Design and Architecture of Collaborative Online Communities: A Quantitative Analysis

Reuven Aviv
Zippy Erlich
Gilad Ravid
Open University of Israel

Abstract

This paper considers four aspects of online communities. Design, mechanisms, architecture, and the constructed knowledge. We hypothesize that different designs of communities drive different mechanisms, which give rise to different architectures, which in turn result in different levels of collaborative knowledge construction. To test this chain of hypotheses, we analyzed the recorded responsiveness data of two online communities of learners having different designs: a formal, structured team, and an informal, non-structured, Q&A forum. The designs are evaluated according to the Social Interdependence Theory of Cooperative Learning. Knowledge construction is assessed through Content Analysis. The architectures are revealed by Statistical Analysis of p Markov Models for the communities. The mechanisms are then identified by matching the predictions of Network Emergence Theories with the observed architectures. The hypotheses are strongly supported. Our analysis shows that the minimal-effort hunt-for-social-capital mechanism controls a major behavior of both communities: negative tendency to respond. Differences in the goals, interdependence and the promotive interaction features of the designs of the two communities lead to the development of different mechanisms: cognition balance and peer pressure in the team, but not in the forum. Exchange mechanism in the forum, but not in the team. In addition, the pre-assigned role of the tutor in the forum gave rise to its responsibility mechanism in that community, but not in team community. These differences in the mechanisms led to the formation of different sets of virtual neighborhoods, which show up macroscopically as differences in the cohesion and the distribution of response power. These differences are associated with the differences in the buildup of knowledge in the two communities. The methods can be extended to other relations in online communities and longitudinal analysis, and for real-time monitoring of online communications.*

Introduction

Building communities is recognized as an essential strategy for online learning. An online community consists of actors who develop certain relations among themselves. For example, some actors only read what others write; some respond to queries posted by others and some influence others to do something (for example to access a web page), etc. This simple observation led us to adopt a network abstraction to describe online communities. A network is a set of actors – members of communities, groups, web-pages, countries, genes, etc., with certain possible relations between pairs of actors. The relations may – or may not – be hierarchical, symmetrical, binary, or other. Network abstraction is thus very flexible.

Social Network Analysis (Wasserman and Faust 1999) is a useful tool for studying relations in a network. It is a collection of graph analysis methods developed by researchers to analyze networks which consist of precise mathematical definitions of certain network structures and the methods to calculate them. Examples of network structures are *cohesiveness* and *transitivity*: cohesiveness measures the tendency to form groups of strongly interconnected actors; transitivity measures the tendency to form transitive triad relations (if i relates to j and j relates to k, then i necessarily also relates to k). SNA has been utilized to analyze networks in various areas, whose actors include politicians (Faust, Willet et al. 2002), the military (Dekker 2002), adolescents (Ellen, Dolcini et al. 2001), multi-national corporations (Athanassiou 1999), families (Widmer and La Farga 1999), and terrorist networks (van Meter 2002). SNA methods were introduced to online communities research in (Garton, Haythornthwaite et al. 1997). Since then several scholars have demonstrated the applicability of SNA to specific collaborative learning situations (Haythornthwaite 1998; Wortham 1999; Lipponen, Rahikainen et al. 2001; Cho, Stefanone et al. 2002; de Laat 2002; Martinez, Dimitriadis et al. 2002; Reffay and Chanier 2002; Aviv 2003).

Macro-level SNA identifies network macro-structures such as *cohesiveness*. Micro level SNA reveals significant underlying microstructures, or *neighborhoods*, such as transitive triads (Pattison and Robbins 2000;

Pattison and Robbins 2002). The identified neighborhoods are the basis for revealing theories that explain their emergence (Contractor, Wasserman et al. 1999). For example, the theory of cognitive balance explains the emergence of transitive triads, which underlies the macroscopic phenomenon of cohesiveness. The precise definition of a neighborhood will be given in section 2.

We examine online communities of learners according to the constructivist perspective (Jonassen, Davidson et al. 1995). Rafaeli (Rafaeli 1988) emphasized that constructive communication is determined by its responsiveness. Accordingly, we analyze the network structures of the responsiveness relation between actors in the online communities. Previous work (Aviv, Erlich et al. 2003) demonstrated that certain macrostructures (cohesion, centrality and role groups) are correlated with the design of the communities and with the quality of the constructed shared knowledge. In this study, we extract the micro-level neighborhoods of the same communities. Our goal is to reveal the underlying theoretical mechanisms that control the dynamics of the communities and to correlate them with the design parameters and with the quality of the knowledge constructed by the communities.

Architecture of a Community

An online community is modeled as a network of actors. Every ordered pair of actors has a potential *response tie relation*. The response tie between actor i and actor j is *realized* if i responded to at least one message sent by j to the community; otherwise the response tie is not realized. In addition, a (non-directed) *viewing relation* is realized between a pair of actors if they read the same messages. In a broadcast community, a realized response tie relation is also a realized viewing tie. The reverse is not necessarily true.

A *virtual neighborhood* (VN) is a sub-set of actors, endowed with a set of prescribed possible response ties between them, all of which are pair-wise statistically dependent. We identified the significant VNs of a community by fitting a p* stochastic Markov model (Wasserman and Pattison 1996; Robins and Pattison 2002) to the response tie data. In this model, every pair of response ties in a VN has a common actor, which is why they are interdependent. Same topology VNs are aggregated into a class of VNs. In the model every possible class is associated with a *strength parameter* that measures the tendency of the network to realize VNs of that class. The basic ideas and the formulas of the p* Markov model are elaborated in (Wasserman and Pattison 1996; Robins and Pattison 2002). The model equations are presented in the Appendix. Examp les of Markov VNs are presented graphically in Figure 1.

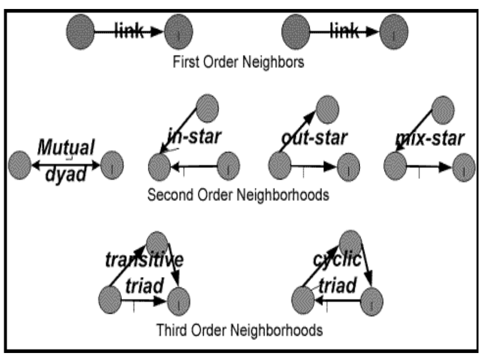

Figure 1. Virtual neighborhoods

In this research we consider the set of Markov classes of VNs listed in Table 1.

VN Class	**Participating Actors & Prescribed Response Ties**
link	All pairs: (i? j) or (j? i)
resp$_i$	All pairs: (i? j) **fixed i**
trigg$_i$	All pairs: (j? i) **fixed i**
mutuality	All pairs: (i? j) and (j? i)
out-stars	All triplets: (i? j) and (i? k)
in-stars	All triplets: (i? j) and (k? j)
mixed-stars	All triplets: (i? j) and (j? k)
transitivity	All triplets: (i? j)and(j? k)and (i? k)

cyclicity	All triplets: (i? j) and (j? k) and (k? i)

Table 1. *Classes of VNs*

Tendencies to form VNs of a certain class are the result of the underlying mechanisms. Several candidate mechanisms, postulated by certain network emergence theories are briefly described below. See (Monge and Contractor 2003) for an extensive survey.

The theory of social capital (Burt 1992; Burt 2002) postulates efficient connectivity in the hunt for a social capital mechanism. In an online *broadcast* community, efficiency means forming zero response ties because a response tie is a redundant viewing tie, so actors prefer to remain passive. This mechanism predicts a tendency for not creating VNs of any class. Thus, other mechanisms are responsible for creating responsiveness.

Exchange and resource dependency theories (Homans 1958; Willer 1999) postulate an information exchange mechanism, in which actors prefer to forge ties with potentially “resource-promising” peers. This mechanism creates tendency for VNs of class *mutuality*.

The theory of generalized exchange (Bearman 1997) postulates an information exchange mechanism via mediators. This theory then predicts tendencies for n-link cycles, in particular VNs from the *cyclicity* class.

Theories of collective action (Marwell and Oliver 1993) postulate a social pressure mechanism that induces actors to contribute to the goal of the community if threshold values of “pressing” peers, existing ties, and central actors are met (Granovetter 1983; Valente 1996). In that case, actors will respond to several others, forging *out-stars* VNs.

Contagion theories (Burt 1987; Contractor and Eisenberg 1990) postulate that the exposure of actors leads to a contagion mechanism that uses social influence and imitation to create groups of equivalent actors with similar behaviors (Carley and Kaufer 1993). Contagion predicts a tendency for VNs of the various *star* shaped classes.

Theories	**Predicted Tendencies**	**Hypotheses**
Social capital	Few single tie links	H1: *link* < 0
Collective action	If thresholds met then respond to several others	H2: if thresholds met then *out-stars* > 0
Exchange	Tendency to reciprocate	H3: *mutuality* > 0
Generalized exchange	Tendency to respond cyclically	H4: *cyclicity* > 0
Contagion	Respond to same as others	H5: *out-stars*> 0; *in-stars* > 0; *mixed-stars* > 0
Cognitive consistency	Respond via several paths	H6: *transitivity* > 0
Uncertainty reduction	Attract many responses	H7: *in-stars* > 0
Exogenous factors: Students	No tendencies to respond/trigger	H8: $\{resp_i = 0 \mid i\ ?\ \text{students}\}$ H9: $\{trigg_i = 0 \mid i\ ?\ \text{students}\}$
Exogenous factors: Tutors	Personal tendencies to respond/trigger	H10: $\{resp_i > 0 \mid i = \text{tutor}\}$ H11: $\{trigg_i > 0 \mid i = \text{tutor}\}$

Table 2*: Research Hypotheses*

Theories of cognitive balance (Cartwright and Harary 1956; Festinger 1957; Harary, Norman et al. 1965) postulate a cognition balance mechanism with a drive to overcome dissonance and achieve cognition consistency among actors. This drive is implemented by *transitivity* VNs.

The uncertainty reduction theory (Berger 1987) postulates drives in actors to forge links with many others to reduce the gap of the unknown between themselves and their environment; this theory predicts a tendency to create *in-stars* (responses to triggering actors) VNs.

Finally, responsibilities of actors influence their residual personal tendencies toward response ties. In this study, students did not have pre-assigned responsibilities, predicting that the students' VNs $resp_i$ and $trigg_i$ will be insignificant. The tutors' residual tendencies will be significant, due to their roles.

The theories, and predicted tendencies stated as Research Hypotheses, are presented in Table 2.

The Analysis

We analyzed recorded transcripts of two online communities – two communities of students at the Open University of Israel. These communities were established for 17 weeks during the Fall 2000 semester (19 participants) and the Spring 2002 semester (18 participants) as part of an academic course in Business Ethics. Each community included one tutor. The designs of the activities of the two communities were different. The

Fall 2000 community was designed as a goal-directed collaborative team, whereas the Spring 2002 community was a Q&A forum. Hence we have labeled the communities "team" and "forum," respectively. The data is available at http://telem.openu.ac.il/courses (password protected).

The *team* community engaged in a formal debate. Participants registered and committed to active participation, with associated rewards in place. Students took the role of an "advisory committee" that had to advise a company on how to handle the business/ethical problem of cellular phone emissions. The debate was scheduled as a 5-step process of moral decision-making, with predefined goals (Geva 2000). A unique feature of the team community was that the goals of the debate were to reach consensus up to the point of writing a joint proposal to an external agency. The *forum* community was open to all students in the course. Participants were asked to raise questions on issues relating to the course. We followed the social interdependence theory of cooperative learning (Johnson and Johnson 1999) to characterize the communities according to four groups of parameters: interdependence, promotive interaction, pre-assigned roles, and reflection. The two communities differ in most of the design parameters. Table 3 summarizes the differences between the designs of the two communities.

Parameter	Team	Forum
Registration & commitment	Yes	No
Interdependence: deliverables	Yes	No
Interdependence: tasks & schedule	Yes	No
Interdependence: resources	Yes	No
Reward mechanism	Yes	No
Interdependence: reward	No	No
Promotive interaction: support & help	Yes	No
Promotive interaction: feedback	Yes	No
Promotive interaction: advocating achievements	No	No
Promotive interaction: monitoring	Yes	No
Pre-assigned roles: tutor	No	Yes
Pre-assigned roles: students	No	No
Reflection procedures	No	No
Individual accountability	Yes	No
Social skills	Yes	Yes

Table 3*: Design of Communities*

Previous analysis (Aviv, Erlich et al. 2003) analyzed the constructed knowledge and the macro-structures of the communities. The analysis revealed that the team community exhibited high levels of constructing knowledge, developed a mesh of interlinked cliques, and that many participants took on bridging and triggering roles while the tutor remained on the side. The forum community did not construct knowledge, cohesion was dull, and only the tutor had a special role. In the team community, many students participate in many cliques, but the tutor belongs to only one clique. In the forum community, only one student and the tutor belong to the two cliques that developed. In addition, participants in the team community shared the role of responders among them, whereas in the forum community only the tutor was a central responder.

The p* model of the team community has 43 classes of virtual neighborhoods, each with its explanatory and parameter: 18 $resp_i$, 18 $trigg_i$, *link*, *mutuality*, *transitivity*, *cyclicity*, and the three *stars*. Similarly, the model of the forum community includes 45 classes of virtual neighborhoods: 19 $resp_i$, 19 $trigg_i$, *link*, *mutuality*, *transitivity*, *cyclicity*, and the three *stars*. The explanatories count the number of virtual neighborhoods that were completely realized in the networks. The strength parameters represent the tendency to create (or not) neighborhoods from the classes.

The analysis of the p* model consists of two steps: In the first step we calculate the explanatories. This was performed using the PREPSTAR program (Anderson, Wasserman et al. 1999). The second step involves solving the binary logistic regression (equation A5). The solution provides an approximate estimate for the strength parameters. This step was performed with SPSS. Details are provided in the Appendix. We configured the SPSS binary logistic procedure to work in forward steps, adding one class of virtual neighborhoods (i.e., its explanatory) at a time, according to its significance, where significance was assessed by the decrements in the PLLD (Pseudo Log-Likelihood Deviance). The analysis stops when no more significant explanatory variables can be identified.

The analysis revealed three significant classes of virtual neighborhoods for the team community, and four significant classes of virtual neighborhoods for the forum community. The PLLD estimates of the strength parameters are presented in Table 4.

Class	θ_K	*SE*	Wald	p	$\exp(\theta_K)$
Team					
link	-3.13	.32	97.5	.000	.043
out-star	.18	.06	9.6	.002	1.199
transitivity	.31	.06	23.9	.000	1.366
Forum					
link	-2.6	.8	10.29	.001	.076
$resp_{18}$	6.1	.12	26.78	.000	456.28
mutuality	6.2	1.38	20.61	.002	519.92
in-stars	-3.2	.91	12.39	.000	.041

Table 4: *Revealed VNs*

In Table 4, θ_K is the MPLE (maximal pseudo-likelihood estimator) for the strength parameter of class K of VNs; *SE* is an estimate of its associated standard error, $\exp(\theta_K)$ measures the increase (or decrease, if θ_K negative) in the conditional odds of creating a response tie between any pair of participants if that response tie completes a new VN of class K.

We tested the hypotheses that $\theta_K = 0$ by the Wald parameter $(\theta_K/SE)^2$ which is assumed to have chi square distribution. Table 3 shows that all these null hypotheses were rejected with extremely small *p* values. The statistical distributions of the MPLEs and the Wald parameters are unknown (Robins and Pattison 2002), so inferences are not precise in the pure statistical sense.

Results

Few classes of VNs are significant: 3 in the Team, 4 in the Forum. In particular, the personal classes of VNs of students, $resp_i$ and $trigg_i$, are *not* significant. This corroborates hypotheses H8 and H9.
The relative importance of the classes of VNs is depicted by their contributions to the goodness of fit of the Markov models. These are presented in Figure 2.

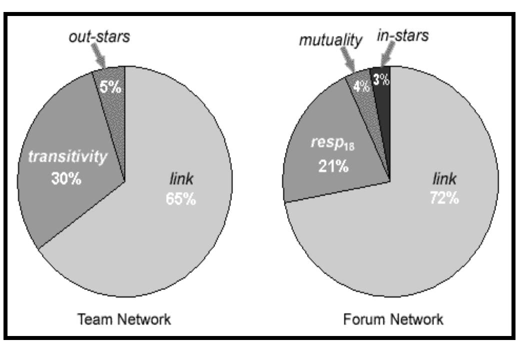

Figure 2. Relative importance of VNs

Figure 2 shows that the global class *link* of the single response tie virtual neighborhoods is the most significant in both communities: 65% and 72% of the goodness of fit are explained by the tendencies associated with this class. Table 5 shows that in both communities the strength parameter θ of the *link* class is negative. This means that the major observed phenomenon in both communities is a significant tendency for not creating single response tie neighborhoods – the phenomenon of lurking. As elaborated above, this can be explained by basic self interest – minimizing the effort required to forge a response tie vs. the possible social capital reward, given that every response tie is a redundant viewing tie. This observation is in accordance with hypothesis H1. Note that the fact that response ties are redundant viewing ties is a feature of every broadcast community, irrespective of the design of the community.

By itself, the negative tendency to create virtual neighborhoods of class *link* will give rise to a community of non-responsive isolates. The actual responsiveness is formed by the other neighborhoods. These neighborhoods are quite different in the two communities. The significant virtual neighborhoods in the team community are from the global classes *transitivity* and *out-stars*. The significant virtual neighborhoods in the forum community are from the personal class resp_{18}, and from the global classes *mutuality* and *in-stars*. In the subsections below, we will consider each of these virtual neighborhoods.

The goodness of fit of the Markov model for the team community increases by 30% when the *transitivity* class of virtual neighborhoods is included. In this community there is a positive tendency to create transitive virtual neighborhoods. This means that in the team community, the likelihood of setting up a response tie from any actor i to any other actor j is enhanced (by 1.37) if that tie completes a transitive triangle virtual neighborhood. This is relative to the likelihood of setting up any other neighborhood. No such preference exists in the forum community. Hypothesis H6 – the tendency for creating virtual neighborhoods of the *transitivity* class is positive – is therefore accepted for the team community but rejected for the forum community.

The tendency to create transitive structures can be explained by cognitive balance theory. It seems that the design of the team community leads to the cognition balance mechanism, by which dissonance between actors and between their perceptions of objects is resolved by balanced paths of communication. This can be attributed to the interdependence built into the design of the community and to the particular goal which forced the participants to reach consensus during the online debate (in order to submit joint proposals). The forum community, on the other hand, was a series of typical Q&A sessions. Here each of the students participating was interested at a certain point in time in a specific issue usually related to an assignment. The scope of the issue was, in many cases, limited; it interested few students. Other issues, or even related concepts not directly connected to the query, were less important to the student who asked the question, let alone to other students. The lifetime of each issue was short (usually until the assignment due date). There was no drive to settle conceptual inconsistencies regarding past issues, or dissonance in perceptions regarding others. Thus, no cognition balance mechanism was needed and none was established.

Introducing the personal class resp_{18} to the model of the forum community increases its goodness of fit by 21%. This class includes all the virtual neighborhoods of single response ties initiated by N18 – the tutor. This means that the residual tendency of N18 to respond – above and beyond the common tendency accounted for by *link* – is significant. Specifically, in the forum community the odds of setting up a response tie (i ? j) increases (by 1,280) if actor i is N18, the main responder in this community. In contrast, the personal class of the tutor's responses in the team community, resp_1, is statistically insignificant. resp_1 neighborhoods are therefore not significant in the explanation of the behavior of the team community. This simply means that the tutor of the team community, P1, showed no tendency to respond. Hypothesis H10 – the tutor’s residual responsiveness is significant – is accepted for the forum community but rejected for the team community.

This difference is attributed to the difference in the role-assignment design of the two communities, which leads to different responsibility mechanisms. The tutor of the forum community was assigned the job of responder. The tutor of the team community was – deliberately – not assigned that role. This results in a difference in their responsibility mechanisms which leads to the difference in their tendency to create the personal class of virtual neighborhoods. A similar observation, mentioned above, is that none of the students in either community showed a significant personal residual tendency to respond, which supports hypothesis H8. This again is attributed to the fact that students in both communities were not controlled by responsibility mechanisms because they were not assigned any particular role. Similarly, in both communities every actor could trigger others by posting a question. No student was pre-assigned the role of trigger. This is reflected in the insignificance of the trigg_i class of neighborhoods (consisting of a single response tie towards actor i), in agreement with hypothesis H9.

We see that the tutors in both communities had no significant tendency to trigger others, contrary to assumption H11. Checking the designs of the two communities, we see that the tutors' behavior was not controlled by responsibility mechanisms, but by other factors. In the forum community, the tutor served only as a helper or responder; no initiation of discussion was designed; accordingly, no triggering role was assigned to the tutor. In the team community, discussion was initiated by the tutor, but the design of the collaborative work dictated that the tutor should step aside. The tutor was therefore not responsible for triggering others.

Incorporating the *out-stars* class increases the goodness of fit of the Markov model for the team community by 5% but has no significance for the forum community. This means that in the team community the likelihood of forging a response tie from any actor i to an actor j is enhanced (by 1.2) if the tie completes an *out-star.* No such tendency is observed in the forum community.

The tendency to create *out-stars*, that is, to forge more than one response tie can be explained by the contagion theory (hypothesis H5) and the theory of collective action (hypothesis H2). The theory of contagion predicts tendencies toward both *in-stars* and *mixed-stars*, but these predictions were not supported by the data of either community. Thus, hypothesis H5 was rejected for both communities. In general, contagion by exposure, as found in friendship relations, is a time consuming process, which presumably could not be developed during the short lifetime of the communities (one semester).

The hypothesis concerning the theory of collective action, H2, was accepted for the team community but rejected for the forum community. This theory assumes the development of peer pressure, provided that the community parameters of density and centrality are above threshold values. This condition is fulfilled for the team community, but not for the forum community, as can be seen in Figures 2 and 3. In general, developing peer pressure is not trivial, as it has to overcome the basic tendency for lurking. In the team community, appropriate initial conditions – commitments, interdependence, and in particular promotive interactions – were set up, and peer pressure was maintained by the tight schedule of common sub-goals imposed on the community. None of these features were designed into the forum community, hence the density and the number of central actors did not reach the thresholds required for peer pressure to work. In the absence of peer pressure, no drive for collective action arouse, which is the reason for the non-significance of the *out-stars* class of virtual neighborhoods in the forum community. The differences between the two communities in the tendencies for *out-stars* is explained quite well by the theory of collective action.

The *mutuality* class of virtual neighborhoods accounts for 4% of the goodness of fit of the Markov model for the forum community. It has no significance for the team community. This means that in the forum community the likelihood of setting up a response tie from any actor i to any actor j is enhanced (by 5,000) if that tie closes a mutual tie. (As stated elsewhere in this paper, the actual number is not precise). This is relative to the likelihood of setting up a tie which is not part of a mutual tie. No such tendency for *mutuality* neighborhoods exists in the team community. Thus, hypothesis H3 is accepted for the forum community but rejected for the team community.

Hypothesis H3 predicts a tendency for *mutuality* virtual neighborhoods on the basis of the exchange mechanism postulated by the theories of exchange and resource dependency. Actors select their partners for response according to their particular resource-promising state. In the forum community the actors prefer to forge response ties (if at all) with partner(s) who usually respond to them – which in this community is the tutor. The tutor is an a priori resource-promising actor as result of her pre-assigned role. This kind of exchange calculus is not developed in the team community because actors in that community cannot identify a priori resource-promising actors. Instead, actors in the team community chose another response policy, governed by the cognition balance mechanism, of responding via transitive triads, as we saw above.

The *in-stars* class of neighborhoods accounts for 3% of the goodness of fit of the Markov model to the forum community but has no significance in the team community. From Table 5 we see that in the forum community *in-stars* is negative. In the forum community, the likelihood of setting up a response tie from i to j decreases if this tie complements an *in-star* neighborhood, that is, if some other actor already has a response tie with j. Contagion theory and the theory of uncertainty reduction both predict a positive tendency for *in-stars* virtual neighborhoods. This prediction is not fulfilled. Hypotheses H5 and H7 are rejected for both communities. As mentioned above, the fact that a contagion process did not develop can probably be attributed to the short lifetime of the communities (one semester). In addition, it seems that there was no need in either community to reduce uncertainties by attracting responses from several sources: in the forum community, the tutor was assigned this role; in the team community, the rules of the game were clearly explained in the document detailing the design of the forum.

We have yet to understand the negative tendency toward *in-stars* virtual neighborhoods in the forum community. This negative tendency means that participants deliberately avoid responding again to the same actor. This phenomenon is explained by the theory of social capital: responding again to an actor is a waste of energy; it decreases the structural autonomy of the responder.

Neither community shows a tendency for *mixed-stars* or *cyclicity* classes of virtual neighborhoods. *mixed-stars* is predicted by the contagion theory, hypothesis H5; the tendency for *cyclicity* is predicted by the theory of generalized exchange, hypothesis H4. Both hypotheses were rejected for both communities. As mentioned above, it is plausible that the contagion mechanism could not develop during the short lifetime of the communities. The theory of generalized exchange relies on knowledge transfer through intermediaries, who seem to be unnecessary in online broadcast communities.

Our findings are summarized in Table 5.

Predicted Hypotheses and Tendencies	**Results**
H1: *link* < 0 Few single tie links	Supported for both communities; feature of every broadcast community independent of design
H2: If large density, centralization, and size, then out-stars > 0 Respond to several others	Supported in team, but not in forum; difference in meeting threshold conditions due to built-in/lack of promotive interactions
H3: *mutuality* > 0 Tendency to reciprocate to resource promising partners	Supported in forum but not in team; difference in existence/non-existence of a priori resource-promising actors due to pre-assigned roles
H4: cyclicity > 0 Tendency to respond cyclically to resource-promising partner	Not supported in either community, probably because there was no need for information exchange via mediators
H5: *out-stars* > 0; *in-stars* > 0; *mixed-stars* > 0; *transitivity* > 0 Respond to same as other equivalent actors	Not supported in either community, probably because contagion process could not develop in the short lifetime of the communities
H6: *transitivity* > 0 Respond via several paths	Supported in team, but not in forum; difference due to difference in consensus reaching requirements and interdependence
H7: *in-stars* > 0 Attract responses from several others	Not supported in either community; uncertainties were clarified by the design (in team) and by the tutor (in forum)
H8: {respi = 0 | i ? students} H9: {triggi = 0 | i ? students} H10: {respi > 0 | i = tutor} H11: {triggi > 0 | i = tutor} Residual personal tendencies to respond or trigger only to actors with pre-assigned roles	H8, H9: Supported for both communities; no pre-assigned role of responders to students H10: Supported in forum, but not in team; differences due to differences in pre-assigned roles of the tutor H11: not supported for either community; no pre-assigned role of triggers to students

Table 5*: Summary of Results*

Discussion

Our analysis shows that the minimal-effort hunt-for-social-capital mechanism, predicted by the theory of social capital & transaction costs controls a large part of the behavior of both communities: a negative tendency to respond. This is a feature of every broadcast community, independent of design.

Differences in the goals, interdependence, and the promotive interaction features of the designs of the two communities lead to the development of different mechanisms: cognition balance, predicted by the balance theory, and peer pressure, predicted by the collective action theory developed in the team community, but not in the forum community. An exchange mechanism developed in the forum community, but not in the team community. In addition, the unique pre-assigned role of the tutor in the forum community gave rise to the responsibility mechanism in that community, but not in the team community. The differences in the mechanisms led to the formation of different sets of virtual neighborhoods, which show up macroscopically as differences in cohesion and in distribution of response power. These differences are associated with the differences in the construction of knowledge in the two communities (Aviv et al., 2003).

It should be noted that the important contagion mechanism did not develop in either community. This mechanism, if developed, would have led to social influence and imitation in attitudes, knowledge, and behavior, which would have developed all kinds of *star* virtual neighborhoods. The required design parameters – promotive interaction – were in place in the team community, but it seems that the lifetime of the community was too short for the development of this mechanism.

There are obvious limitations to the conclusions drawn here. First, we have considered only two communities. In order to capture the commonality, as well as the differences in design, neighborhoods, and mechanisms of online communities, one needs to consider a larger set of communities of different sizes, topics, and, in particular, with different designs. Furthermore, one should consider a set of relations embedded in these communities. One possibly relevant relation between actors is common interest, which can be captured by common keywords in the transcripts and/or common sets of visited web-pages.

Another limitation lies in restricting ourselves to Markov neighborhoods. Pattison and Robbins (Pattison and Robbins 2002) emphasized the possible importance of non-Markovian neighborhoods and brought initial evidence for the empirical value of models that incorporate such neighborhoods. Thus, the dependence structures can, and perhaps should, be treated as a hierarchy of increasingly complex dependence structures.

It seems that SNA can be a useful research tool for revealing community architectures and mechanisms of online communities. There are numerous directions for future research. One direction is "community-covariates interaction." Several studies, such as (Lipponen, Rahikainen et al. 2001), revealed that certain participants take on the roles of influencers (who trigger responses) or of celebrities (who attract responses). Others are isolated – no-one responds to them or is triggered by them. The question is whether this behavior depends on individual attributes or whether this is universal and found across communities. Another direction is "community dynamics," an inquiry into the time development of community structures. When do cliques develop? Are they stable? What network structures determine this behavior? Yet another direction is "large group information overload." It is well known that the dynamics of large groups leads to boundary effects that occur when the group and/or the thread size increases (Jones, Ravid et al. 2002). How are these manifested in online communities?

One practical implication of the methodology used here is the possibility of online monitoring and evaluation of online communities, by embedding SNA tools into community support environments. But a word of caution is necessary: There are various definitions of network structures. Experience shows that applying different definitions may lead to different, even contradictory, results. Further research is needed to determine the stability of network structures under such redefinitions.

Appendix: Key Ideas of the p* Markov Model and the Estimation Procedure

In this research we construct parameterized p* Markov models for the two networks, assuming isomorphism invariance, thereafter extracting the parameters via the MPLE (maximum pseudo-likelihood estimation) procedure (Strauss and Ikeda 1990). Details of the precise formulation of the models, assumptions and the analysis are presented in a series of papers (Wasserman and Pattison 1996; Anderson, Wasserman et al. 1999; Pattison and Robbins 2002). In this section we present the key ideas required for understanding the results and their interpretation.

Any ordered pair of actors in a network has a potential response tie relation. We say that the response tie relation between actor i and actor j is in the realized state if i responded to at least one message sent by j to the network. Otherwise a response tie is in the un-realized state. The state of the network of g actors is then defined by the gXg response matrix $\mathbf{r}$: $r_{ij} = 1$ if a response tie between i and j is realized, otherwise $\mathbf{r}_{ij} = 0$. The states of the response ties are assumed to be the result of stochastic mechanisms operating between pairs of actors. Furthermore, we assume that the probability that the response matrix will actually develop into a state $\mathbf{r}$, $\Pr(\mathbf{r})$, is an exponential function of a linear combination of p state dependent explanatory variables or explanatories, $\{z_1(\mathbf{r}), z_2(\mathbf{r}), \ldots, z_p(\mathbf{r})\}$. Each explanatory $z_i(\mathbf{r})$ has an associated unknown strength parameter θ_i. Estimates of the strength parameters are obtained by fitting the observed states of the response-ties in the forum to the predictions of the probability function $P(\mathbf{r})$.

$$\Pr(\mathbf{r}) = \exp\{\theta_1 z_1(\mathbf{r}) + \theta_2 z_2(\mathbf{r}) + \ldots + \theta_p z_p(\mathbf{r})\} / K(\theta_1, \theta_2, \ldots, \theta_p) \qquad \text{(A1)}$$

The Hamersley-Clifford theorem (Besag, 1974, 1975) states that each explanatory $z_N(\mathbf{r})$ and its associated strength parameter θ_N are associated with one *virtual neighborhood* **N**. A virtual neighborhood is a sub-set of actors and prescribed possible response ties between them, all of which are pair-wise statistically dependent. Actors in a neighborhood may be physically far apart (which is why we call it virtual), but due to certain mechanisms, their possible response ties are all statistically interdependent. Note that the interdependency of the prescribed possible response ties is an inherent property of the virtual neighborhood which in principle is unrelated to the actual realization states of the response ties. A virtual neighborhood may be completely or partially realized, or not realized at all. According to this definition, two possible response ties between pairs of actors in different virtual neighborhoods are statistically independent. The Hamersley-Clifford theorem states that each virtual neighborhood is associated with one explanatory and its strength parameter. The explanatory measures whether the virtual neighborhood is completely realized, in which case it is 1. Otherwise it is zero. The strength parameter quantifies the probabilistic tendency to realize the virtual neighborhood completely.

A subset of a virtual neighborhood is also a virtual neighborhood. Any single pair of actors with a single prescribed possible response tie between them is by definition a virtual neighborhood. Different virtual neighborhoods might have the same set of actors but different prescribed response ties. Holland and Leinhardt (Holland and Leinhardt 1981) considered models in which virtual neighborhoods included only dyads of actors with mutual ties. This implies that dyads are independent, which is an oversimplification. Markov models incorporate a larger variety of various sizes of virtual neighborhoods. In a Markov neighborhood, every two prescribed response ties have one actor in common (which is why they are dependent). Examples of Markov virtual neighborhoods are graphically presented in Figure 1. Markov dependency was introduced by Frank and Strauss (Frank and Strauss 1986). It is a natural assumption in an online community: Forging response ties is an effort, so an actor’s response ties are conceivably interdependent.

The isomorphism invariance (or homogeneity) approximation aggregates same-topology virtual neighborhoods into isomorphism classes, each having one common strength parameter and one explanatory. The explanatory is then a simple counter: It counts the number of virtual neighborhoods of the particular class that are realized in the network. For example, the explanatory associated with the class of transitive triads counts the number of such triads that are realized in the network. The strength parameter quantifies the probabilistic tendency of the network for realizing virtual neighborhoods of the class. In this research we consider the set of Markov isomorphism classes listed in Table A1. The three left-hand columns in the table define the membership of actors in each class and the prescribed possible response ties, the name of the associated strength parameter (which also serves as the name of the class itself), and the formula for deriving the explanatory variables (counters) from the response matrix **r**.

Table A1*: Classes of Virtual Neighborhoods and Explanatories used in Study*

Isomorphism Class of Virtual Neighborhoods: Participating Actors & Prescribed Response ties	Strength Parameter θ	Explanatory z_N(**r**) (counter)	Effects: If $\theta > 0$ is significant ? enhanced tendency to create
All pairs {i, j}| (i? j) or (j? i)	*link*	L(**r**)? **?** $_i$? $_j r_{ij}$	links (either direction)
All pairs {i, j}| (i? j) fixed i	*resp*$_i$	R_i(**r**) = ? $_j r_{ij}$	responses
All pairs {j, i}| (j? i) fixed i	*trigg*$_i$	T_i(**r**) = ? $_j r_{ji}$	triggers
All pairs {i, j}| (i? j) AND (j? i)	*mutuality*	M(**r**) ? **?** $_i$? $_j r_{ij} r_{ji}$	mutual responses
All triplets {i, j, k}| (i? j) AND (i? k)	*out-stars*	OS_2(**r**) ? ?? $_i$? $_j$? $_k r_{ij} r_{ik}$	star-responses
All triplets {i, j, k}| (i? j) AND (k? j)	*in-stars*	IS_2(**r**) ? ?? $_i$? $_j$? $_k r_{ij} r_{kj}$	star-triggers
All triplets {i, j, k}| (i? j) AND (j? k)	*mixed-stars*	MS_2(**r**) ? **?** $_i$? $_j$? $_k r_{ij} r_{jk}$	mixed trigger-responses
All triplets {i, j, k}| (i? j) AND (j? k) AND (i? k)	*transitivity*	TRT(**r**) ? ? $_i$? $_j$? $_k r_{ij} r_{jk} r_{ik}$	transitive triads
All triplets {i, j, k}| (i? j) AND (j? k) AND (k? i)	*cyclicity*	CYT(**r**) ? ? $_i$? $_j$? $_k r_{ij} r_{jk} r_{ki}$	cyclic triads

The probability function then takes the following form:

$$\Pr(\mathbf{r}) = \exp\{\underline{\theta}' \bullet \underline{z}(\mathbf{r})\} / k(\underline{\theta}?) \qquad \text{(A2)}$$

Where the vector of explanatories consists of the counters listed in Table A1:

$$\underline{z}(\mathbf{r}) = \{L(\mathbf{r}), Ri(\mathbf{r}), Ti(\mathbf{r}), M(\mathbf{r}), OS_2(\mathbf{r}), IS_2(\mathbf{r}), MS_2(\mathbf{r}), TRT(\mathbf{r}), CYT(\mathbf{r})\} \qquad \text{(A3)}$$

and the strength parameters measure the tendencies for realizing the virtual neighborhoods of the corresponding Markov classes:

$$\underline{\theta} = \{link, resp_i, trigg_i, mutuality, out\text{-}stars, in\text{-}stars, mixed\text{-}stars, transitivity, cyclicity\} \quad (A4)$$

The L(**r**) explanatory counts the number of neighborhoods of class *link* that were actually realized in the network whose response matrix is **r**. Its strength parameter *link* measures the common tendency to form single response ties; that is, to respond or to trigger. If *link* is negative, it measures the tendency not to form response ties.

The R_i(**r**) explanatory counts the number of neighborhoods of the *resp*$_i$ class that were realized in the network, and its strength parameter. *resp*$_i$ measures the residual tendency (or non-tendency) of actor i to respond, above and beyond the common tendency measured by *link*. Similarly, T_i(**r**) counts the number of neighborhoods of class *trigg*$_i$ that were actually realized, and the *trigg*$_i$ strength parameter measures the residual capability of actor i to attract responses to his/her previous messages; that is, to trigger others to respond, above and beyond the common capability measured by *link*. The M(**r**) explanatory counts the number of realized mutual dyads. The strength parameter *mutuality* measures the global tendency (or non-tendency) of a network to form such dyads. The OS_2(**r**), IS_2(**r**), and MS_2(**r**) variables count the number of realized *star* virtual neighborhoods of the three global classes (see Fig. 3). The corresponding strength parameters measure the tendency (or non-tendency) to forge response ties with (and/or from) two partners. The *transitivity* and *cyclicity* global classes include all triad virtual neighborhoods that are transitive or cyclic, respectively. The associated explanatories, TRT(**r**) and CYT(**r**) count the number of virtual neighborhoods from these classes that were actually realized, and the corresponding strength parameters – *transitivity* and *cyclicity* – measure the tendency to realize virtual neighborhoods of these classes.

It should be emphasized that the explanatories count only completely realized neighborhoods: a virtual neighborhood must have all its prescribed response ties realized in order to be counted.

Wasserman and Pattison (Wasserman and Pattison 1996) reformulated the exponential form of Pr(**r**) into a logit form, which provides both an insight into the precise meaning of "tendency" and a useful procedure for estimating the strength parameters. The logit form of the Markov model is presented in equation A5:

$$w_{ij} = \log [\, \Pr(r_{ij} = 1 | r^c_{ij}) / \Pr(r_{ij} = 0 | r^c_{ij}) \,] = ?_N \theta_N d_N(r^c_{ij}, ij) \quad (A5)$$

The left hand side is the logit – the log of the conditional odds of a pair of actors (i, j) to realize a response tie (i ? j). Here the odds (the ratio between the probability for realizing and not realizing a response tie) is conditioned on all other response tie states, denoted by r^c_{ij}, held fixed. The logit w_{ij} is a linear combination of the changes in the values of the explanatories when the response tie (i ? j) jumps from a not realized to a realized state, when all other response ties, r^c_{ij} are held fixed:

$$d_N(r^c_{ij}, ij) = z_N(r^c_{ij}, r_{ij} = 1) - z_N(r^c_{ij}, r_{ij} = 0) \quad (A6)$$

The change statistic $d_N(r^c_{ij}, ij)$ counts the increase in the number of virtual neighborhoods of class N when the response tie (i ? j) flips from "non-realized" to "realized." It is 1 if (i ? j) completes a whole virtual neighborhood; otherwise it is zero.

The logit form (A5) provides a simple interpretation of the strength parameters. Suppose that an explanatory z_N(**r**) with strength parameter θ is significant. If this happens then the conditional odds for the realization of the response tie (i ? j) from any actor i to any actor j will be enhanced by e^{θ} if this envisaged response tie will make a new virtual neighborhood of class N realized completely. This will happen if the network already has an almost complete realization of the neighborhood: only (i ? j) is missing. Otherwise the conditional odds do not change. Note that if the strength parameter θ is negative, the conditional odds will be decreased. This means that the network has the opposite tendency. For example, if the *transitivity* explanatory is significant, then the conditional odds for forming a response tie (i ? j) is multiplied by $e^{transitivity}$ if the response tie (i ? j) completes a transitive triad (i responds to j, AND j responds to k AND i also responds to k). This will be larger or smaller than 1 depending on the *transitivity* sign.

The logit form (A5) is the basis for one method of estimating the strength-parameters. This method – the MPLE procedure – treats A5 as a binary logistic regression equation: the response tie variable is the dependent variable. There are g(g-1) cases: each ordered pair of actors (i, j) is one case. The values of r_{ij} (1 or 0) for all cases are the observed response ties. The independent variables in the regression equation are the "change statistics" $d_N(r^c_{ij}, ij)$ associated with the explanatories. The coefficients of the change statistics in the regression equation, θ_N, are the unknown strength parameters of the corresponding explanatories.

To solve A5 and estimate the strength parameters, one constructs the pseudo log likelihood function:

$$PL(\underline{\theta}) = ?_{ij} \log [\, Pr(r_{ij} = 1 \mid r^c_{ij}) / Pr(r_{ij} = 0 \mid r^c_{ij}) \,] = ?_{ij} ?_N \theta_N d_N(r^c_{ij}, ij) \qquad \text{(A7)}$$

$PL(\underline{\theta})$ is the log of the product of all the conditional probabilities. It is considered a function of the unknown strength parameters $\underline{\theta}$? ? θ_1, θ_2? ?... θ_p? ?with the response tie states **r** fixed at the observed values. The estimators of the strength parameters are then the values of $\theta_1, \theta_2, \ldots \theta_p$ that maximize $PL(\underline{\theta})$. These are the Maximum Pseudo Likelihood Estimators (MPLEs). The problem with this method is that the statistical distributions of these estimators are not known. One cannot assume that they have the same statistical (chi squared) distributions as their MLE (maximum likelihood estimator) counterparts. Significance intervals based on this assumption can at best be considered defendable approximations, not precise statistical statements. This study attempts to identify the relative strength of the most important explanatories, with no claim to provide precise numerical values for the actual values of their strength parameters.

In this research the actual values of the change-statistics $d_N(r^c_{ij}, ij)$ were calculated from the observed response **r** matrix using PREPSTAR (Anderson et al., 1999). The MPLEs for the strength parameters were then obtained by solving equation (A7) using the binary logistic procedure of SPSS. See (Crouch and Wasserman 1998; Contractor, Wasserman et al. 1999) for examples and details.

Note that once we have estimates for the strength parameters, we can estimate the value of the pseudo log likelihood function, $PL(\underline{\theta})$, itself. This value, to be precise -2* $PL(\underline{\theta})$, can serve as an estimate for the goodness of fit of the model. The best fit is when the product of the conditional probabilities is 1, so that -2* $PL(\underline{\theta})$ is zero. In practice, this is a positive number called Pseudo Log Likelihood Deviance (PLLD) signifying that the model is not perfect. What we are interested in, however, are the decrements in the PLLD caused by introducing more explanatories into the model. A decrement, denoted by $?_?$, measures the contribution of the virtual neighborhood N to the goodness of fit of the model. If one can conjecture that PLLD and $?_?$ have chi square distributions (as do their counterparts in MLE procedure), then one has precise numerical estimates for the relative importance of their contributions to the goodness of fit. As stated above, this assumption is approximate at best. Therefore the decrements $?_?$ serve as a guide to the relative importance of the virtual neighborhoods, but do not provide the range of the true values.

References

Anderson, C. J., S. Wasserman, et al. (1999). "A p* Primer: Logit Model for Social Networks." Social Networks 21: 37-66.

Athanassiou, N. (1999). "International Management Research and Social Networks." Connections 22(2): 12-21.

Aviv, R., Z. Erlich, et al. (2003). "Network Analysis of Knowledge Construction in Asynchronous Learning Networks." Journal of Asynchronous Learning Networks 7(3).

Aviv, R., Z. Erlich, et al. (2003). "Network Analysis of Knowledge Construction in Asynchronous Learning Networks." Journal of Asynchronous Networks 7(3): 1-23.

Aviv, R., Erlich, Z., Ravid, G & Geva, A. (2003). "Network analysis of knowledge construction in asynchronous learning networks." Journal of Asynchronous Learning Networks 7(3): 1-23.

Bearman, P. (1997). "Generalized Exchange." American Journal of Sociology 102: 1383-1415.

Berger, C. R. (1987). Communication Under Uncertainty. Interpersonal Processes: New Directions in Communication Research. E. Roloff and G. R. Miller. Newbury Park, CA, Sage: 39-62.

Burt, R. S. (1987). "Social Contagion and Innovation: Cohesion versus Structural Equivalence." American Journal of Sociology 92: 1287-1335.

Burt, R. S. (1992). Structural Holes. Cambridge, MA, Harvard University Press.

Burt, R. S. (2002). The Social Capital of Structural Holes. New Directions in Economic Sociology. M. F. Guillen, R. Collins, P. England and M. Meyer. New York, NY, Russell Sage Foundation.

Carley, K. M. and D. S. Kaufer (1993). "Semantic Connectivity: An Approach for Analyzing Symbols in Semantic Networks." Communication Theory 3(3): 183-213.

Cartwright, D. and F. Harary (1956). "Structural Balance: A Generalization of Heider's Theory." Psychological Review 63: 277-293.

Cho, H., M. Stefanone, et al. (2002). Social Network Analysis of Information Sharing Networks in a CSCL Community. Proceedings of Computer Support for Collaborative Learning (CSCL) 2002 Conference

Jan. 7-11, Boulder, CO. G. Stahl. Mahwah, NJ, Lawrence Erlbaum: 43-50.
Contractor, N., S. Wasserman, et al. (1999). Testing Multi-Level, Multi-Theoretical Hypotheses about Networks in 21 Century Organizational Forms: An Analytic Framework and Empirical Example.
Contractor, N. S. and E. M. Eisenberg (1990). Communication Networks and New Media in Organizations. Organizations and Communication Technology. J. Fulk and C. W. Steinfeld. Newbury Park, CA, Sage: 143-172.
Crouch, B. and S. Wasserman (1998). "A Practical Guide to Fitting p* Social Network Models via Logistic Regression." Connections 21: 87-101.
de Laat, M. (2002). Network and Content Analysis in an Online Community Discourse. Proceedings of the Computer Support for Collaborative Learning (CSCL 2002) Conference. G. Stahl. Mahwah, NJ, Lawrence Erlbaum: 625-626.
Dekker, A. (2002). "Applying Social Network Analysis Concepts to Military C4ISR Architectures." Connections 24(3): 93-103.
Ellen, J. M., M. M. Dolcini, et al. (2001). "Social Network Centrality and Sexual Experience Among a Household Sample of Urban African-American Adolescents." Connections 24(2): 76-81.
Faust, K., K. Willet, et al. (2002). "Scaling and Statistical Models for Affiliation Networks: Patterns of Participation among Soviet Politicians during the Brezhnev Era." Social Networks 24: 231-259.
Festinger, L. (1957). A Theory of Cognitive Dissonance. Evanston, IL, Row, Preston & Co.
Frank, O. and D. Strauss (1986). "Markov Graphs." Journal of the American Statistical Association 81: 832-842.
Garton, L., C. Haythornthwaite, et al. (1997). "Studying Online Social Networks." Journal of Computer Mediated Communications 3(1).
Geva, A. (2000). "The Internet and the Book: Media and Messages in Teaching Business Ethics." Teaching Business Ethics 4(1): 85-106.
Granovetter, M. (1983). "Threshold Models of Diffusion and Collective Behavior." Journal of Mathematical Sociology 9: 165-179.
Harary, F., R. Z. Norman, et al. (1965). Structural Models: An Introduction to the Theory of Directed Graphs. New York, NY, John Wiley & Sons.
Haythornthwaite, C. (1998). Network Structure among Computer-Supported Distance Learners: Multiple Relations, Multiple Media, and Time.
Holland, P. W. and S. Leinhardt (1981). "An Exponential Family of Probability Distributions for Directed Graphs (with Discussion)." Journal of the American Statistical Association 76: 33-65.
Homans, G. C. (1958). "Social Behavior as Exchange." American Journal of Sociology 19: 22-24.
Johnson, D. W. and R. T. Johnson (1999). Learning Together and Alone: Cooperative, Competitive and Individualistic Learning. Needham Heights, MA, Allyn and Bacon.
Jonassen, D., M. Davidson, et al. (1995). "Constructivism and Computer Mediated Communication in Distance Education." American Journal of Distance Education 9(2): 7-25.
Jones, G. Q., G. Ravid, et al. (2002). An Empirical Exploration of Mass Interaction System Dynamics: Individual Information Overload and Usenet Discourse. 35rd Annual Hawaii International Conference on System Sciences, Big Island, Hawaii, IEEE,.
Lipponen, L., M. Rahikainen, et al. (2001). Analyzing Patterns of Participation and Discourse in Elementary Students' Online Science Discussion. European Perspectives on Computer-Supported Collaborative Learning. Proceedings of the First European Conference on CSCL. P. Dillenbourg, A. Eurelings and K. Hakkarainen. McLuhan Institute, University of Maastricht: 421-428.
Martinez, A., Y. Dimitriadis, et al. (2002). Studying Social Aspects of Computer-Supported Collaboration with a Mixed Evaluation Approach. Proceedings of Computer Support for Collaborative Learning (CSCL 2002) Conference. G. Stahl. Mahwah, NJ, Lawrence Erlbaum: 631-632.
Marwell, G. and P. E. Oliver (1993). The Critical Mass in Collective Action: A Micro Social Theory. Cambridge, MA, Cambridge University Press.
Monge, P. R. and N. Contractor (2003). Theories of Communication Networks. New York, NY, Oxford University Press.
Pattison, P. E. and G. L. Robbins (2000). Neighborhood-Based Models for Social Networks. the International Conference on Social Networks, Vancouver, Canada.
Pattison, P. E. and G. L. Robbins (2002). "Neighborhood-Based Models for Social Networks." Sociological Methodology 22(1): 301-337.
Rafaeli, S. (1988). From New Media to Communication. Sage Annual Review of Communication Research:

Advancing Communication Science. Beverly Hills, CA, Sage: 110-134.
Reffay, C. and T. Chanier (2002). Social Network Analysis Used for Modeling Collaboration in Distance Learning Groups. Lecture Notes in Computer Science (LNCS). S. A. Cerri, G. Guarderes and F. Paraguaco**:** 31-40.
Robins, G. and P. Pattison (2002). Workshop on Exponential Random Graph (p*) Models for Social Networks. Kennedy School of Government, Harvard University.
Strauss, D. and M. Ikeda (1990). "Pseudo-Likelihood Estimation for Social Networks." Journal of the American Statistical Association 85: 204-212.
Valente, T. W. (1996). "Social Network Thresholds in the Diffusion of Innovation." Social Networks 18: 69-89.
van Meter, K. M. (2002). "Terrorists/Liberators: Researching and Dealing with Adversary Social Networks." Connections 24(3): 66-78.
Wasserman, S. and K. Faust (1999). Social Network Analysis: Methods and Applications. Cambridge, UK, Cambridge University Press.
Wasserman, S. and P. Pattison (1996). "Logit Models and Logistic Regression for Social Networks. Part I. An Introduction to Markov Graphs and p*." Psychometrika 60: 401-426.
Widmer, E. D. and L. A. La Farga (1999). "Boundedness and Connectivity of Contemporary Families: A Case Study." Connections 22(2): 30-36.
Willer, D. (1999). Network Exchange Theory. Westport, CN, Praeger.
Wortham, D. W. (1999). Nodal and Matrix Analysis of Communication Patterns in Small Groups. Computer Support for Collaborative Learning (CSCL), Stanford University, Palo Alto, CA, Lawrence Erlbaum.