THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

What is the source of social capital? the association between social network position and social presence in Communities of Inquiry

What is the Source of Social Capital?

The Association Between Social Network Position and Social Presence in Communities of Inquiry

Vitomir Kovanovic*
School of Interactive Arts and Technology
Simon Fraser University
250 - 13450 102nd Avenue
Surrey, BC, V3T0A3 Canada
vitomir_kovanovic@sfu.ca

Srecko Joksimovic
School of Interactive Arts and Technology
Simon Fraser University
250 - 13450 102nd Avenue
Surrey, BC, V3T0A3 Canada
sjoksimo@sfu.ca

Dragan Gasevic
School of Computing Science
Athabasca University
1 University Drive
Athabasca, AB, T9S 3A3
Canada
dgasevic@acm.org

Marek Hatala
School of Interactive Arts and Technology
Simon Fraser University
250 - 13450 102nd Avenue
Surrey, BC, V3T0A3 Canada
mhatala@sfu.ca

ABSTRACT

It is widely accepted that the social capital of students – developed through their participation in learning communities – has a significant impact on many aspects of the students' learning outcomes, such as academic performance, persistence, retention, program satisfaction and sense of community. However, the underlying social processes that contribute to the development of social capital are not well understood. By using the well-known Community of Inquiry (CoI) model of distance and online education, we looked into the nature of the underlying social processes, and how they relate to the development of the students' social capital. The results of our study indicate that the affective, cohesive and interactive facets of social presence significantly predict the network centrality measures commonly used for measurement of social capital.

General Terms

Social Network Analysis, Community of Inquiry, Social Presence

1. INTRODUCTION

Asynchronous online discussions have been frequently used both in blended and fully online learning [41]. However, with the broader adoption of social-constructivist pedagogies and the shift towards the collaborative learning [2], they are viewed as one of the important study tools for the computer-supported collaborative learning (CSCL) within the online learning environments. Their use has produced an enormous amount of data about the interactions between students and instructors [21]. The distance education and CSCL research communities have tried to use these data for gaining insights into the very complex nature of the learning phenomena. Among the different ways of researching students' social interactions *Quantitative Content Analysis* (QCA) [38, 19] and *Social Network Analysis* (SNA) [52, 46] represent two commonly used methods.

A widely accepted model of distance education which makes a use of QCA is the *Community of Inquiry* (CoI) model [28]. According to Garrison and Arbaugh [30], it is one of the leading models of distance education that describes the key constructs of the overall educational experience. The CoI model provides the in-depth assessment of teaching, cognitive and social dimensions of learning phenomena, and how those three dimensions affect: i) the overall success of the learning process, and ii) the attainment of learning objectives [28]. Empirical research showed that the social dimension of learning plays an important role in the learning communities by mediating the relationship between the teaching and cognitive dimensions [31]. Still, the CoI model does not explicitly address the question of student social networks, their structure, or the effects they have on the overall educational experience and learning outcomes. Given the amount of evidence from the studies of student social networks [46], this warrants further investigation.

One of the central aspects in the study of social networks is the idea of the *social capital* [13, 12]. Generally speaking, social capital can be defined as a value resulting from occupying a particularly advantageous position within a social network [12]. Over the years, the study of social capital has become increasingly popular in the field of education [14]. The large number of studies in the distance education field indicated an important connection between the students' social capital and many important aspects of education and learning including academic performance [33, 15, 7, 49, 43], retention [23], persistence [50], program satisfaction [7], and sense of community [17]. Still, research of the student social networks have involved mostly isolated studies that were focused on the understanding of the relationship between a particular set of constructs selected by the researchers and the students' network position. Likewise, the underlying mechanisms responsible for the observed social structure are typically not addressed, which is understandable given the lack of educational theories that explicitly

*Corresponding Author

take into the consideration student social networks.

In this paper, we present the results of the study which explored the links between the CoI model and the social network analysis of student networks. With the current advancement within the CoI research and most recent validations of the model [31], the model is mature enough and empirically sound to provide this missing theoretical foundation for understanding the structure of students' social networks. Likewise, the understanding of the structure of social networks can provide a more comprehensive overview of the social dimension of learning that it is already accounted for in the research of the CoI model.

Given the exploratory nature of this study, we focused on the relationship between social capital and social processes which are indicative of the student social presence development. The main question we aim to answer, in this paper is *which social processes, and to what extent, are indicative of the development of the social capital in a communities of inquiry*? Given the detailed characterization of social aspects of learning in the CoI model through the construct of *social presence*, we explored how this construct relates to the students' social capital, as characterized by their position in social networks formed around communities of inquiry. As the community of inquiry provides characterization of different sociological processes that constitute social presence, we looked how each of them contributed to the development of social capital withing students' social network.

2. THEORETICAL BACKGROUND

2.1 Social network analysis

2.1.1 Social capital

The study of social networks has attracted much attention in social and behavioral sciences [17, 14]. The focus in social network analysis is on the study of *relationships*, also known as *ties*, between a set of *actors*, or *participants* [14]. Through the relationships, members of a network engage in sharing, exchange or delivery of various resources including information [36]. Social network analysis draws much of its ideas from the mathematical graph theory and the sociometric studies of the human relationships [52].

An important concept in the study of social networks is the idea of *relation strength* [34], which is used to make a distinction between *strong social ties*, which require a substantial commitment (e.g., family, close friends), and *weak social ties* which do not obligate a strong commitment (e.g., acquaintances). Likewise, the idea of *network brokerage* builds on the fact that in a large network, the density of relationships is not uniform, which indicates the existence of smaller sub-communities within a large social network [12, 13]. In his seminal paper, Granovetter [34] stressed the tremendous importance of weak social ties, as they provide access to novel information from different parts of a social network and provide pathways of information exchange between sub-communities. An individual who possesses a large number of weak ties in many different sub-communities is able to take advantage by combining diverse information coming from different sub-communities, and to even control to a certain degree the spread of information from one sub-community to another [12]. This ability to create a value from occupying a particular position in a social network is known as *social capital* [13]. To study and assess values of different network positions, the principles of graph theory are the most commonly used [52]. The notion of *centrality* is particularly important. This notion captures the relative importance of individuals in social networks [52]. Given the complexity of measuring actors' relative importance, a large number of centrality measures were proposed over the years out of which degree, closeness and betweenness centralities are the most frequently used [26].

2.1.2 Social network analysis in education

While social network analysis has been widely adopted in social and behavioral sciences, its adoption in the field of education was initially very limited [14]. According to Carolan [14], the main reasons for this are *"overemphasis on individual explanations of educational opportunities and outcomes, a quest for scientific legitimacy, and a preference for experimental designs that estimate the causal effects of 'educational interventions' "* [14, 32]. Nevertheless, over the years, the number of studies that indicated the importance of social connections on the overall academic experience has grown considerably. A good example is the study of students' overall academic experience from early 1990s by Astin [5] in which he concluded that: i) the environment made by the instructors and students is crucial, and ii) the single most important environmental influence is *peer group*.

In the context of distance education, there have been many studies recently that looked at the connection between several important learning constructs and social capital of students. Likewise, in the fields of educational data mining (EDM) [6] and learning analytics [40], the interest in SNA has been growing. The recent review of the EDM field by Romero and Ventura [44] noted a growing interest in SNA; likewise, in the learning analytics community, SNA was recognized as one of the most important techniques of social learning analytics [11, 25].

As expected, academic performance was the focus of a large majority of the studies [33, 50, 15, 7, 49, 43] that have found positive effects of student positions in social networks on academic performance. Still, academic performance was not the only construct that was examined. The study of retention by Eckles and Stradley [23] found that for each friend that leaves an academic degree program makes a student five times more likely to leave as well, while every friend who stays makes a student 2.25 times more likely to also stay in college. The study of student persistence and integration by Thomas [50] found that students with a broader set of acquaintances are more likely to persist in the academic program of a higher education institution, and that students with a higher proportion of ties outside their peer group also perform better academically. This is aligned with the findings of Dawson [17] who showed that students' sense of community membership was positively related to their closeness and degree centrality measures. Similarly, in the study of a team-based MBA program by Baldwin et al. [7], it was found that the high embeddedness in the friendship network increased students' perception of learning and enjoyment in the program; as well, the centrality in the communication networks was found to be positively linked with the student grades.

One important thing to notice is that the majority of the studies did not draw their theoretical foundations of network formation from the established educational theories. As pointed out by Rizzuto et al. [43], there is a lack of *"theory of academic performance that combines individual characteristics as well as social and infrastructural factors"* (p180). The main exception is the use of retention theories by Tinto [51] and Bean [8] in the study of student persistence and retention. The other notable theories that are adopted, such as Feld's theory of focused choice [24], or Lin's theory of social resources [39] are general sociological theories that do not take into the account the specific of learning processes and educational contexts.

2.2 The community of inquiry (CoI) model

2.2.1 Overview

The Community of Inquiry (CoI) model is a general model of distance education which explains the constructs that contribute to the overall learning experience. It is rooted in the social constructivist philosophy, most notably in the work of John Dewey [20], and is particularly well suited for understanding different aspects of learning within the learning communities. The main goal of the CoI model was to define the constructs that characterize a worthwhile educational experience, and a methodology for their assessment. The CoI model consists of the three interdependent constructs, also known as *presences*, that together provide a comprehensive coverage of the distance learning phenomena:

1) **Cognitive Presence** explains different phases of students' knowledge construction process through social interactions within a learning community [28].
2) **Teaching Presence** describes the instructor's role in course delivery and during course design and preparation [3].
3) **Social Presence** explains the social relationships and the social climate within a learning community that have a significant effect on the success and quality of social learning [45].

The CoI model is well-researched and widely accepted within the distance learning research community as shown by a recent two-part special issue of The Internet and Higher Education journal [1]. The model defines its own coding schemes that are used to assess the levels of the three presences through the QCA in transcripts of asynchronous online discussions. More recently, instead of relying on the QCA, a CoI survey instrument [4] was developed as an alternative way of assessing the levels of the three presences.

2.2.2 Social presence

Social presence is defined as the *"ability of participants in a community of inquiry to project themselves socially and emotionally, as "real" people (i.e., their full personality), through the medium of communication being used"* [28, p3]. Critical thinking, social construction of knowledge and the development of the cognitive presence are more easily developed in the cases where the appropriate levels of social presence have been established [28].

Given the form of delivery in distance education, face-to-face communication that is typical for more traditional forms of education delivery is not possible. Hence, establishing and sustaining social presence is more challenging. Distance education was often criticized as being inferior to more traditional forms of education, particularly because of the inability to create social presence between the members of a learning community [2]. However, according to Garrison et al. [28], the form of communication is not the solely factor determining the development of social presence. A key aspect of establishing social presence in face-to-face settings are visual cues, while participants in online communities use different techniques – such as emoticons – to convey the affective dimension of communication that lacks in typical text-based communications.

As described by Rourke et al. [45], the origins of social presence can be found in the work of Mehrabian [42] and his notion of *immediacy* which is defined as *"the extent to which communication behaviors enhance closeness to and nonverbal interaction with another"* [42, p203]. This, and the set of follow-up studies by communication theorists, defined the theoretical background on which the construct of social presence was based [45]. The social presence in the CoI model is defined as consisting of three different dimension of communication:

1) **Affectivity and expression of emotions**: Since emotions are strongly associated with motivation and persistence, they are indirectly connected to critical thinking and communities of inquiry. More formally, emotional expression has been indicated by the *"ability and confidence to express feelings related to the educational experience"* [28, p99].
2) **Interactivity and open communication**: In order to promote the development of higher-order critical thinking skills, the notion that the other side is listening and attending is crucial [45]. Thus, activities such as praising of the student work, actions, or comments contribute to the teacher immediacy, which in turn leads to affective, behavioral and cognitive learning [45]. Similarly, open communication is defined as *"reciprocal and respectful exchanges of messages"* [28, p100] and together with interactivity provide a basis on which productive social learning can be established.
3) **Cohesiveness**: The activities that *"build and sustain a sense of group commitment"* [28, p101] define cohesiveness. The goal is to create a group where the members possess strong bonds to both i) each other and ii) the group as a whole. This in turn stimulates productive learning and the development of critical thinking skills.

Given that there are three different dimensions of social presence, the coding scheme for social presence (see Table 1) defines a list of indicators for each dimension. By looking at the content and the timing of each message, it is possible to see how the social climate unfolded during the course delivery. This provides a way of understanding and evaluating the different pedagogical interventions with respect to the development of a productive social climate in a learning community which enables for the meaningful social interactions [53].

2.3 Research Question: Characterization of social capital through social presence

As indicated in the previous sections, there is a strong evidence that social capital plays an important role in the shaping of the overall learning experience. The main research question that we investigate in this paper:

> What is the relationship between the students' *social capital*, as captured by social network centrality measures, and students' *social presence*, as defined by the three categories in the Community of Inquiry model?

The higher the social capital of a learner is, the more capable the learner is in terms of learning opportunities, information exchange, or integration within the academic environment. Still, the origins of social capital are not fully understood. Why certain students occupy advantageous positions in social networks? What are the social processes that enable them to take advantage of their social relationships? As for now, not a single theory of learning addresses the question of social capital directly, even though the impact of social context on learning is widely acknowledged.

As indicated by the previous study by de Laat et al. [18], content analysis techniques can be used in combination with SNA to provide a more comprehensive view of the social learning processes. In this paper, we propose the use of the Community of Inquiry model, given its holistic view of educational experience and extensive empirical evaluation by the research community [29], with the aim to characterize the origins of social capital in communities of inquiry. The CoI model description of important behavioral indices that contribute to the development of the positive social climate could be used to interpret the observed differences among students positions in a social network.

Likewise, the synergistic effect of using those two perspectives on student interactions provide a value for the CoI model by emphasizing the effects of the theorized social processes. For example, are interactivity and open communication important for the development of social capital? Are the students who show group cohesion the ones who take brokerage positions? Recently, there have been

Table 1: Social Presence Categories and Indicators as defined by Rourke et al. [45]

Category	Code	Name	Definition
Affective	A1	Expression of emotions	Conventional expressions of emotion, or unconventional expression of emotion, includes repetitions punctuation, conspicuous capitalization, emoticons.
	A2	Use of humor	Teasing, cajoling, irony, understatements, sarcasm.
	A3	Self-disclosure	Presenting details of life outside of class, or express vulnerability.
Interactive or Open Communication	I1	Continuing a thread	Using reply feature of software rather than starting a new thread.
	I2	Quoting from others' messages	Using software features to quote others entire messages or cutting and pasting selections of others' messages.
	I3	Referring explicitly to others' messages	Direct references to contents of others' posts
	I4	Asking questions	Students ask questions of other students or the moderator.
	I5	Complementing, expressing appreciation	Complimenting others or contents of others' messages.
	I6	Expressing agreement	Expressing agreement with others or content of others' messages.
Cohesive	C1	Vocatives	Addressing or referring to participants by name.
	C2	Addresses or refers to the group using inclusive pronouns	Addresses the group as *we,us, our, group*.
	C3	Phatics, salutations	Communication that serves a purely social function: greetings, closures.

some attempts [47, 48] that make use of SNA in conjunction with the CoI model to provide insights into particular aspects of learning, such as self-regulation [9]. Still, the central question of social capital is left unexplored and that is the goal in our study.

3. METHODS

3.1 Dataset

For our study, we used the dataset consisting of six offers (Winter 2008, Fall 2008, Summer 2009, Fall 2009, Winter 2010, Winter 2011) of the masters level software-engineering course offered through the fully online instructional condition at a Canadian open public university. The course is 13 weeks long, research-intensive, and focuses on understanding of current research trends and challenges in the area of software engineering. Students were requested: i) to participate in online discussions for which they received 15% of their final grade (see details in [32]), and ii) to work on a four tutor marked assignments. Overall, 81 student created the total of 1747 discussion messages which were then used as the main data source for this study. The total number of students and messages for all six course offerings are shown in Table 2.

Table 2: Course offering statistics

	Student count	Message count	Graph density
Winter 2008	15	212	0.52
Fall 2008	22	633	0.69
Summer 2009	10	243	0.84
Fall 2009	7	63	0.58
Winter 2010	14	359	0.84
Winter 2011	13	237	0.77
Average	13	291	0.71
Total	81	1747	

Table 3: Descriptive statistics of social network metrics

	$Mean$	SD	Min	Max
Betweenness	9.04	14.51	0.00	74.20
In-degree	19.84	8.62	4.00	42.00
Out-degree	19.86	9.37	3.00	44.00
In-closeness	0.09	0.04	0.04	0.17
Out-closeness	0.08	0.04	0.03	0.18

3.2 Social network measures

In order to measure students' social capital we extracted student social network graphs from the interactions on the discussion boards. We extracted *directed* social graphs, so that whenever a student $X1$ responded to a message from another student $X2$, we created a direct relationship between the two of them ($X1 \Rightarrow X2$). Since two students can exchange more than one message, we extracted a *weighted* graph where the weights corresponded to the number of exchanges between a given pair of students. We created a separate social graph for each of the course offerings independently and the graph densities for each offering are shown in Table 2.

From the constructed social network graphs, we extracted the three network centrality measures which are most frequently used for the study of the educational social networks [14]:

1) **Betweenness centrality** captures brokerage opportunities of actors in a network and is the most directly related to the social capital construct [13, 12]. For a given actor A, it is mathematically defined as the number of shortest paths between any two other actors that "pass through" the actor A [26].
2) **Degree centrality** measures the total number of relationships that each participant has [26]. Given that we constructed the directed social graphs, we considered separately the in-degree and out-degree centrality measures. They represent the total number of incoming and outgoing relations for a given individual, respectively. Degree is the simplest centrality measure, very easy to calculate, as it takes into account only the direct relationships between the actors [52].
3) **Closeness centrality** represents the distance of an individual participant in the network from all the other network participants [26]. It is defined as the inverse of the sum of the distances to all other participants [14], and hence takes into account both direct and indirect relationships [52]. Much like degree centrality, given that the student graphs are directed, we calculated the in-closeness and the out-closeness centrality measures. For a given actor A, in-closeness centrality measures how many indirect steps are needed for all other actors to reach the actor A, while out-closeness measures how many indirect steps the actor A requires in order to reach all the other actors in the network.

Table 3 shows the descriptive statistics for all five extracted centrality measures. We can see that on average the students wrote around 20 messages, and also received on average around 20 responses. This level of activity was expected, as by the course design the students were expected to spend a significant amount of time on the online discussions. Still, from the descriptive statistics reported in Table 3, we can observe the large differences between the individual students in the case of all five centrality measures.

3.3 Message coding

In order to assess students' social presence, all messages were manually coded by two coders in accordance with the coding scheme defined by Rourke et al. [45]. As the individual messages can

Table 4: Social Presence Indicators

Category	Code	Indicator	Count	Percent Agreement
Affective	A1	Expression of emotions	288 (16.5%)	84.4
	A2	Use of humor	44 (2.52%)	93.1
	A3	Self-disclosure	322 (18.4%)	84.1
Interactive	I1	Continuing a thread	1664 (95.2%)	98.9
	I2	Quoting from others messages	65 (3.72%)	95.4
	I3	Referring explicitly to other's messages	91 (5.21%)	92.7
	I4	Asking questions	800 (45.8%)	89.4
	I5	Complementing, expressing appreciation	1391 (79.6%)	90.7
	I6	Expressing agreement	243 (13.9%)	96.6
Cohesive	C1	Vocatives	1433 (82%)	91.8
	C2	Addresses or refers to the group using inclusive pronouns	144 (8.24%)	88.8
	C3	Phatics, salutations	1281 (73.3%)	96.1

Table 5: Social Presence Categories.

Category	Count	Percent Agreement
Affective	530 (30.3%)	80.8
Interactive (Excluded I1 and I5)	1030 (59%)	86.2
Cohesive (Excluded C1)	1326 (75.9%)	93.4

be simultaneously classified into more than one category of social presence, each message was coded with three binary codes indicating whether the message belongs to a particular social presence category. However, early in the coding process, we observed an extremely high frequency of some of the indicators in the cohesive and interactive categories. Because of this, almost all of the messages could be classified as both interactive and cohesive, which would limit the discriminatory power of those two categories. Thus, to resolve this issue, instead of coding on the levels of categories, the coding was done on the levels of the individual indicators, so that each message was coded with the twelve binary codes (i.e., three indicators of the affective category, six indicators of the interactive category and three indicators of the cohesive category) each indicating an occurrence of a particular social presence indicator within a given message. This enabled us to look at the distribution of the individual indicators and to be more selective in the type of the indicators that we wanted to investigate. Overall, the coding agreement was high, with all of the indicators reaching percent agreement of at least 84%, and all the coding disagreements were resolved through discussion between the coders in a follow-up meeting, after they first coded the messages independently. The coding results are shown in Table 4. The results show that some of the indicators were recorded in a disproportionately large number of messages. Thus, in order to evaluate different aspects of social presence captured by those three categories, we omitted some of the indicators from our analysis: i) Continuing a thread, ii) Complementing, expressing appreciation, and iii) Vocatives. We intentionally kept the "Phatics, salutations indicator" as its removal would render the cohesive category in only 8.24% of the messages. By using the remaining nine indicators, we categories all of the messages in the corpus, and the final results are shown in Table 5.

3.4 Statistical analysis

In order to investigate the relationships between the three categories of social presence, as defined by the CoI model, and *social capital*, as operationalized through the five network centrality measures, we conducted backward-stepwise multiple linear regression analyses [35] for each of the five extracted network centrality measures. To evaluate different regression models for a particular centrality measure, we used the popular Akaike Information Criterion (AIC) [35]. In order to control for the inflation of the Type-I error rate due to multiple statistical significance testing, we used the Holm-Bonferroni correction [37], also known as the sequential rejective Bonferroni correction. It provides a control for Type-I errors at a prescribed significance level – in our case $\alpha = 0.05$ – while providing a substantial increase in the statistical power over the commonly used Bonferroni correction [22]. In the case of testing the family of N null-hypothesis and significance level α, the Holm-Bonferroni method proceeds as follows:

1) Hypothesis with the smallest observed p-value, is tested using the adjusted significance level $\alpha' = \alpha/N$, in the same manner as in the traditional Bonferroni procedure.
2) However, the next smallest observed p-value is tested using differently adjusted significance level $\alpha' = \alpha/(N-1)$.
3) The same process repeats up to the hypothesis with the highest observed p-value which is tested using the unadjusted significance level α.
4) The important additional rule is that if any of the hypothesis in the family gets rejected, then *all the subsequent* hypotheses are rejected as well regardless of their observed p-values.

By using differently adjusted statistical significance levels, Holm-Bonferroni method guarantees that the family-wise error rate is kept at the prescribed level, while providing a significant increase in the statistical power over the more commonly used simple Bonferroni correction [22]. We used the Holm-Bonferroni correction for testing the overall significance of the regression models, and for testing the significance of the individual predictor variables. In our case, with five hypothesis tests, the values of the adjusted statistical significance levels were $\alpha = [0.01, 0.0125, 0.0167, 0.0250, 0.05]$.

We also inspected the QQ-Plots for the signs of the severe deviation from the normality of residuals, and we assessed the multicollinearity of the three predictor variables using the variance-inflation factors (VIFs). The QQ-Plots did not reveal deviations from the normality of the residuals and VIF values were substantially lower than the typically used thresholds such as 4 or 10 [10]. Thus, we considered the use of the multiple linear regression appropriate for our study.

4. RESULTS

The results of the regression analyses are shown in Table 6. The models for betweenness, in-degree, out-degree and in-closeness centralities were significant, while the model for out-closeness was marginally significant.

In the case of betweenness centrality, the multiple regression model explained 32% of the variability in the students scores of betweenness centrality. The backwards-stepwise regression analysis selection using the (AIC) criterion resulted in a regression model consisting of the affective and interactive categories of social presence, and both variables were found to be statistically significant predictors of betweenness centrality. In terms of their relative importance, the interactive category had a slightly larger standardized β coefficient than the affective category of social presence, indicating a slightly larger effect on the students' betweenness centrality scores.

With respect to degree centrality, the regression models explained 86% and 83% of the variability in the measures of in-degree and out-degree centralities, respectively. All three predictors were positively associated with the degree centrality measures, and all three reached the statistical significance. In terms of their relative importance, in both models, the interactive category of social presence

Table 6: Regression results for selected centrality measures after stepwise model selection using AIC criterion.

	Betweenness			In-degree			Out-degree			In-closeness			Out-closeness		
	β	SE	p	β	SE	p	β	SE	p	β	SE	p	β	SE	p
Affective	**0.27**	**0.12**	**0.024**	**0.18**	**0.054**	**0.001**	**0.23**	**0.059**	**<0.001**						
Interactive	**0.38**	**0.12**	**0.002**	**0.65**	**0.064**	**<0.001**	**0.65**	**0.07**	**<0.001**	**0.27**	**0.11**	**0.015**	0.37	0.15	0.017
Cohesive				**0.2**	**0.061**	**0.001**	**0.14**	**0.066**	**0.041**				-0.23	0.15	0.137
$F(3,77)$	**19.6**		**<0.001**	**159**		**<0.001**	**130**		**<0.001**	**6.24**		**0.015**	3.03		0.054
Adjusted R^2	**0.32**			**0.86**			**0.83**			**0.061**			0.048		

had the largest standardized β coefficient, while the affective and cohesive categories had roughly the same standardized coefficients.

Regarding the two closeness centrality measures, the regression model for in-closeness was statistically significant, explaining 6.1% of the variability in the students' in-closeness centrality scores, while the model for out-closeness failed to reach the significance by a very small margin. The model for in-closeness consisted of only the interactive category, which was found to be a statistically significant predictor of in-closeness centrality. Similarly, the regression model for out-closeness consisted of the interactive and cohesive social presence categories, and explained 4.8% of the variation in the students' out-closeness centrality scores. In the model for out-closeness centrality, the only statistically significant predictor was the interactive category of social presence, while interestingly, the cohesive category of social presence was negatively associated with the change in the out-closeness centrality values, although statistically insignificantly.

5. DISCUSSION

One finding immediately stands out of the regression analyses results: *Interactive social presence is the most strongly associated with all of the network centrality measures, indicating a significant relation with the development of the students' social capital.* A possible explanation of this lies to some degree in the nature of students' social networks. Given that the primary goal of social networks in online courses is to serve as a communication medium for fostering of collaborative learning [27], it is reasonable to expect that interactivity in communication can explain a significant proportion of the differences in network positions, and ultimately the differences in the development of students' social capital. The reason why the interactive category is had the strongest association might be that only after the students have gott familiar with each other through focused, on-task interactions, and after they have started developing trust within a learning community, the expression of emotions and the sense of group belonging begins to emerge. This is aligned with the findings of Garrison [27] who suggested that interactive social presence is dominant at the beginning of a course, but decreases over time, while affective and cohesive social presence increase over time [27]. However, as Garrison [27] points out, too much of the interpersonal and affective interactions undermine the productivity of the collaborative learning activities. There is a certain amount of social interactions that is beneficial for learning [27], and the focus of the instructional interventions should be on: i) stimulating the right amount of the different social interactions that support productive and purposeful collaborative learning activities, and ii) the development of trust and the sense of community among the group of learners [17].

One practical implication of these results is that they suggest the effective way for fostering the productive social climate – and that is *focusing on the student interaction and open communication*. In order to guide the development of the social relationships in a learning community, it seems that the instructional emphasis should be on the interventions that require engaging in an open exchange of ideas and opinions, that would in turn lead to more affective expression, and eventually to the development of the sense of community belonging. Still, this hypothesis warrants further investigation, and in the future we plan to analyze the evolution of the students' social presence and the corresponding social network structures over time, which would shed new light on this important question.

The results of individual network centrality measures revealed that both in-degree and out-degree centrality measures were significantly predicted by all the three categories of students' social presence. By looking at the description (Section 2.2.2) and the indicators (Table 1) of the interactive category of social presence, we can see that interactive social presence is mainly about stimulating open and direct communication between the students. Thus, the students who exhibit a high level of interactive social presence have higher chances of "provoking" a response from the other students. Activities such as asking questions, explicitly referring to other students by name, quoting their messages, complementing them or agreeing with their messages, are all activities associated with an interactive and open communication, and can be used to elicit a response from the other students. It would be interesting to further investigate the relationship between different indicators of social presence and social capital, as certain indicators – such as `I4` "Asking questions" – seem to have more impact than the other indicators. Besides the interactive category, the regression model revealed that the affective and cohesive categories of social presence were also significant predictors of in-degree and out-degree centralities. These findings are even more interesting, as affective and cohesive exchanges are not directly stimulating discussions in the same manner as the interactive category. Further investigation is needed to examine particular time periods over the duration of a course in which those different dimensions of social presence contribute to the degree centrality measures of students.

With respect to betweenness centrality that is most closely related to the notion of social capital [13, 12], the regression model was statistically significant and explained 32% of the variability in the betweenness centrality scores. This corresponds to Cohen's $f^2 = 0.47$ effect size, which is considered to be a large effect size [16]. Both the interactive and affective categories of social presence were statistically significant predictors of the betweenness centrality, with the interactive category having a bit greater standardized β coefficient. This might be due to the nature of student communication networks and their focus on collaborative learning, which resulted in the emphasis on information exchange. Still, these are very intriguing findings, given that betweenness centrality is not directly related to the number of interactions the student has, but more to the overall diversity of the interactions within a group of learners. In a follow-up study, it would be very interesting to investigate whether there are any particular ways in which the students with the high betweenness centrality differ from the other students (e.g., asking many questions or exhibiting higher self-disclosure).

Regarding the closeness centrality measures, the regression model for in-closeness was also statistically significant. The model explained 6.1% of the variability, and the stepwise model selection using the AIC criteria resulted in a simple regression model with only the interactive category of social presence. In contrast to degree centrality, which considers only direct relationships, closeness centrality also considers the indirect relationships. Such indirect relationships could be the reason why only interactive category was rendered as important. The affective and cohesive exchanges between students A and B, although very important, provide very little, or no influence on the indirect relations of student B and the rest of the students. The similar findings we could see in the model for out-closeness, which was marginally significant with the p-value of 0.054. However, it could be expected that the significance of this model would be conformed in a larger replication study.

The major limitations of this study is the sample size and the use of the single course from a single institution. Even though there were six offerings of the course taught by the two instructors, there might still be significant effects of the adopted pedagogical approach, which could have shaped a specific social dynamics, and thus, potentially distort the findings of our study. Likewise, we considered all interactions among the students as contributing to their social capital, it is very likely that the certain interactions (e.g., adversarial interactions) might have a negative effect on the student social capital. In the future work, we plan on replicating our findings on a bigger sample and with more diverse courses from different subject matter domains. Finally, we plan to investigate the temporal aspects of the relationship between social capital and the social presence, which might give us a deeper insight into the complexity of the social interactions in learning communities.

6. CONCLUSIONS

The study presented in this paper investigated some of the social processes that can contribute to the development of students' social capital. We have looked at the relationship between students' *social presence*, operationalized through the Community of Inquiry model, and students' *social capital*, operationalized through the three network centrality measures. The implications of our findings are twofold: *First*, our results indicate that a significant part of the variability in network centrality scores can be explained using the three dimensions of the social presence, and this in turn indicates the existence of the relationship between the development of social presence and social capital. All three categories of social presence were significant predictors of in-degree and out-degree centrality measures while interactive and affective categories were significant predictors of the betweenness centrality. Also, interactive category of social presence was significantly predictive of the in-closeness and out-closeness centrality measures, although the overall regression model for out-closeness was marginally significant. A possible explanation is that given the task-oriented nature of discussions in online courses, students' social presence develops mostly through interactions focused on learning, and then over time, with the development of trust among a group of learners, the other dimensions of social presence start to emerge. *Second*, the study shows the significant relationship between the interactive category of social presence and betweenness, in-degree, out-degree, and in-closeness network centrality measures. This provides an empirical basis for fostering the productive social climate in discussions through interventions that increase interactivity and open communication among the students. By engaging students to participate in discussions with the clearly defined expectations, students develop social relationships which can in turn have positive impact on the attainment of the learning objectives and their overall academic experience.

References

[1] Special issue on the community of inquiry framework: Ten years later. *The Internet and Higher Education*, 13(1–2), 2010.

[2] T. Anderson and J. Dron. Three generations of distance education pedagogy. *The International Review of Research in Open and Distance Learning*, 12(3):80–97, 2010.

[3] T. Anderson, L. Rourke, D. R. Garrison, and W. Archer. Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks*, 5:1–17, 2001.

[4] J. Arbaugh, M. Cleveland-Innes, S. R. Diaz, D. R. Garrison, P. Ice, J. C. Richardson, and K. P. Swan. Developing a community of inquiry instrument: Testing a measure of the community of inquiry framework using a multi-institutional sample. *The Internet and Higher Education*, 11(3–4):133–136, 2008.

[5] A. W. Astin. *What Matters in College: Four Critical Years Revisited*. Jossey-Bass, 1 edition edition, 1997.

[6] R. S. Baker and K. Yacef. The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1):3–17, 2009.

[7] T. T. Baldwin, M. D. Bedell, and J. L. Johnson. The social fabric of a team-based M.B.A. program: Network effects on student satisfaction and performance. *The Academy of Management Journal*, 40(6):1369–1397, 1997.

[8] J. P. Bean. Conceptual models of student attrition: How theory can help the institutional researcher. *New Directions for Institutional Research*, 1982(36):17–33, 1982.

[9] R. A. Bjork, J. Dunlosky, and N. Kornell. Self-regulated learning: beliefs, techniques, and illusions. *Annual review of psychology*, 64:417–444, 2013.

[10] B. L. Bowerman and R. T. O'Connell. *Linear Statistical Models: An Applied Approach*. Duxbury Press, 1990.

[11] S. Buckingham Shum and R. Ferguson. Social learning analytics. *Journal of Educational Technology & Society*, 15(3):3–26, 2012.

[12] R. S. Burt. Structural holes versus network closure as social capital. In N. Lin, K. Cook, and R. S. Burt, editors, *Social Capital: Theory and Research*. Aldine Transaction, 2001.

[13] R. S. Burt. The social capital of structural holes. In M. F. Guillen, R. Collins, P. England, and M. Meyer, editors, *The New Economic Sociology: Developments In An Emerging Field*. Russell Sage Foundation, 2005.

[14] B. V. Carolan. *Social Network Analysis and Education: Theory, Methods and Applications*. SAGE Publications, Inc., 2014.

[15] H. Cho, G. Gay, B. Davidson, and A. Ingraffea. Social networks, communication styles, and learning performance in a CSCL community. *Computers & Education*, 49(2):309–329, 2007.

[16] J. Cohen. The analysis of variance. In *Statistical power analysis for the behavioral sciences*, pages 273—406. L. Erlbaum Associates, Hillsdale, N.J., 1988.

[17] S. Dawson. A study of the relationship between student social networks and sense of community. *Journal of Educational Technology & Society*, 11(3):224–238, 2008.

[18] M. F. De Laat, V. Lally, L. Lipponen, and R.-J. Simons. Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *International Journal of Computer-Supported Collaborative Learning*, 2(1):87–103, 2007.

[19] B. De Wever, T. Schellens, M. Valcke, and H. Van Keer. Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46(1):6–28, 2006.

[20] J. Dewey. My pedagogical creed. *School Journal*, 54(3):77–80, 1897.

[21] R. Donnelly and J. Gardner. Content analysis of computer conferencing transcripts. *Interactive Learning Environments*, 19(4):303–315, 2011.

[22] O. J. Dunn. Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293):52–64, 1961.

[23] J. E. Eckles and E. G. Stradley. A social network analysis of student retention using archival data. *Social Psychology of Education*, 15(2):165–180, 2011.

[24] S. L. Feld. The focused organization of social ties. *American Journal of Sociology*, 86(5):1015–1035, 1981.

[25] R. Ferguson and S. B. Shum. Social learning analytics: five approaches. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, LAK '12, page 23–33, New York, NY, USA, 2012. ACM.

[26] L. C. Freeman. Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215–239, 1978.

[27] D. R. Garrison. *E-Learning in the 21st Century: A Framework for Research and Practice*. Routledge, New York, 2 edition edition, 2011.

[28] D. R. Garrison, T. Anderson, and W. Archer. Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2–3):87–105, 1999.

[29] D. R. Garrison, T. Anderson, and W. Archer. The first decade of the community of inquiry framework: A retrospective. *The Internet and Higher Education*, 13(1–2):5–9, 2010.

[30] D. R. Garrison and J. Arbaugh. Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, 10(3):157–172, 2007.

[31] R. Garrison, M. Cleveland-Innes, and T. S. Fung. Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education*, 13(1–2):31–36, 2010.

[32] D. Gasevic, A. Olusola, S. Joksimovic, and V. Kovanovic. Externally-facilitated regulation scaffolding and role assignment to develop cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, (submitted), 2014.

[33] D. Gasevic, A. Zouaq, and R. Janzen. "Choose your classmates, your GPA is at stake!": The association of cross-class social ties and academic performance. *American Behavioral Scientist*, 2013.

[34] M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[35] T. J. Hastie, R. J. Tibshirani, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer, New York, NY, 2013.

[36] C. Haythornthwaite. Social network analysis: An approach and technique for the study of information exchange. *Library & Information Science Research*, 18(4):323–342, 1996.

[37] S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

[38] K. H. Krippendorff. *Content Analysis: An Introduction to Its Methodology*. Sage Publications, 2003.

[39] N. Lin. Social resources and instrumental action. In P. V. Marsden and N. Lin, editors, *Social structure and network analysis*, pages 131—145. Sage Publications, 1982.

[40] P. Long and G. Siemens. Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5):31–40, 2011.

[41] R. Luppicini. Review of computer mediated communication research for education. *Instructional Science*, 35(2):141–185, 2007.

[42] A. Mehrabian. Some referents and measures of nonverbal behavior. *Behavior Research Methods & Instrumentation*, 1(6):203–207, 1968.

[43] T. Rizzuto, J. LeDoux, and J. Hatala. It's not just what you know, it's who you know: Testing a model of the relative importance of social networks to academic performance. *Social Psychology of Education*, 12(2):175–189, 2009.

[44] C. Romero and S. Ventura. Educational data mining: A review of the state of the art. *Trans. Sys. Man Cyber Part C*, 40(6):601–618, 2010.

[45] L. Rourke, T. Anderson, D. R. Garrison, and W. Archer. Assessing social presence in asynchronous text-based computer conferencing. *The Journal of Distance Education*, 14(2):50–71, 1999.

[46] J. Scott and P. J. Carrington. *The SAGE Handbook of Social Network Analysis*. SAGE Publications, 2011.

[47] P. Shea, S. Hayes, S. U. Smith, J. Vickers, T. Bidjerano, M. Gozza-Cohen, S.-B. Jian, A. Pickett, J. Wilde, and C.-H. Tseng. Online learner self-regulation: Learning presence viewed through quantitative content- and social network analysis. *The International Review of Research in Open and Distance Learning*, 14(3):427–461, 2013.

[48] P. Shea, S. Hayes, J. Vickers, M. Gozza-Cohen, S. Uzuner, R. Mehta, A. Valchova, and P. Rangan. A re-examination of the community of inquiry framework: Social network and content analysis. *The Internet and Higher Education*, 13(1–2):10–21, 2010.

[49] R. A. Smith and B. L. Peterson. "Psst ... what do you think?" the relationship between advice prestige, type of advice, and academic performance. *Communication Education*, 56(3):278–291, 2007.

[50] S. L. Thomas. Ties that bind: A social network approach to understanding student integration and persistence. *The Journal of Higher Education*, 71(5):591–615, 2000.

[51] V. Tinto. *Leaving College: Rethinking the Causes and Cures of Student Attrition*. University of Chicago Press, 1993.

[52] S. Wasserman. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.

[53] Y. Woo and T. C. Reeves. Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and Higher Education*, 10(1):15–25, 2007.