Author(s): Toikkanen, Tarmo & Lipponen, Lasse

Title: The applicability of social network analysis to the study of networked learning

Year: 2011

Version: Post print

Please cite the original version:

Toikkanen, Tarmo & Lipponen, Lasse. 2011. The applicability of social network analysis to the study of networked learning. Interactive Learning Environments. Volume 19, Issue 4. 365-379. ISSN 1049-4820 (printed). DOI: 10.1080/10494820903281999.

This is an Accepted Manuscript of an article published by Taylor & Francis in Interactive Learning Environments on 12 Oct 2009 available online: http://www.tandfonline.com/10.1080/10494820903281999.

RESEARCH ARTICLE

The Applicability of Social Network Analysis to the Study of Networked Learning

Tarmo Toikkanen[a*] and Lasse Lipponen[b]
[a]Media Lab, University of Art and Design, Helsinki, Finland; [b]Department of Applied Sciences of Education, University of Helsinki, Helsinki, Finland

Studying networked learning (NL) by applying social network analysis (SNA), has gained popularity in recent years. However, it appears that in the context of networked learning the choice of SNA indices is very often dictated by using easily achievable SNA tools. Most studies in this field only involve a single group of students and utilize simple indices, such as density and Freeman's degree centrality. This study uses data collected from 23 groups of pupils and correlates various SNA indices with the pupils' experiences of the learning process, thus identifying SNA indices that actually relate to the experiences of a learning process. The results show that density is not very useful in studying networked learning, and Freeman's degree centrality is meaningful only in certain cases. Further, the study points out several potentially better suited indices for use in further studies of networked learning.

Keywords: networked learning; social network analysis; progressive inquiry; density; Freeman's degree

1 Introduction

The method of social network analysis (SNA) (Scott, 2000; Wasserman & Faust, 1998) has attracted considerable interest in social sciences in recent decades. Lately it has also been adopted in the field of learning sciences, and especially in the context of networked learning (NL) (see examples in Cho, Stefanone & Gay, 2002; Haythornthwaite, 1999; Lipponen, Rahikainen, Lallimo & Hakkarainen, 2003; Martinez, Dimitriadis, Rubia, Gomez, & de la Fuente, 2003; Garton, Haythornthwaite & Wellman, 1997; Nurmela, Lehtinen & Palonen, 1999). In networked learning, social network analysis is applied to collective-level analysis such as studying relationships and interaction processes that go on beyond single participants, demonstrating how information circulates in a networked learning environment. With social network analysis it is, for example, possible to identify key interaction and participation patterns of networked learning.

It appears that in studies of networked learning the choice of SNA indices is very often dictated by what is easily achievable using entry level SNA tools. A typical scenario found in NL studies using SNA consists of a single group of pupils; SNA is used to calculate some basic indices of the scenario, which are then either analyzed descriptively, or with a graphical projection (for examples, see De Laat, Lally, Lipponen & Simons, 2007; Lipponen et al., 2003). No proper multi-group studies have been conducted using SNA (Lipponen, Rahikainen, Hakkarainen, & Palonen, 2002).

The best example is the density index, which is one of the most applied SNA indexes. It is almost invariably included in studies on networked learning. Density is a property of the whole network and describes the general level of linkage among the nodes. It describes the extent to which all participants (or events, if one is analyzing events) of a particular network are interconnected. The density of a network is at a maximum when all the nodes are connected to each other. It is well known that the larger the group, the harder it is to maintain a fully connected network, where everyone is directly interacting with everyone else. Thus there is a strong negative correlation between density and network size: the larger the network, the smaller the density value tends to be. However, this result highlights the true challenge. If the density measure regularly produces these results, how can one, for instance, compare the density values of networks of different sizes, or is it even reasonable to try? And if it is not reasonable to compare the density of networks of different sizes, what is the function of the density index in SNA studies? What would be a goal density value

* Email: tarmo.toikkanen@iki.fi

for a group of a certain size?

In addition to density, Freeman's degree centrality is often used in analyzing interaction patterns in networked learning. Freeman's degree is basically the number of connections to a single actor. Inbound and outbound connections can be differentiated, forming indegree and outdegree indices. Current literature in SNA lists over a dozen variations of centrality indices with their special characteristics (Scott, 2000). The main characteristic of Freeman's degree centrality is that it was the first centrality measure, and easiest to calculate without computers. Other centralities will look at the patterns of actors, and the routes between them, to give more refined measurements of the prominence of a single actor, but these indices are hardly ever used in the field of NL.

SNA is a complex method for analyzing learning interactions, and needs to be applied beyond the descriptive level to analyses. In addition the selection of SNA indices to use in a study should be grounded either on a theoretical basis or on empirical evidence, not on ease of use.

The aim of the paper is to analyze the usefulness of SNA in the study of networked learning. We are especially interested to see which SNA indices would produce meaningful results, and thus, would be useful for researchers of networked learning. To this end, it was sought to understand the following:

(1) Which group level SNA indices are related to meaningful networked learning experience? Specifically, is density a viable index?

(2) Which actor level SNA indices are related to meaningful networked learning experience? Specifically, are Freeman's degree centrality or the betweenness centrality viable indices?

2 Method

In this study, pupils' experiences were collected with a questionnaire and analyzed with factor analysis. These analyses were combined with pupils' networked learning activities analyzed by SNA, thus identifying the SNA indices that actually relate to the meaningful experiences of a learning process.

2.1 Participants and design

The study is part of a larger research project, that aimed to study the implementation and use of pedagogical practices and models of collaborative networked learning in European schools (see Dean and Leinonen, 2003). The present study took place in suburban elementary and secondary schools in an urban district of the city of Helsinki, Finland. The participants of this study were 392 pupils from 17 classes in elementary schools and 99 pupils from 6 classes in secondary schools. Class sizes varied between 7 and 34 pupils, with most classes having between 24 and 30 pupils. In each class, the research period spread over one course, which varied from 2.5 weeks to 15 weeks. although the longest durations were due to holidays. The number of notes written varied from 67 to 363. Table 1 provides some important characteristics of the courses.

Table 1. Characteristics of the 23 analysed courses. Each course has a numeric ID. Active participants include pupils and teachers – some courses had more than one teacher, and some teachers participated in several courses.

Course ID	Active participants	Duration in days	# of notes written	Grade level (1-9; 10-12)
583	31	85	306	6
893	28	21	107	6

1002	27	55	143	10-11
1237	15	51	99	9
1338	13	28	173	10
1518	15	42	67	6
1590	7	39	79	11
1683	22	60	278	9
2214	9	45	167	10
2387	26	83*	258	10
2659	18	90	559	8
3577	19	106*	244	9
3840	20	79	331	3
4314	51	47	363	5
4762	34	30	291	5
5063	26	34	211	5
5279	33	29	220	5
5508	17	15	287	11
5800	26	15	247	9
6054	24	54	193	3
6496	30	50	420	6
6929	13	28	72	6
7003	28	36	176	4
	532	*: includes winter holiday of 14 days	5291	

All courses followed the same general procedure even if their schedules and grade levels did vary. During the course, the teachers of the classes guided and instructed pupils in collaborative networked learning, specifically in the Progressive Inquiry method (Muukkonen, Lakkala & Hakkarainen, 2005) as well as learning with the online Synergeia learning environment. The Progressive Inquiry learning unit consisted of the following four phases. During phase 1, the teacher created a context for the learning process and guided the pupils in generating research questions for the topics being learnt, asking for instance, “Tell me what do you wonder about oceans?” This phase was conducted as a whole-class discussion. During phase 2, pupils were asked to answer their research questions, constructing explanations on the basis of their prior knowledge. In phase 3, pupils were instructed to assess the explanations they had generated or test them empirically. In phase 4, they were tutored to search for new scientific information using books and the Internet for the improvement and refinement of their previous explanations. To facilitate collaborative learning, pupils were encouraged to post their research questions and explanations to Synergeia and to comment there on others' work in order to give and receive assistance and feedback. Thus, the aim was to share all phases of the learning unit from the setting up of research questions to information search through the Synergeia environment. During the course period pupils studied topics such as oceans, aquatic plants and animals, ecology, pressure, and so on. The teachers of the classes received pedagogical training for collaborative networked learning, and technical training for the Synergeia platform. As for the pupils, this was for the most part their first experience with collaborative networked learning and the Synergeia environment.

Synergeia provides teachers and pupils with a shared, structured, web-based work space in which collaborative learning can take place, documents and ideas can be shared, discussions can be stored, and knowledge artifacts can be developed and presented. Teachers and pupils are offered a whiteboard, accompanied by a chat window, to work together simultaneously. With Synergeia, pupils and teaches are allowed to work both asynchronously and synchronously.

2.2 Data

The following sources were used in data collection: we used material from Synergeia log files, and pupils' messages from the Synergeia database. The Synergeia environment archived all conversations for later analysis. In this archive were stored the author of each message, its time of posting, title, content, and position in the conversation thread, and all the times that the message was read by other pupils. 491 pupils were active in the learning environment.

After the course (the total number of courses was 23), the pupils were given a questionnaire. Part II of the questionnaire focused on the actual learning process and the learning environment. (The questionnaire contained many questions related to other issues as well). The questionnaire was returned by 417 pupils, but only 365 cases had complete answers to each question. In 38 cases which only had 1 missing value, that missing value was replaced by the neutral value 3 of the 1-5 Likert scale. Finally, 403 complete answers were used for analysis, and 14 were left out due to incomplete answers.

2.3 Analysis of the data

The analysis of the data had three phases: analyzing the questionnaire, SNA analysis of Synergeia mediated activities, and combined analysis of the questionnaire and SNA.

2.3.1 Analysis of the questionnaire

In order to analyze the questionnaire, 27 Likert-scale items from the second part of the questionnaire were used, and an orthogonal explorative principal axis factor analysis was applied. Each pupil was given factor point values from these factors. The normality of the factors was checked with Shapiro-Wilks's test and visual analysis of Q-Q plots. This is a standard procedure to reduce a large dataset to a few representative data points.

2.3.2 SNA analysis of the Synergeia mediated activities

The database of Synergeia was used to provide information about pupils' interaction with each other during the courses. This information was used to create sociomatrices for social network analysis. While the traditional method of creating sociomatrices is to ask participants about their contacts, this method of creating sociomatrices based on communication data in a virtual environment has been the most common method in previous publications in this field.

In previous studies (Lipponen et al., 2003; Martinez et al, 2003; Nurmela, Palonen, Lehtinen, Hakkarainen, 2003) sociomatrices have been constructed from communication information using varying methods, and no single best solution exists. For this reason, the following matrices were constructed:

- A reply sociomatrix that shows how many times each person has answered other people's messages.
- A read sociomatrix that shows how many messages written by other people each person has read.
- A two-mode affiliation network containing pupils and conversation threads. By calculating both its squares, two sociomatrices were derived: a participant matrix that shows the connections between people who have written into the same conversation threads, and a thread matrix that shows which threads have common participants. The actors of this matrix aren't people but the threads themselves. Neither matrix has directed links, unlike the read and reply matrices.

These four sociomatrices were constructed for each of the 23 courses. From each matrix, several egocentric and graph level indices were calculated (see table 2) by using R (R Development Core Team, 2003) and its SNA library (Butts, 2004). In addition, all matrices were converted into dichotomous matrices (rule: values >0 form a dichotomous connection), because many SNA properties are applicable only to dichotomous matrices (Scott, 2000; Wasserman & Faust, 1998).

Table 2. SNA indices that were calculated for the dataset.

Egocentric / Graph level	Type	Name	Description
Egocentric	Centrality	Freeman's indegree	Sum of incoming links
		Freeman's outdegree	Sum of outgoing links
		Freemans' degree	Sum of incoming and outgoing links
		Closeness	Calculated from the lengths of an actor's all geodesics (closest routes between two actors)
		Freeman's betweenness	Calculated based on the actor being positioned along many geodesics
		Information centrality	Generalization of betweenness, taking into account all routes in the graph, not just the geodesics
		Stress centrality	Count of how many geodesics' path the actor is (simplified version of betweenness)
	Prestige	Degree	Simplest possible prestige measure; same as Freeman's indegree
		Bonacich's power index	Complex measure that calculates prestige by taking into account the prestige of an actor's neighbors.
Graph	Descriptive	Size	Number of nodes in the graph
		Density	Number of different connections compared to highest possible amount (range: 0..1)
		Components	Number of components (groups of interconnected nodes) in the graph that have no connections to each other (range: 1..size)

	Isolates	Number of nodes in the graph that have no connections to other nodes
Centralization	same as for egocentric centralities	Centralization shows the uneven distribution of the corresponding centrality measure (range: 0..1, 0 meaning all actors are equally central, 1 meaning a single actor is central while all others are peripheral)
Krackhardt's hierarchy	Connectedness	Portion of actors that are at least weakly connected to all others
	Efficiency	Amount of superfluous connections in addition to a spanning tree structure
	Lubeness	Do each pair of actors (dyad) have a common information source
	Hierarchy	Portion of non-symmetrics dyads to all non-empty dyads.
Reciprocal	Reciprocity	Portion of symmetric dyads to all non-empty dyads.
	Mutuality	Amount of symmetric dyads.

2.3.3 Combined analysis of the questionnaire and SNA

For the combined analysis of the questionnaire and SNA, course-specific averages from each pupils' factor points were used to measure pupils' experiences about the courses. The SNA indices in turn were used as metrics for the learning interactions that took place in the Synergeia environment during the course. The course-specific factor point averages were then compared with the SNA graph level indices, and the individual pupils' factor points were compared with the egocentric SNA indices.

Spearman's correlation was used to identify potentially significant measures of learning interactions (SNA indices) connected to pupils' experiences of the courses (factor points), both at group and individual levels. All significant correlations were graphed and visually checked, and any correlations that appeared to be caused by a few anomalous pupils were removed, leaving only those that were caused by a general trend.

3 Results

The data from the questionnaire was analyzed with factor analysis and the data from the Synergeia database was analyzed using SNA. These two results were then merged to see what kinds of learning interactions are connected to pupils' experiences about the course.

3.1 Questionnaire results

An orthogonal explorative principal axis factor analysis of the 27 questionnaire items produced 3 factors which mainly included 15 items that upon closer inspection were related to the learning process, teachers' and pupils' activities, and outcomes. The 12 other items were related to questions such as user interface, and practicalities of the course, and were clearly less interesting for this study. Thus the factor analysis was repeated with just those 15 variables. Five factors with eigenvalues of over 1 were found, but the last two factors consisted principally of only one variable, and added little to the result. The final factor analysis was conducted with three factors, which are in table 3. The factors were named on the basis of the variables loaded onto them.

Table 3. Results of the factor analysis. All loadings of over .30 have been highlighted.

No	Question	Meaningfulness of the learning process	Pupils' understanding of the learning process	Teacher's activity
12.	I summarized my thoughts in Synergeia.	**.63**	.18	.09
13.	Synergeia facilitated finding new connections between ideas.	**.62**	.16	.08
8.	Writing my ideas and notes into Synergeia helped my own thinking.	**.62**	.14	.05
9.	I believe other pupils benefitted from being able to read my notes in Synergeia.	**.62**	.16	.02
6.	Writing notes into Synergeia helped me to understand the topics being studied better.	**.54**	.05	.20
10.	Being able to read other pupils' notes from Synergeia helped me.	**.54**	.11	.26
17.	I understood how Progressive Inquiry works while working in Synergeia.	**.52**	**.35**	.09
11.	Being able to read the teacher's notes from Synergeia helped me.	**.50**	.12	.19
7.	I explained my ideas and concepts to other pupils using Synergeia.	**.48**	.18	**.31**
15.	I wrote progressing questions about my research into Synergeia during the project.	**.39**	.09	**.40**
19.	It was easy for me to conceptualize my research project in Synergeia.	**.37**	**.34**	.11
18.	I was able to follow my own research project's progress in Synergeia.	.30	**.95**	.05
14.	The teacher instructed us to write research problems about our work into Synergeia.	.13	.19	**.42**
16.	The teacher encouraged us to put summaries into Synergeia.	.10	-.08	**.56**
26.	The teacher encouraged us pupils to collaborate with each other.	.03	.05	**.61**
		3.28	**1.35**	**1.30**

Many questions loaded onto the factor of 'Meaningfulness of the learning process', but most strongly those which conserned the benefit or help the pupil had received. Questions 18, 17, and 19, which asked about the understanding of the progressive inquiry learning process, were most strongly loaded onto the factor of 'Pupils' understanding of the learning process'. The third factor, 'Teacher's activity', consisted mainly of questions 26, 16, and 14, which asked about the tutoring and guidance given by the teacher.

According to a Shapiro-Wilks test, none of the factors are strictly normal, but on the other hand most normality tests are overly critical with large datasets. Additionally, both the F-test and variance analysis can handle non-normal data: as long as all categories have multiple values, the distribution of the averages is normal even though the variables are not (eg. StatSoft Inc., 2003). In addition a Q-Q plot was used to check that each factor was sufficiently normal, with only a few outliers (see figure 1).

3.2 SNA results

Some SNA indices were not calculable for certain pupils since some calculations require matrix operations that simply cannot be done with certain datasets, for example an inversion of a matrix can result in a division by zero. Information centrality of thread matrices failed for 32 pupils, Bonacich's power index on the dichotomous reply matrix failed for 23 pupils, and betweenness and stress centrality for the participant matrix failed for 58 pupils. These problematic indices were left out of further analysis.

Course sizes varied between 7 and 34 pupils. As detailed in the previous section, several sociomatrices were calculated for each course. Some basic SNA results will be reported briefly, as they are usually mentioned in other NL studies using SNA. Densities of read matrices varied between .27-1 and densities of reply matrices between .082-.72. The size of the course was negatively correlated to the densities (-.71 for read and -.68 for reply matrices), which was to be expected, as it has been commonly shown that density tends to be lower for larger groups (Scott, 2000; Wasserman & Faust, 1998). The read matrices all formed a single component, while the reply matrices formed 1-13 components. Thus all pupils were interconnected by reading each other's posts, but when writing new posts, many courses formed several isolated groups. This indicates that pupils in several courses either through pedagogical design or by themselves formed smaller groups, which monitored others groups' work, but did not comment on them.

3.3 Results from combining SNA and questionnaire results

In this section the results of social network analysis are combined with the factor solution, thus eliciting connections between the pupils' interaction in Synergeia during the course and the experiences they reported after the course.

Of the 403 questionnaire answers and 491 actors in the learning environment, finally 362 had complete data on both datasets and were used in this combined analysis for examining individual actor results. All 23 courses were included, as they were represented by the average factor points and average SNA indices of their respective pupils.

3.3.1 Correlations between factor points and SNA group level indices

The correlations of the factor points and SNA indices of 23 courses are provided in table 2. Correlations only reached .6, which doesn't provide a very high explanation ratio. Any correlations that upon visual inspection were due to a few outliers and not due to a general trend were removed. The first column shows the sociomatrix that was analyzed, and the second column shows the name of the SNA index that was calculated. Only those indices that achieved .05 significance levels or better are shown. It should be remembered that many centralization indices have high cocorrelations among each other.

Table 4: Correlations between factors and SNA group level indices. Correlations that were caused by outliers were removed after a visual inspection of the distributions. Dichotomous matrices are denoted by (D).

Sociomatrix	SNA index	Meaningfulness of the learning process	Pupils' understanding of the learning process	Teacher's activity
Read	Density		-.48*	
	Reciprocity		.56**	
	Efficiency		.47*	
Read (D)	Density		-.46*	
	Indegree		.54**	
	Outdegree		.45*	
	Degree		.54*	
	Betweenness		.65***	
	Stress		.59**	
	Efficiency		.46*	
Reply	Outdegree			.50*
	Reciprocity	-.40*		
	Bonacich	-.53*		
Reply (D)	Components	-.53*		
	Hierarchy	-.41*		
Thread	Density	.50*		
	Degree	-.41*		
	Stress			.46*
	Efficiency			.41*

Note: n = 22, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

3.3.2 Correlations between factor points and SNA egocentric indices

Table 5 is formatted similarly to table 4, but provides egocentric indices and their correlation to individual factor points. The correlations don't surpass .2, which means that on the individual level SNA accounts for only a small portion of the variance in the factors, even though that portion is statistically significant. Only indices that achieved .01 significance or better are shown, as those achieving .05 were too numerous to analyze meaningfully.

Table 5: Correlations between factors and SNA egocentric indices. There were no correlations due to outliers upon visual inspection of the distributions. Dichotomous matrices are denoted by (D).

Sociomatrix	SNA index	Meaningfulness of the learning process	Pupils' understanding of the learning process	Teacher's activity
Read	Degree	.16**		.18***
	Indegree	.15**		.20***
	Outdegree			.16**
	Information			.16**
Read (D)	Degree			.19***
	Outdegree			.15**
	Information			.16**
Reply	Degree	.14**		.18***
	Indegree			
	Outdegree	.14**		
	Information	.14**		
Reply (D)	Degree	.14**		
	Indegree			.17**

Note: n = 362, ** p<0.01, *** p<0.001

4 Conclusions

In this paper, we analyzed the relation between group and actor level SNA indices and meaningful networked learning experiences. Twenty-three groups of elementary and upper secondary school pupils were studied. Their learning interactions in a virtual learning environment were analyzed using Social Network Analysis, SNA. Their experiences about the courses were analyzed with a factor analysis of their questionnaire answers. These results were combined to answer the key question of how SNA should be used in the field of networked learning (NL).

All findings are compared to the findings of previously published SNA studies of NL, highlighting which findings are supported by this study and which are not. As this study included numerous correlation checks between SNA indices and factor point values, significance levels should be treated with some skepticism. All significant correlations have been found with this method, but it is possible that random variance has produced some false positives as well. Many of the indices are strongly cocorrelated, so this problem is not critical, but still these findings need to be verified in a follow-up study.

4.1 Group level indices

The density of a network is the simplest and most often used group level SNA index in the field of networked learning. Often a high density is assumed to be a positive thing. In this study of 23 groups of pupils, density of reading or reply activity had no correlation with the meaningfulness of the course, and was in fact negatively correlated to the pupils' understanding of the learning process. Thus, if the reading density was high (meaning most pupils read messages from most other pupils), their grasp of the Progressive Inquiry learning method was lower. Part of the Progressive Inquiry is for the pupils to form small groups for intensive work, but also to monitor the other groups. It is unclear without a qualitative analysis, whether the groups with high reading densities were more unstructured and chaotic, or if pupils found it easier to just focus on their own group's activities, or if the group sizes were simply too large for proper work. However, density seems to be unrelated to the quality or meaningfulness of a learning process.

Several other group level indices were negatively correlated to the first factor, meaningfulness of the learning process:

Reciprocity: If pupils engaged in one-on-one conversations, replying directly to one another, the value of the conversation suffered. Better conversation was achieved when several pupils participated and replied to others.

Bonacich's power index: This is a centralization index, measuring to what extent the act of writing replies was centralized to a few key actors. A low index, meaning more evenly distributed activity, was beneficial to the conversation.

Components: If the replies were split into several distinct components, the value of conversation suffered. Several components means that pupils formed groups which discussed things by themselves.

Hierarchy: This index estimates how much of a hierarchical structure the replies exhibit. A high hierarchy index could mean that the teacher plays a central role in the conversation, or that pupils mainly discuss among their own small group, with only a few participating in other groups, thus forming a hierarchy. A low hierarchy, meaning freely replying to other people, was beneficial to the learning process.

Thus a knowledge building process as part of a collaborative method, namely Progressive Inquiry, should not be hierarchical, should involve many pupils, its activity should be evenly distributed, and all groups should be interacting with one another.

For the second factor, the best correlate seems to be the betweenness centralization of the dichotomous read matrix. Thus Progressive Inquiry is understood better when some actors become mediators between several groups, passing information along to their own group, while others have a less central role. Other indices, such as stress or degree centralization, show the same general trend that having the interaction distributed unevenly helps in understanding the process.

These two results are in some way conflicting with each other. While a meaningful learning experience would seem to correlate with an even distribution of activity and position, understanding of the process is more difficult. This result may indicate that the method of Progressive Inquiry was itself quite challenging to grasp, or that some courses decided to follow the method more closely while others let the pupils decide their working methods more freely, opting for a more familiar way of working.

In previous studies, Lipponen et al. (2002) analyzed reply activity using network indices of density and Freeman's degree centralization. Martinez et al. (2003) analyzed reading activity using density and degree centralization of four sub-phases of a course. In this study, neither index was related to themeaningful experiences of the learning process, but both were related to the pupils' understanding of the process. However, density was negatively correlated, and degree was only correlated in the dichotomous read matrix, which the studies in question did not use.

4.2 Actor level indices

A commonly used index in NL, Freeman's degree centrality, appears to have a relation to the usefulness of the course. Closer analysis of the different matrices and indices of table 5 suggests that meaningfulness increases when pupils reply to messages of many other pupils, and they read messages from many other pupils. These two activities are of course tightly linked, since in order to reply to a message, it has to be read first. Thus the conclusion could be that the act of writing many messages to several other pupils is the best correlate for meaningfulness of a learning process. In terms of research, this would mean looking at Freeman's outdegree on reply matrixes.

Pupils of active teachers read more widely other pupils' messages. They also received more replies from other pupils. Thus pupils that received active assistance from teachers wrote either high quality notes, or lots of notes, or in some other way succeeded in getting most of the other pupils to reply to them. While their activity was not directly related to their individual success, they were the instigators that encouraged others to write more notes, thus improving their learning experience, and that of the entire class.

In a previous study, Cho et al. (2002) analyzed reply activity with Freeman's indegree and outdegree, as well as Bonacich's power index. Both Lipponen et al. (2002) and Lipponen et al. (2003) analyzed reply activity using Freeman's degree and betweenness. In this study, outdegree and degree had a relation to meaningful learning experiences, while the other indices had no relation.

4.3 Implications

The results of this study are the first empirical data from a large population of pupils (23 groups, n=403) to compare various SNA indices' applicability in the field of networked learning. Previous SNA studies in this field analysing single groups of learners and using SNA indices that find no support in this study may need to be re-evaluated, and the data used in these studies might benefit from a revised analysis using different indices.

The most often used index in NL studies, density, holds no bearing on the quality of the learning process, and as it is strongly correlated to group sizes, using it to compare groups of different sizes is actually misleading. If it has some bearing on collaborative networked learning, its correlation seems to be in fact negative, the reverse of what most previous studies assume.

Another commonly used index, betweenness, measures actor centrality by valuing those positions in the network that serve as conduits between parts of the network. This is intuitively a meaningful valuation to make, but this study showed no correlation between meaningfulness of learning experience and betweenness for reply activity. However, betweenness in reading activity was very significantly correlated with understanding of the learning process, implying that when some pupils acted as messengers between groups, the whole class benefited from it.

The positive findings on Freeman's outdegree and information centrality on reply activity would benefit from a further study to verify them. Studies using Freeman's degree should carefully consider which activity to measure, and whether to use indegree or outdegree. This study found outdegree of replies and indegree of reading to be most strongly correlated with the meaningfulness of the learning experience.

The general findings also provide empirical support for collaborative learning, and the Progressive Inquiry model. Interaction patterns following those recommended in Progressive Inquiry were related to better meaningfulness of learning experience. This result strengthens previous similar finds (Lipponen et al., 2002).

Finally, it should be remembered that these results were derived in a pedagogical setting utilizing the Progressive Inquiry model, and the Synergeia learning environment, and its results may not generalize to other settings directly, but can provide a starting point for analysis. This issue will be pursued in future studies.

Notes on contributors

Tarmo Toikkanen PsM is a researcher in the Learning Environments research group at the Media Lab of the University of Art and Design, Helsinki, and a doctoral student at the University of Helsinki, Department of Psychology. His main research area is the analysis of interaction patterns in collaborative learning scenarios.

Lasse Lipponen Ph.D is a university lecturer in educational psychology and works at the University of Helsinki, in the Department of Applied Sciences of Education. His main research area is collaborative and technology mediated practices of learning and working.

Bibliography

Butts, C. T. (2004). *Tools for Social Network Analysis.* Retrieved 15.10.2004.

Cho, H., Stefanone, M., Gay, G. (2002). Social information sharing in a CSCL community. In G. Stahl (Ed.), *Proceedings of the 2002 CSCL*, 43–50

Dean, P. and Leinonen, T. (2003). *ITCOLE final report*. Technical Report IST-2000-26249, European Commission, ITCOLE Project.

De Laat, M., Lally, V., Lipponen, L., & Simons, R-J. (2007).Online teaching in networked learning communities: A multi-method approach to studying the role of the teacher Journal. *Instructional Science*, 35, 257–286.

Garton, L. Haythornthwaite, C., & Wellman, B. (1997). Studying online social networks. *Journal of Computer-Mediated Communication*, 3, 1–26.

Haythornthwaite, C. (1999). Networks of information sharing among computer-supported distance learners. *Proceedings of the 1999 CSCL,* 218–222, Stanford, CA.

Lipponen, L., Rahikainen, M., Hakkarainen, K. & Palonen, T. (2002). Effective participation and discourse through a computer network: Investigating elementary students' computer supported interaction. *Journal of Educational Computing Research*, 27(4), 355–384.

Lipponen, L., Rahikainen, M., Lallimo, J., and Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction*, 13:487–509.

Martinez, A., Dimitriadis, Y., Rubia, B., Gomez, E., and de la Fuente, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions. *Computers & Education*, 41:353–368.

Muukkonen, H., Lakkala, M., & Hakkarainen, K. (2005). Technology-mediation and tutoring: how do they shape progressive inquiry discourse? *The Journal of the Learning Sciences*, 14(4), 527–565.

Nurmela, K., Lehtinen, E., & Palonen, T. (1999). Evaluating CSCL Log Files by Social Network Analysis. *Proceedings of The 1999 CSCL*, 434–444. December 12-15, 1999. Palo Alto, California.

Nurmela, K., Palonen, T., Lehtinen, E., and Hakkarainen, K. (2003). Developing tools for analyzing cscl process. *In CSCL Conference*. University of Turku and University of Helsinki.

R Development Core Team (2003). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3.

Scott, J. (2000). *Social Network Analysis: a handbook*. SAGE Publications, London, UK.

StatSoft Inc. (2003). *Electronic Statistics Textbook*. Retrieved 7.11.2004.

Wasserman, S. & Faust, K. (1998). *Social Network Analysis: Methods and Applications*. Cambridge University Press.