11.765	0.063
13	9.000	10.588	0.056
16	9.000	10.588	0.056
2	8.000	9.412	0.050
3	8.000	9.412	0.050
5	8.000	9.412	0.050
11	7.000	8.235	0.044
10	7.000	8.235	0.044
7	7.000	8.235	0.044
15	7.000	8.235	0.044
14	5.000	5.882	0.031
12	5.000	5.882	0.031
9	0.000	0.000	0.000
4	0.000	0.000	0.000
17			

Sess4
Density (matrix average) = 0.3464
Standard deviation = 0.7906
Network Centralization = 39.85%
Heterogeneity = 9.93%. Normalized = 4.63%

	Degree	NrmDegree	Share
1	23.000	21.296	0.189
[Facilitator]	17.000	15.741	0.139
7	12.000	11.111	0.098
19	12.000	11.111	0.098
13	9.000	8.333	0.074
16	8.000	7.407	0.066
9	8.000	7.407	0.066
10	6.000	5.556	0.049
8	5.000	4.630	0.041
17	5.000	4.630	0.041
3	4.000	3.704	0.033
5	3.000	2.778	0.025
11	3.000	2.778	0.025
15	3.000	2.778	0.025
12	2.000	1.852	0.016
14	1.000	0.926	0.008
4	1.000	0.926	0.008
18	0.000	0.000	0.000
2	0.000	0.000	0.000
6			

Appendix C

Transitional Probability Matrices

[Note: Green = statistically significant]

For Sess4+Sess2

Transitional probabilities matrix

	1	2	3	4	5	Replies	No Replies	Givens	Reply Rate
1	**.59**	.18	.06	.12	.06	17	73	88	.17
2	.33	.12	.10	.31	.12	48	26	55	.53
3	.28	**.50**	.06	.11	.06	18	9	23	.61
4	.27	.20	.07	.35	.10	40	34	61	.44
5	.29	.29	.14	.29	.00	14	12	22	.45
	46	30	12	37	12	138	155	251	.17

Z-Scores identify the probabilities that are higher/lower than expected (p value = 2.32)

Z-score = 2.32 at .01 significance
Z-score = 1.64 at .05 significance
Z-score = 1.28 at .10 significance
p108-109 Bakeman & Gottman (1997)

	1	2	3	4	5
1	**2.38**	-0.44	-0.44	-1.50	-0.44
2	0.00	-1.92	0.52	0.86	1.16
3	-0.54	**3.12**	-0.51	-1.61	-0.51
4	-0.93	-0.32	-0.32	1.39	0.35
5	-0.40	0.65	0.78	0.16	-1.22
	46	30	12	37	12

For Sess4+Sess2+Sess3

Transitional probabilities matrix

	1	2	3	4	5	Replies	No Replies	Givens	Reply Rate
1	**.61**	.18	.06	.12	.03	33	104	134	.22
2	.41	.11	.10	.28	.10	61	34	72	.53
3	.24	**.57**	.05	.10	.05	21	11	26	.58
4	.26	.23	.07	.33	.12	43	41	71	.42
5	.31	.25	.19	.25	.00	16	13	25	.48
	66	39	15	41	13	175	204	330	.17

Z-Scores identify the probabilities that are higher/lower than expected (p value = 2.32)

	1	2	3	4	5
1	**3.01**	-0.63	-0.57	-1.70	-1.07
2	0.65	-2.51	0.44	1.01	0.89
3	-1.40	**4.09**	-0.66	-1.60	-0.50
4	-1.89	0.18	-0.43	1.63	1.21
5	-0.56	0.27	1.53	0.16	-1.19
	66	39	15	41	13

For all ATL sessions

Transitional probabilities matrix

	1	2	3	4	5	Replies	No Replies	Givens	Reply Rate
1	**.60**	.15	.10	.10	.05	40	139	176	.21
2	.42	.12	.09	.28	.09	65	41	83	.51
3	.20	**.56**	.04	.08	.12	25	17	35	.51
4	.33	.20	.08	.29	.10	49	48	84	.43
5	.37	.21	.21	.21	.00	19	17	32	.47
	79	42	19	42	16	199	263	412	.16

Z-Scores identify the probabilities that are higher/lower than expected (p value = 2.32)

	1	2	3	4	5
1	**2.94**	-1.06	0.11	-1.93	-0.79
2	0.37	-2.12	-0.11	1.59	0.43
3	-2.15	**4.57**	-1.01	-1.72	0.78
4	-1.16	-0.14	-0.38	1.48	0.64
5	-0.27	-0.01	1.79	-0.01	-1.36
	79	42	19	42	16