CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Edinburgh Research Explorer

THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Social presence in online discussions as a process predictor of academic performance

Citation for published version:
Joksimovic, S, Gasevic, D, Kovanovic, V, Riecke, B & Hatala, M 2015, 'Social presence in online discussions as a process predictor of academic performance' Journal of Computer Assisted Learning. DOI: 10.1111/jcal.12107

Digital Object Identifier (DOI):
10.1111/jcal.12107

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal of Computer Assisted Learning

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 05. Apr. 2019

Social presence in online discussions as a process predictor of academic performance

Srećko Joksimović[a], Dragan Gašević[a,b], Vitomir Kovanović[b], Bernhard E. Riecke[c], Marek Hatala[c]

[a]Moray House School of Education, University of Edinburgh, United Kingdom

[b]School of Informatics, University of Edinburgh, United Kingdom

[c]School of Interactive Arts and Technology, Simon Fraser University, Canada

Abstract

With the steady development of online education and online learning environments, possibilities to support social interactions between students have advanced significantly. This study examined the relationship between indicators of social presence and academic performance. Social presence is defined as students' ability to engage socially with an online learning community. The results of a multiple regression analysis showed that certain indicators of social presence - i.e., *continuing a thread* and *complimenting, expressing appreciation* – were significant predictors of final grades in a master's level computer science online course. Moreover, the study also revealed that teaching presence – operationalized through the changes in instructional design - moderated the association between social presence and academic performance, indicating that a course design that increased the level of meaningful interactions between students had a significant impact on the development of social presence and thus, could positively affect students' academic performance. This is especially important in situations when discussions are introduced to promote the development of learning outcomes (e.g., critical thinking) assessed in courses. Another implication of our results is that indicators of social presence can be used for early detection of students at risk of failing a course. This finding informs research and practice in the emerging field of learning analytics by prompting the opportunities to offer actionable insights into the reasons why certain students are lagging behind. These insights are grounded in indicators of the quality of social knowledge construction rather than a simple quantity of discussion posts.

Keywords: Social presence, Teaching presence, Content analysis, Discussion forums

Introduction

A common assumption for small and medium-sized classrooms is that instructors can observe learning progression of their students and detect the extent to which they are engaged (Ready, 2008). However, in a large classroom setting and/or with the incorporation of information and communication technologies into teaching and learning experience (i.e., blended learning), the challenge of evaluating learning and observing students' engagement becomes more evident (Ginns & Ellis, 2007). Moreover, online learning assumes that students are geographically dispersed, that they learn at different times (i.e., asynchronously), and that visual contact is limited or does not exist (Dabbagh, 2007; Kear, 2010). In such an environment, teachers have very limited possibilities to determine the level of students' engagement, thus, not being able to focus on a particular student (or group of students) in order to increase the level of their engagement.

Many researchers have stressed the importance of social presence as a crucial element in order to understand factors that influence the formation of online learning communities (e.g., Kear, 2010; Ubon & Kimble, 2004; Vrasidas & McIsaac, 1999). Consequently, different frameworks have been proposed in order to analyze social presence within online learning communities (e.g., Rourke, Anderson, Garrison, & Archer, 1999; Tu, 2002; Ubon & Kimble, 2004; Weaver & Albion, 2005). The initial application of the social presence construct originates in the communication literature, where Short, Williams, and Christie (1976) defined social presence as the level of "salience" that emerges from the interaction between two (or more) persons. However, provenance of the social presence construct comes from the work of Mehrabian (1968), and his definition of "immediacy". This definition was later adapted by

Garrison, who defined social presence as “the degree to which participants are able to project themselves affectively within a medium” (1999, p.6). However, without any face-to-face interaction, it is a challenging task to reveal students' sense of social presence within an online community.

Following the definition of social presence suggested by Garrison (1999), and later extended in Garrison and Akyol (2013), this study aims at revealing whether (and how) different indicators of social presence (Table 1) are associated with academic performance. Providing teachers with the information about the level of social engagement and possible implications based on that finding, might help identifying students at risk and guide externally-facilitated interventions (Gašević, Adescope, Joksimović, & Kovanović, 2015). Moreover, informing students of their activities and increasing awareness of the level of social presence of their peers, should lead to enhanced self-regulation of social interaction (Järvelä & Hadwin, 2013). Hence, many researchers revealed that students' sense of belonging to an online learning community might indicate the level of motivation (Rovai, 2002; Tao & Florida, 2009; Weaver & Albion, 2005), satisfaction with course content and teachers (Cobb, 2009; Kear, 2010; So & Brush, 2008), course retention (Liu, Gomez, & Yen, 2009), as well as the degree of perceived learning (Richardson & Swan, 2003; Wise, Chang, Duffy, & del Valle, 2004). Therefore, it seems promising to study the connection between students’ social presence and academic performance.

Insert Table 1 about here

Highlighting the importance of analyzing relationships between three presences within the CoI model, Garrison and Arbaugh (2007) confirmed a causal relationship from teaching presence to both social and cognitive presence. Therefore, this study aims at investigating the moderating role from teaching presence to social presence within the community of inquiry

model (Garrison & Arbaugh, 2007; Garrison, 1999). Given that the theoretical model of teaching presence is conceptualized through *instructional design*, *facilitation*, *direct instruction* (Anderson, Rourke, Garrison, & Archer, 2001), we examined how changes in teaching presence might affect the development of sociological processes that determine indicators of social presence and its predictive power of academic performance.

Theoretical Background

Community of Inquiry Model

The Community of Inquiry (CoI) model presents an effective approach to assessing the quality of knowledge construction in online discussions, and has become one of the most researched models that fosters higher-order learning (Garrison & Arbaugh, 2007; Swan & Ice, 2010). The CoI model is social constructivist in nature and is concerned with deep and meaningful learning through computer-mediated interaction among students and between students and instructors. Therefore, the main focus of the CoI model is the learning process (Akyol et al., 2009). According to Garrison (1999), three interleaved components are essential in order to achieve an effective educational experience in a CoI: i) cognitive presence, ii) teaching presence, and iii) social presence.

Cognitive presence is operationalized through the notion of practical inquiry, which represents a process of constructing and validating meaning through critical thinking and extended communication within educational environments (Garrison, Anderson, & Archer, 2001; Kozan & Richardson, 2014). Therefore, cognitive presence is usually identified as the main construct of the CoI model. Cognitive presence is comprised of four iterative (and occasionally recurrent) phases: triggering, exploration, integration, and resolution (Garrison et al., 2001), where each phase represents a different socio-cognitive process in a practical inquiry.

Anderson et al. defined **teaching presence** as a "design, facilitation, and direction of cognitive and social processes for the purpose of realizing personally meaningful and educationally worthwhile learning outcomes" (2001, p.5). The instructor's responsibility for establishing and maintaining a high level of learning quality begins with a course design and continues throughout direct course instruction and/or facilitation. Therefore, teaching presence is defined through the following three components: design and administration, facilitating discourse, and direct instruction (Anderson et al., 2001).

In order to support a practical inquiry and development of higher order thinking, the substantial level of interaction among students and between students and instructors needs to be established (Garrison & Akyol, 2013; Kozan & Richardson, 2014). Students' ability to project themselves within an online learning community and the level of their communication with peers and instructors was initially identified as **social presence** (Rourke et al., 1999). However, social presence in a community of inquiry is not limited to establishing social relationships between their members. The most important purpose of social presence – that supports higher order thinking – is in creating a healthy support that provides a comfortable place for students to exchange ideas freely, explore different perspectives, and solve problems collectively. More precisely, the focus of social presence is on the progressive development of relationships among community members, through a meaningful and purposeful collaboration (Garrison & Akyol, 2013).

According to the original definition by Rourke et al. (1999), the main categories that constitute social presence are the affective, interactive and cohesive communication types. However, as a result of over a decade of sustained research related to the CoI constructs in general, and social presence in particular, a revised definition of social presence recognizes

interpersonal, open and cohesive communication as the main factors of social presence (Garrison & Akyol, 2013; Garrison, 2011). Specific indicators of each communication type (Table 1), should evolve over time, while the theoretical model also predicts (Garrison & Akyol, 2013) that an open communication should be higher at the beginning of an interaction, and decrease towards the end of the interaction. On the other hand, indicators of interpersonal and cohesive communication are likely to increase as group members engage more effectively into the learning process. Moreover, the three groups of indicators of social presence are highly interleaved, in a sense that the level of interpersonal communication influences the climate for an open communication while open, challenging, and respectful communication contributes to the higher group cohesion in a community (Garrison & Akyol, 2013).

Social presence in online learning communities

Analyzing factors that influence the development of social presence within online learning communities, researchers mostly agree that student behavior, interactions between participants and characteristics of communication medium are the most significant prerequisites needed to establish and sustain social presence (Kear, 2010; So & Brush, 2008). On the other hand, understanding the importance of social presence for the development of a healthy and successful learning community, researchers in distance education found the association between social presence and different constructs, such as motivation (Jorge, 2010; Rovai, 2002; Tao & Florida, 2009; Weaver & Albion, 2005), satisfaction (Cobb, 2009; Kear, 2010; So & Brush, 2008), perceived learning (Richardson & Swan, 2003; Wise et al., 2004), course retention (Liu et al., 2009), enrollment intent (Reio & Crim, 2013), interaction between course participants (Picciano, 2002; Tu, 2002), and performance in online courses (Liu et al., 2009; Picciano, 2002).

Similar to our study, Picciano (2002) and Liu et al. (2009) analyzed relationship between social presence and learning outcomes in fully online courses. While Picciano (2002) examined a graduate course in an education administration program (*N*=23), Liu et al. (2009) analyzed several courses (including mathematics, science, business, English, history, and psychology) at a suburban community college (*N*=108). However, both studies relied on a social presence questionnaire developed by Tu (2001, 2002). In our study, we relied on the quantitative content analysis of discussion transcripts by coding messages using indicators of social presence described in Table 1, as another well-accepted instrument for measuring social presence in communities of inquiry (Garrison & Arbaugh, 2007).

Picciano (2002) analyzed the association between perceived social presence and two measures of academic performance: scores on a final examination and a written assignment. Interestingly, those two measures revealed different correlations with the student perceptions of social presence. While the correlation between perceived social presence and the written assignment was positive and significant, perceived social presence and the final examination yielded a negative and not statistically significant correlation. Picciano (2002) concluded that this might be explained with the specific structure of the course, where students communicated with each other using online discussions. Therefore, Picciano (2002) assumed that the characteristics of activities in a written assignment were similar to those on the discussion board, as opposite to the examination which included multiple-choice questions and was considered as rather an asocial activity. We could expand this conclusion of Picciano (2002) and propose that it is not crucial whether the assessment activity is social or not for the association between social presence and academic performance. We posit if the development of skills promoted with meaningful social interaction is well aligned with the learning outcomes assessed, the association

between social presence and academic performance scores will likely be found. In the case of the Picciano (2002) study this proposition would be interpreted as follows - given that quizzes are not typically suitable for assessment of critical thinking skills, the association between social presence (that promotes higher order and critical thinking) and quiz scores was unlikely.

Analogous to our study, Liu et al. (2009) also used final course grade as a performance measure. Using an ordinal logistic regression model, the study by Liu et al. (2009) revealed a statistically significant correlation between perceived social presence and the final course grades. Moreover, Liu et al. (2009) also showed that social presence is a significant predictor of online course retention, and further defined guidelines for building effective online learning communities and early identification of students at risk in order to increase their chances to successfully complete the course work. Liu et al. (2009) also concluded that a higher perception of social presence led to a higher interaction between participants, increasing students' likelihood to attain better grades.

The study reported in this paper focuses on the social presence within online learning communities, however, there is an extensive research on how (and whether) other types of social activities might predict learning outcome. For example, Agudo-Peregrina, Iglesias-Pradas, Conde-González, and Hernández-García (2014) analyzed six online courses (N=138) and two face-to-face courses that were supported with learning management system (N=218). A multiple linear regression analysis performed by Agudo-Peregrina et al. (2014) showed that the quantity of student-student and student-teacher interactions significantly predicted success in online courses. Using classification via clustering approach, Lopez, Luna, Romero, and Ventura, (2012) and Romero, López, Luna, and Ventura (2013) analyzed whether student participation in a discussion forum (N=114) can predict the final learning outcome. Their studies revealed that the

number of messages, number of words, average score messages, degree of centrality and prestige were the most significant predictors. Their approach also allowed for early prediction of students at risk of failing a course. Yoo and Kim (2012) applied a multiple linear regression to examine whether speech acts, conversational dialogue dynamics and emotional features, as defined by the Linguistic Inquiry and Word Count (LIWC) tool (Tausczik & Pennebaker, 2010), of online discussions can predict student success. Their results indicate that project grades were significantly correlated with the number of answers provided to peer students, the quantity of words that belonged to the positive emotion LIWC category, and the time before the deadline and before the initiation of collaboration (Yoo & Kim, 2012). Finally, Macfadyen and Dawson (2010) showed that the total number of discussion messages posted and the number of reply discussion messages posted had positive and statistically significant correlations with grades. Similar positive associations between student grades and the volume of the activity in asynchronous online discussions were found in several other studies (Alstete & Beutell, 2004; Finnegan, Morris, & Lee, 2009; Palmer, Holt, & Bray, 2008).

Interconnectedness of social, teaching, and cognitive presence

Gašević et al., (2015) analyzed the relationship between teaching presence and cognitive presence, observing the same course over several offerings, with the different levels of teachers' engagement. Consistent with the previous research (Garrison & Cleveland-Innes, 2005), the findings indicate that instructional conditions, operationalized through the three dimensions of teaching presence (Anderson et al., 2001), have a significant impact on the facilitation of cognitive presence. Moreover, Garrison and Cleveland-Innes (2005) also revealed that engagement into deep and meaningful learning depends on the level of teaching presence. Social presence by itself and shallow interaction are not enough for establishing an environment for

meaningful learning; course design with clear guidelines, expectations, and scaffolds for participation in online discussions as well as a high level of leadership by a course teacher, are necessary for students to take a deep approach to learning.

Analyzing emerging issues in the research literature based on the community of inquiry model, Garrison and Arbaugh (2007) stressed the importance of social presence for establishing and maintaining the community and creating a climate for an open and trustworthy communication within a cohesive learning group. More recent studies of Shea and Bidjerano (2009) and Garrison et al. (2010) showed that students' perception of teaching presence can predict perception of cognitive presence, while social presence should be observed as a mediating variable between teaching and cognitive presences. Both studies relied on the CoI survey responses in order to analyze the interconnectedness between three presences in a fully online program, with the difference that Shea and Bidjerano (2009) applied a factor analysis and structural equation modeling (N=2159), while Garrison et al. (2010) focused on the factor analysis only (N=205). On the other hand, Shea et al. (2010) analyzed discussions in two online courses (N_{post}=944) with different levels of teaching presence, and suggested more complex relationships between presences. Specifically, they observed teaching presence through two components – teachers' teaching presence and teachers' social presence. Using social network and content analysis Shea et al. (2010) revealed a complex correlation between students' social presence and two components of teaching presence. For example, the course with a higher teaching presence yielded a strong correlation between teachers' teaching presence and students' social presence, while the association was even stronger between teacher's and students' social presence for the same course. On the other hand, the course with a lower level of teaching presence showed a weaker correlation in both cases. Nevertheless, findings in studies by Shea

and Bidjerano (2009), Shea et al. (2010), and Garrison et al. (2010) clearly support the importance of effective instructional scaffold for the development of social presence. It seems that in Shea et al.'s (2010) study, the course with a higher teaching presence provided a better support for meaningful interactions to occur (Vrasidas & McIsaac, 1999), leading towards a higher quality discourse and increased students' social presence.

Following conclusions from the previous studies (Garrison et al., 2010; Shea & Bidjerano, 2009; Shea et al., 2010), we aim at gaining a theoretical and practical insight into the nature of relationships between teaching and social presence within the community of inquiry. Moreover, we also investigated how changes in a course design and the level of teachers' leadership influence the extent to which social presence is associated with the final course grade.

Research questions

Since most of the studies showed the importance of social presence as a necessary factor for building a healthy online community, which leads to improvements in the quality of knowledge construction (e.g., Cobb, 2009; Liu et al., 2009; Rovai, 2002; Swan & Shih, 2005; Vrasidas & McIsaac, 1999), we analyzed whether indicators of social presence, as defined by Garrison (2011) and presented in Table 1, were associated with final course grades. In other words, we examined whether the higher level of social presence leads to better academic performance. Therefore, we define our first research question:

RQ 1: *Is the higher degree of social presence within an online learning community associated with higher academic performance?*

Following the recommendations made by Garrison and Arbaugh (2007), and building on the findings from similar studies (Garrison & Cleveland-Innes, 2005; Kozan & Richardson, 2014; Shea & Bidjerano, 2009; Shea et al., 2010), we also analyzed a potential moderation role

of teaching presence in the association of social presence and academic performance within the community of inquiry model. Gašević et al. (2015) showed that the statistical interaction between externally-facilitated regulation scaffolding and role assignment (i.e., increased level of teaching presence) had a significant effect on cognitive presence. Expanding on those results, we intend to further explain how different instructional conditions influence the predictive power of social presence on academic performance. Thus, we define the following research question:

RQ 2: *Does instructional design moderate the association between social presence and academic performance? That is, whether and if so, to what extent, the change in instructional design alters the significance and strength of the association between individual indicators of social presence and students' grades?*

The final research questions is a direct consequence of investigated research question 1, provided that the moderation effect of instructional design was found in research question 2:

RQ 3: *Are students with a lower level of social presence at higher risk of failing a course under different instructional conditions?*

The third research question might unveil certain practical implications in terms of specifying instructional conditions that will enhance potential knowledge construction benefits.

Method

In this section, we describe the study design, the data collection process, the measures used in the study, and the analysis performed on the collected data.

Study Setting and Data Collection

For the purpose of this research, we used a dataset obtained from a research intensive software engineering course of a master's of science (information systems) program in a public

online university in Canada[1]. The dataset contained 1,747 messages of students' online discussions within an asynchronous discussion forum with 84 different topics. The course for this study was selected as it allowed for investigating the effects of the moderating role of teaching presence on the association between social presence and academic performance. In particular, the course had a revision of teaching presence operationalized through the changes in its instructional design: i) minimal guidance for social interaction provided focusing primarily on quantitative expectations (control group) and ii) externally-facilitated scaffold tailored to increase the level of students' engagement – i.e., social and cognitive presence - (treatment group) observed in the control group. Following the proposed structure for designing effective online discussions (Rovai, 2007), participation guidelines in both the control and treatment groups covered the following elements: motivation (i.e., student participation in online discussion was graded), expectations (i.e., the expected contributions in terms of the quality of posts was defined) and task-oriented discussions (i.e., discussions were focused on specific topics being presented by each student in the course). The difference between the control and the treatment group was only in terms of the expectations, where this component for the treatment group introduced an externally-facilitated regulation scaffold. For example, in the control group, the students were asked to participate in discussions on the presentations of their peers in order develop a constructive discussion around the topics presented. However, in the control group, it was not specified what "constructive discussion" meant, nor there were clear guidelines related to the quality of messages expected. In the treatment group, participation guidelines clearly

[1] Being part of an online university, the program is specifically tailored for part-time students who are working full-time. The program has two intakes of students annually. Students on average enroll between two to three courses annually, each course is valued three credits as a norm across Canada. The students are requested to complete nine to twelve three-credit courses to fulfill the requirements for granting the degree. The difference stems from the diversity of the background of the students, as students who do not have undergraduate degrees in the area are requested to take up to additional three foundational courses; otherwise, the students would be waived from the requirement to complete those courses. Completion of the foundational courses is a prerequisite for the course investigated in this study. The allowed time for completion of the program is five years.

asked for specific types of quality of messages in order to provide students with clear expectations regarding the quality of their posts. The details about the scaffold, the text of the scaffold itself and its theoretical background are described by Gašević et al. (2015).

The data used in the study were collected from the course offerings in: i) control group – Winter 2008 (*N*=15) and Fall 2008 (*N*=22), and ii) treatment group – Summer 2009 (*N*=10), Fall 2009 (*N*=7), Winter 2010 (*N*=14), and Winter 2011 (*N*=13). The control group and the first offering of the treatment group were taught by one instructor and the last three offerings of the treatment group were taught by another instructor. All the students enrolled in the course under the study were part-time students (being employed while study). Descriptive statistics for grade average point (GPA), number of courses enrolled and number of days in program before enrolling the course under the study are presented in Table 2. There was no statistically significant difference in the three variables between the control and treatment group (Table 2).

Insert Table 2 about here

In the control and treatment groups, the course design requested students to participate in asynchronous online discussions in weeks 3-5; the participation in the discussions was graded and contributed 15% towards the final grade. Specifically, the students were requested to engage into discussions of the presentations given by each of their peers, who a) first recorded their presentations of a selected peer-reviewed paper published in the leading software engineering journal/conference proceedings, and b) then shared the video recordings through a university-hosted video streaming website. The presentations constituted assignment one of the courses and contributed 15% towards the final grade of the course. In those discussions, the students were requested to ask questions about the presentations, draw connections between the presentations and the other readings they had in the course, and generate the ideas and methods for their

literature reviews and research projects requested to be performed in assignments two and four contributing to the final grade 25% and 30%, respectively. Assignment three (15% of the final grade) of the course requested students to provide written answers to six questions expecting students to synthesize and contrast different perspectives found in the course readings. Assignments two, three and four were all graded based on the quality of the arguments to critically assess different perspectives and to justify decisions made and/or positions taken methodologically in addition to the assumed expectation for theoretical and/or empirical correctness of the statements made. The course had no final examination. It is also important to note that Instructor A taught the first three offerings and Instructor B taught the following three offerings. Instructor A designed the intervention and Instructor B did not change the instructional design at all. Instructor B followed the original design even to the point that they used all the emails of instruction that Instructor A created and communicated to the students in the third offering of the course (i.e., the offering when the intervention was introduced).

Variables and Content Analysis

According to the community of inquiry model (Garrison & Akyol, 2013; Garrison & Anderson, 2000), social presence within an online learning community is described by using 12 indicators grouped into three categories, presented in Table 1. Counts of each indicator of social presence per student are *predictors* or *independent variables*, while academic performance (i.e., the final grade) is the *outcome* or *dependent* variable.

To assess the level of social presence of students in the asynchronous online discussions, two independent coders coded student discussion transcripts, according to the three categories of social presence proposed within the coding scheme developed by Rourke et al. (1999). An individual message was identified as a unit of analysis, as an objectively identifiable unit that

produces a manageable set of cases, and it is a unit whose parameters are defined by the author of the message (Rourke, et al., 2001; Strijbos, Martens, Prins, & Jochems, 2006). As each unit of analysis can be simultaneously coded with more than one category of social presence, we assigned three binary indicators to each message, indicating the interpersonal, open and/or cohesive communication categories of social presence. However, very early in the coding process we observed the high frequency of some of the indicators of open and cohesive communication, causing that most of the messages would be classified with both of these categories. Having the most of the messages classified as open and cohesive communication would rather limit the discriminative power of those two categories; therefore we decided to repeat the coding process at the level of indicators of social presence, resulting with the 12 binary codes for each unit of analysis. The overall observed agreement for each category was satisfying, since coders reached the percent agreement of at least 84% (Table 1).

The course under study was three credits worth according to the course credit hours commonly used in Canadian courses (i.e., the course had a typical credit value), while grades in the program were from F to A+. According to the university policy, the grades were converted into grade points in the range from 0 to 4 where each grade increase resulted in an increment of .33 grade point (e.g., from B- to B), except for A and A+ where both worth 4 grade points. However, the minimal grade for the course to be counted towards the degree program was B- (i.e., 2.66). Thus, the students who received B- or lower can be considered not receiving a passing grade.

Analysis

To investigate research questions 1 and 2, we performed a multiple regression analysis. However, the individual contribution of variables in a regression model depends on the sequence

of how variables were introduced in the model (Field, Miles, & Field, 2012). Considering that there is limited knowledge on which of the indicators of social presence is a better predictor of academic performance, we first performed a Pearson's correlation analysis for each indicator with the final grades.

Using a version of the backward elimination algorithm (Kutner, Nachtsheim, & Neter, 2004), with the Akaike Information Criterion (AIC) (Hastie, Tibshirani, & Friedman, 2013) as a selection criterion, we reduced a model size and selected the regression model with the best fit to our data. The complete procedure was conducted through several steps: (i) the initial regression model contained all the variables (i.e., all 12 indicators of social presence); (ii) after the analysis was performed, the variable with the highest *p-value* was removed from the regression model; (iii) the process iteratively continued until we obtained the model with one independent variable; (iv) for each iterative model, the AIC value was calculated, and the model with the lowest AIC value was selected as optimal. Further, normal Q-Q plots of standardized residuals were used to check normality of the distribution of errors, which was confirmed in all the regression analyses reported in the paper. The assumptions of independent errors and multicollinearity between predictors were also tested and no regression model, reported in the paper, violated these two assumptions.

To test the moderation effect of instructional design on the association between indicators of social presence and academic achievement (research question 2), we first compared the differences in the level of each indicator of social presence between the control and treatment groups. Due to a non-normal distribution of the variables, which was confirmed using the Shapiro-Wilk test and P-P plots, we performed analysis by using the Mann-Whitney test. Moreover, we also performed multiple linear regression analyses to investigate the effect of

interaction of the group assignment (control vs. treatment) and each indicator of social presence. The effect of interaction on academic performance was found significant. The significant differences and interactions reinforced the validity to conduct separate regression analysis, by using the backward elimination algorithm, for each group - control and treatment - separately.

Finally, to investigate research question 3, for each group (control and treatment), we divided students into quartiles based on the indicators of social presence that were significantly associated with the student grades. The comparison between quartiles was performed by the Kruskal-Wallis test (due to the non-normal distribution of the variables), which was followed by the Mann-Whitney test with the Bonferroni correction. Results are considered significant if p is less than .05.

Results

Research question 1

The descriptive statistics for the variables used in the analysis of our first research question are presented in Figure 1 (mean values and 95% confidence intervals). The descriptive statistics showed that *continuing a thread* was the most commonly used indicator of social presence (M=19.87, SD=10.62). Fairly high mean values were also obtained in the case of the *vocatives* (M=17.06, SD=10.48), *complimenting, expressing appreciation* (M=16.57, SD=10.22), *phatics, salutation* (M=15.18, SD=10.84), and *asking questions* (M=9.54, SD=6.14) indicators. On the other hand, the indicators of *self-disclosure* (M=3.84, SD=3.55), *affective expression* (M=3.46, SD=4.48), *expressing agreement* (M=2.94, SD=2.26), *referring to the group using inclusive pronouns* (M=1.68, SD=1.88), *referring to others' messages* (M=1.06, SD=1.32), *quoting from others' messages* (M=0.79, SD=1.65), and *use of humor* (M=0.54, SD=0.99) were much less represented.

Insert Figure 1 about here

Insert Table 3 about here

The results of the correlation analysis presented in Table 3 showed that the *continuing a thread*, *asking questions*, *complimenting-expressing appreciation*, and *vocatives* indicators were strongly, positively correlated with the final grade, while the correlation with *self-disclosure* tend to be marginally significant.

The backward elimination procedure for the optimal model selection with the AIC criterion resulted in a regression model which included only two indicators of social presence: *continuing a thread* and *complimenting, expressing appreciation*. The results of the multiple regression analysis (Table 4) indicated that these two indicators explained 16% of the variance (R^2=.16, F(2,79)=7.398, p<.01) in final grades. Further, it was found that the both indicators *continuing a thread* and *complimenting, expressing appreciation* significantly predicted final grades (β=.9, p=.001, and β=-.61, p=.05, respectively), with the *complimenting, expressing appreciation indicator* being negatively associated with the final grade. More precisely, the results revealed that the final grades increased by 0.9 points, as a result of one standard deviation change (i.e., SD=10.62) in the count of messages classified as *continuing a thread* (i.e., almost an entire letter grade such as from B to A). This is quite significant having in mind that the grades ranged from 0 to 4. On the other hand, in the case of *complimenting, expressing appreciation* messages, the final grades decreased by 0.61 points as a result of one standard deviation change (i.e., *SD*=10.22).

Insert Table 4 about here

Finally, the assumptions of independent errors (Durbin-Watson value = 1.76, p=0.29) and multicollinearity between predictors (VIF values in Table 4) in the regression model were not violated.

Research questions 2 and 3

In order to address our second research question, we first tested the difference in the counts of indicators of social presence in the messages between the control (*N*=845) and treatment groups (*N*=902). Figure 2 indicates that the treatment group had higher mean values for each indicator of social presence. A further analysis by using the Mann-Whitney test revealed that those differences were significant for the following indicators of social presence: *self-disclosure* (W=845.5, p=.01), *continuing a thread* (W=831, p=.02), *referring to others' messages* (W=804, p=.03), *asking questions* (W=872, p=.006), *complimenting, expressing appreciation* (W=826, p=.03), *expressing agreement* (W=838.5, p=.02), *vocatives* (W=828, p=.02), and *phatics, salutations* (W=822, p=.03) indicators of social presence. Furthermore, the results of the regression analysis unveiled a significant interaction between the two social presence indicators, found as significant predictors in the analysis reported in Table 4, and the group assignment (control and treatment) variable. Since the interaction can indicate that the regression analysis performed on the total sample may underestimate/overestimate the effect of certain predictors for each of the two groups (Field et al., 2012), we decided to repeat multiple regression analyses with the samples from the control and the treatment groups, independently.

After applying the backward elimination algorithm for optimal model selection to the data from each of the two groups independently, the results were considerably different, although both models (for the control and treatment groups) included the same variables (Table 5) as the model for the complete sample reported in research question RQ1. In the case of the control

group, neither indicator was found as a significant predictor of the final course grades. However, since Cohen's f^2 value of .32 suggests a moderate to large effect size (Cohen, 1992), we are able to talk about practical significance of the *complimenting, expressing appreciation* indicator. This was further confirmed (research question 3) by splitting the control group dataset into quartiles based on the count of messages with the *complimenting, expressing appreciation* indicator of social presence (Table 6). In order to evaluate the differences in the count of the messages with the *complimenting, expressing appreciation* indicator over the quartiles, we performed the Kruskal-Wallis test. The test was significant ($\chi^2(3, 20)=17.85$, $p=.0004$), while the further post-hoc analysis using the Mann-Whitney test with the Bonferroni correction revealed significant differences between non-adjacent quartiles (Q1-Q3, Q1-Q4, and Q2-Q4). Results indicate that students with the highest frequency of the messages with the *complimenting, expressing appreciation* indicator attained the lowest final grades. We can also observe that the mean value of the students' grades with the highest number of messages that complemented others (i.e., Q4) in the control group was slightly over the minimal passing grade (i.e., 2.66 which corresponds to B-) allowed for the course to be counted towards the master's degree program of the institution in which the course was taught.

Insert Table 5 about here

In the case of the treatment group, *continuing a thread* was a significant predictor of the students' final grades. Moreover, the resulted model explained 25% of the variance ($R^2=.25$, $F(2,59)=10.08$, $p<.001$), with almost an identical value of the effect in the predicting the final grades ($\beta=.93$, $p=.008$) compared to the regression model created based on the joined data from both the control and treatment groups as reported earlier in the paper. Dividing the treatment group into quartiles, based on the count of the messages with the *continuing a thread* indicator,

we are able to conclude that the students who showed higher levels of this specific indicator of social presence (*continuing a thread*) in their discussion posts attained higher grades. The Kruskal-Wallis test found that the differences among the quartiles with respect to the count of the messages with the *continuing a thread* indicator were significant ($\chi^2(3, N=62)=57.22$, $p<.001$). The post-hoc Mann-Whitney test with the Bonferroni correction revealed significant differences between the non-adjacent quartiles (Q1-Q3, Q1-Q4, and Q2-Q4). That is, the students with the lowest count of the messages that continued a thread had the lowest grades. We can also observe the students (Q1) with the lowest number of the messages that continued a thread had the mean value of their grades lower than the minimal passing grade.

Insert Table 6 about here

The assumptions of independent errors (Durbin-Watson value=1.83, p=.75 for the control and value=1.83, p=.47, for the treatment group) and multicollinearity between predictors (VIF values in Table 5) were not violated in the regression models.

Discussions and Conclusions

This study analyzed the importance of social presence as a predictor of students' academic performance (i.e., the final grade), in an online Master's level course. The results show that certain indicators of social presence (i.e., *continuing a thread* and *complimenting, expressing appreciation*) were significant predictors of academic performance. Additionally, our study stressed the importance of teaching presence for the development of social presence and the potential impact of teaching presence on the association between social presence and academic performance.

Many researchers (e.g., Garrison & Arbaugh, 2007; Kehrwald, 2008; Swan & Shih, 2005) showed that one of the main prerequisites for establishing a healthy learning community is

to support the development of social presence. Those studies showed that social presence is a good predictor of student motivation, satisfaction, and perceived learning (Section 2 – Theoretical Background), while results of our study (RQ1) further support findings of Picciano (2002) and Liu et al. (2009) that the ability of a student to project themselves within an online learning community is also a significant predictor of academic performance. More importantly, our findings indicate a positive relationship between active participation in asynchronous online discussions (i.e., *continuing a thread*) and final grades, while *complimenting* posts of other students (e.g., "*I found the presentation to be very clear and concise. It was well organized and covered the requirements for the TMA1...*") had a negative effect on academic performance (Table 6). It seems that students, who tended to agree with most of the posts created by their peers, were either superficially engaged into discussions or struggled with the course requirements and sought help, without sufficient learning gains to meet the expectations assessed in the course. Although these results might have been expected, given the findings from Picciano's (2002) and Liu et al.'s (2009) studies, our further analysis (i.e., RQ2) shows that the relationship between social presence and learning outcomes depends on the instructional conditions (Table 5). More precisely, without externally facilitated scaffold towards the quality of the posted messages, there was no positive relationship between active participation in online discussions and final grades. This further confirms the moderating role of teaching presence in the association between social presence and final learning outcome.

Insert Table 6 about here

Findings from Garrison and Cleveland-Innes (2005) highlighted the importance of teaching presence (of either teachers or peers) for creating the transition from social to cognitive presence. On the other hand, Vrasidas and McIsaac, (1999) and Woo and Reeves (2007) argued

that students benefit only from meaningful interaction with their peers and teachers. Analyzing the moderation role of teaching presence in the association between social presence and academic performance (RQ2), we confirmed both views. With the lack of support for meaningful interactions (control group), students tend to compliment the work of their peers, without sufficiently deep engagement and understanding of the problem under study - i.e., students tend to behave politely, but without necessary knowledge construction gains. However, with the course design that offered scaffolding that increased the level of social presence in online discussions as shown in our study and the high level of leadership by a course teacher as shown by Garrison and Cleveland-Innes (2005), it seems that students tend to engage into deep and meaningful learning (treatment group), increase the degree of meaningful interaction, and enhance potential knowledge construction benefits. This would further imply that the students, who were deeply engaged into discussions, by building on the knowledge of their peers, were able to obtain better grades. This conclusion might suggest that the cognitive presence is more dominant as a predictor of academic performance. Building on the theoretical assumptions of the CoI model and theorized interconnectedness between three presences (Garrison et al., 1999), social presence is facilitated by teaching presence while both teaching and social presence, facilitate cognitive presence. Given that cognitive presence is defined as “the extent to which the participants in any particular configuration of a community of inquiry are able to construct meaning through sustained communication” (Garrison et al., 1999, p. 89), the actual learning indicators are expressed through cognitive presence. Further, if the learning outcomes in a course are aligned with online discussions, the association of cognitive presence and grades should exist. Thus, the role of cognitive presence is essential and it is not surprising that the specific indicators of social presence (i.e., continuing a thread) are associated with the grades. However,

as cognitive presence is facilitated by teaching and social presence, it is likely to hypothesize that the association of cognitive presence and grades is moderated by teaching presence, similarly as shown for the association between social presence and grades, as shown in this study. This hypothesis warrants further research to be validated. Finally, as Garrison and Cleveland-Innes (2005) noted, it is not the quantity of interaction that matters, but rather the quality (i.e., higher number of messages with the traces of social presence (Garrison & Akyol, 2013)). Both groups of students (control and treatment) had almost identical numbers of messages (*N*=845 and *N*=902, respectively). However, in the case of increased teaching presence, students generated a significantly higher number of indicators of social presence (Figure 2).

It should be noted that the students in both conditions were graded and the proportion of the grade was identical in both conditions. As Rovai, (2007) suggests, grading is usually recognized as a structural element for designing effective discussions. However, as is has been shown in this and other studies (Pawan, Paulus, Yalcin, & Chang, 2003) without clear guidelines, students tend to exchange commentaries without reflective (critical) responses to peers' posts. Consistent with this, as shown by Gašević et al. (2015), the scaffold helped the students to significantly increase their level of cognitive and this paper also showed an increase in social presence as the result of the scaffold. This is consistent with the theory around and empirical evidence supporting the community of inquiry model (Shea et al., 2010; Shea & Bidjerano, 2009).

In order to identify the extent to which the results of our study could be generalized, certain limitations need to be acknowledged. The study was conducted with a dataset obtained from a fully online course in software engineering. However, in order to increase the external validity of the study, it is important to analyze social presence within a course from other

educational domains (e.g., social sciences or linguistics) and levels of education (undergraduate vs. graduate) (Garrison et al., 2010). Furthermore, it is also important to conduct similar studies with courses with different instructional designs in the same subject domain. More precisely, analyzed courses primarily supported discussion forums. However, the use of some other additional technologies (e.g., video conferencing) might lead to different results (Bernard et al., 2009). Finally, the data was collected at the different period of time (i.e., two course offerings in the controlled group data were collected in 2008 and four course offerings in treatment groups were collected in 2009, 2010, and 2011). These temporal differences in the course offerings and group assignments could compound the results and warrant future research.

The direct implications of the findings of our study are twofold. First, student generated content can provide instructors with a valid and important information about the level of students' social engagement. Based on the level of social engagement within an online learning community, we can offer teachers predictors of their students' academic performance. Based on these indicators, the teachers can tailor instructional interventions for each individual student, if needed. Moreover, all discussions analyzed were generated between weeks three and five of a 12 weeks course. The students in the fourth quartile in the control group and the first quartile in the treatment group revealed had mean values of their grade point around 2.7, which was the minimum requirement for the course to be counted towards the degree program. This further means that such information might help an early identification of students at risk of failing a course. Second, our results show that teaching presence moderates the effect of social presence on academic performance. Unclear expectations regarding online discussions could have a negative effect on students' engagement and course grades, while well-structured discussions and

facilitation could induce knowledge construction benefits, which are associated with the increased academic performance.

Within the study reported, we analyzed online course through several offerings. However, following current research trends in online and distance education (Zawacki-Richter & Anderson, 2013), and building further on our most recent studies (Joksimović et al., 2015; Skrypnyk, Joksimović, Kovanović, Gašević, & Dawson, 2015), it would be highly relevant to analyze applicability of these findings in different settings – e.g., massive open online courses (MOOCs). For example, learning in distributed MOOCs usually assumes use of blogs, Facebook, Twitter and other social media as communication channels (Yeager, Hurley-Dasgupta, & Bliss, 2013), that does not necessarily support (or require) knowledge construction through the four phases of practical inquiry. Moreover, development of social presence in such diverse and large networks is likely harder, and establishing a trustworthy community could be challenging. Further, the existence of teaching presence construct as defined by the CoI model is questionable, having in mind the scale of MOOCs. Therefore, our further research will be focused on assessing applicability of the existing online learning models (e.g., CoI model) for research of teaching and learning within the larger scope (i.e., MOOCs).

References

Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported {F2F} and online learning. *Computers in Human Behavior*, *31*(0), 542 – 550. doi:http://dx.doi.org/10.1016/j.chb.2013.05.031

Akyol, Z., Arbaugh, J. B., Cleveland-Innes, M., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. (2009). A Response to the Review of the Community of Inquiry Framework. *Journal of Distance Education*, *23*(2).

Alstete, J. W., & Beutell, N. J. (2004). Performance indicators in online distance learning courses: a study of management education. *Quality Assurance in Education*, *12*(1), 6–14. doi:10.1108/09684880410517397

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing Teaching Presence in a Computer Conferencing Context. *Journal of Asynchronous Learning Networks*, *5*, 1–17.

Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A., & Bethel, E. C. (2009). A Meta-Analysis of Three Types of Interaction Treatments in Distance Education. *Review of Educational Research*, *79*(3), 1243–1289. doi:10.3102/0034654309333844

Cobb, S. C. (2009). Social Presence and Online Learning: A Current View from a Research Perspective. *Journal of Interactive Online Learning*, *8*(3), 241–254.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Dabbagh, N. (2007). The online learner: Characteristics and pedagogical implications. *Contemporary Issues in Technology and Teacher Education [Online Serial]*, *7*(3).

D. Randy Garrison, & J.B. Arbaugh. (2007). Researching the community of inquiry framework: Review, issues, and future directions. Elsevier Inc.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. SAGE Publications.

Finnegan, C., Morris, L. V., & Lee, K. (2009). Differences by Course Discipline on Student Behavior, Persistence, and Achievement in Online Courses of Undergraduate General Education. *Journal of College Student Retention: Research, Theory and Practice*, *10*(1), 39–54. doi:10.2190/CS.10.1.d

Garrison, D., Anderson, T., & Archer, W. (1999). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, *2*(2-3), 87–105.

Garrison, D. R. (2011). *E-learning in the 21st century: A framework for research and practice*. Taylor & Francis.

Garrison, D. R., & Akyol, Z. (2013). The Community of Inquiry Theoretical Framework. In G. M. Moore (Ed.), *Handbook of Distance Education* (pp. 104–120). New York: Routledge.

Garrison, D. R., & Anderson, T. (2000). Critical inquiry in text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, *2*(2), 87–105.

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

Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating Cognitive Presence in Online Learning: Interaction Is Not Enough. *American Journal of Distance Education*, *19*(3), 133–148. doi:10.1207/s15389286ajde1903_2

Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). 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. doi:http://dx.doi.org/10.1016/j.iheduc.2009.10.002

Gašević, D., Adescope, O., Joksimović, S., & Kovanović, V. (2015). Externally-facilitated Regulation Scaffolding and Role Assignment to develop Cognitive Presence in Asynchronous Online Discussions. *The Internet and Higher Education*.

Ginns, P., & Ellis, R. (2007). Quality in blended learning: Exploring the relationships between on-line and face-to-face teaching and learning. *The Internet and Higher Education*, *10*(1), 53 – 64. doi:http://dx.doi.org/10.1016/j.iheduc.2006.10.003

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

Järvelä, S., & Hadwin, A. F. (2013). New Frontiers: Regulating Learning in CSCL. *Educational Psychologist*, *48*(1), 25–39.

Joksimović, S., Kovanović, V., Jovanović, J., Zouaq, A., Gašević, D., & Hatala, M. (2015). What do cMOOC participants talk about in Social Media? A Topic Analysis of Discourse in a cMOOC. Presented at the The 5th International Learning Analytics and Knowledge (LAK) Conference (Accepted).

Jorge, I. (2010). Social presence and cognitive presence in an online training program for teachers of Portuguese: relation and methodological issues. *IODLandICEM 2010 Joint Conference and Media Days*, 427–435.

Kear, K. (2010). Social presence in online learning communities. In *Proceedings of the 7th International Conference on Networked Learning (NLC2010)* (pp. 541–548).

Kehrwald, B. (2008). Understanding social presence in text based online learning environments. *Distance Education*, *29*(1), 89–106.

Kozan, K., & Richardson, J. C. (2014). Interrelationships between and among social, teaching, and cognitive presence. *The Internet and Higher Education*, *21*(0), 68 – 73. doi:http://dx.doi.org/10.1016/j.iheduc.2013.10.007

Kutner, M., Nachtsheim, C., & Neter, J. (2004). *Applied Linear Regression Models* (4th ed.). New York: McGraw-Hill/Irwin.

Liu, S. Y., Gomez, J., & Yen, C. J. (2009). Community college online course retention and final grade: Predictability of social presence. *Journal of Interactive Online Learning*, *8*(2), 165–182.

Lopez, M. I., Luna, J., Romero, C., & Ventura, S. (2012). Classification via Clustering for Predicting Final Marks Based on Student Participation in Forums. *International Educational Data Mining Society*.

Macfadyen, L. P., & Dawson, S. (2010). Mining {LMS} data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, *54*(2), 588 – 599. doi:http://dx.doi.org/10.1016/j.compedu.2009.09.008

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

Palmer, S., Holt, D., & Bray, S. (2008). Does the discussion help? The impact of a formally assessed online discussion on final student results. *British Journal of Educational Technology*, *39*(5), 847–858. doi:10.1111/j.1467-8535.2007.00780.x

Pawan, F., Paulus, T. M., Yalcin, S., & Chang, C.-F. (2003). Online Learning: Patterns of Engagement and Interaction among In-Service teachers. *Language Learning & Technology*, *7*(3), 119.

Picciano, A. G. (2002). Beyond student perceptions: Issues of interaction, presence, and performance in an online course. *Journal of Asynchronous Learning Networks*, *6*(1), 21–40.

Ready, D. D. (2008). Class-Size Reduction: Policy, Politics, and Implications for Equity. Equity Matters. Research Review No. 2. *Campaign for Educational Equity, Teachers College, Columbia University*.

Reio, T. G., & Crim, S. J. (2013). Social Presence and Student Satisfaction as Predictors of Online Enrollment Intent. *American Journal of Distance Education*, *27*(2), 122–133. doi:10.1080/08923647.2013.775801

Richardson, J. C., & Swan, K. (2003). Examing social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks*, *7*(1), 68–88.

Romero, C., López, M.-I., Luna, J.-M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. *Computers & Education*, *68*, 458–472.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (1999). Assessing Social Presence In Asynchronous Text-based Computer. *Journal of Distance Education*, *14*(2), 50–71.

Rourke, L., Anderson, T., Garrison, D. R., Archer, W., & others. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education (IJAIED)*, *12*, 8–22.

Rovai, A. P. (2002). Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The Internet and Higher Education*, *5*(4), 319 – 332.

Rovai, A. P. (2007). Facilitating online discussions effectively. *Internet and Higher Education*, *10*(1), 77–88.

Shea, P., & Bidjerano, T. (2009). Community of inquiry as a theoretical framework to foster "epistemic engagement" and "cognitive presence" in online education. *Computers & Education*, *52*(3), 543 – 553. doi:http://dx.doi.org/10.1016/j.compedu.2008.10.007

Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., … Rangan, P. (2010). A re-examination of the community of inquiry framework: Social network and content analysis. *The Internet and Higher Education*, *13*(1–2), 10 – 21. doi:http://dx.doi.org/10.1016/j.iheduc.2009.11.002

Short, J., Williams, E., & Christie, B. (1976). *The Social Psychology of Telecommunications*. Hoboken, NJ: Wiley.

Skrypnyk, O., Joksimović, S., Kovanović, V., Gašević, D., & Dawson, S. (2015). Roles of course facilitators, learners, and technology in the flow of information of a cMOOC. *The International Review of Research in Open and Distributed Learning*, submmited for review.

So, H.-J., & Brush, T. A. (2008). Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education*, *51*(1), 318 – 336. doi:http://dx.doi.org/10.1016/j.compedu.2007.05.009

Strijbos, J.-W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education*, *46*(1), 29 – 48.

Swan, K., & Ice, P. (2010). The community of inquiry framework ten years later: Introduction to the special issue. *The Internet and Higher Education*, *13*(1), 1–4.

Swan, K., & Shih, L. F. (2005). On the nature and development of social presence in online course discussions. *Journal of Asynchronous Learning Networks*, *9*(3), 115–136.

Tao, Y., & Florida, U. of C. (2009). *The Relationship Between Motivation and Online Social Presence in an Online Class*. University of Central Florida. Retrieved from http://books.google.ca/books?id=0NO2gxlxq7EC

Tausczik, Y. R., & Pennebaker, J. W. (2010). CoI-1: The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(I), 24–54. doi:10.1177/0261927X09351676

Tu, C.-H. (2001). How Chinese Perceive Social Presence: An Examination of Interaction in Online Learning Environment. *Educational Media International*, *38*(1), 45–60. doi:10.1080/09523980010021235

Tu, C.-H. (2002). The Measurement of Social Presence in an Online Learning Environment. *International Journal on E-Learning*, *1*(2), 34–45.

Ubon, A. N., & Kimble, C. (2004). Exploring Social Presence in Asynchronous Text-Based Online Learning Communities (OLCS). In *Proceedings of the 5th International Conference on Information Communication Technologies in Education*. Samos, Greece.

Vrasidas, C., & McIsaac, M. S. (1999). Factors Influencing Interaction in an Online Course. *American Journal of Distance Education*, *13*(3), 22–36.

Weaver, C. M., & Albion, P. (2005). Momentum in online discussions: the effect of social presence on motivation for participation. In *Proceedings of the 22nd Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ascilite 2005)* (pp. 703–706).

Wise, A., Chang, J., Duffy, T., & del Valle, R. (2004). The Effects of Teacher Social Presence on Student Satisfaction, Engagement, and Learning. In *Proceedings of the 6th International Conference on Learning Sciences* (pp. 568–575). Santa Monica, California: International Society of the Learning Sciences.

Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and Higher Education*, *10*(1), 15–25. doi:10.1016/j.iheduc.2006.10.005

Yeager, C., Hurley-Dasgupta, B., & Bliss, C. A. (2013). cMOOCs and global learning: An authentic alternative. *Journal of Asynchronous Learning Networks*, *17*(2), 133–147.

Yoo, J., & Kim, J. (2012). Predicting learner's project performance with dialogue features in online q&a discussions. In *Intelligent Tutoring Systems* (pp. 570–575). Springer.

Zawacki-Richter, O., & Anderson, T. (2013). *Online Distance Education: Towards a Research Agenda*. AU Press.

Table 1. Indicators of social presence (Garrison, 2011)

Category	Indicators (Label)	Definition	Percent Agreement*
Interpersonal communication	Affective expression (IE)	Conventional expressions of emotion, or unconventional expression of emotion, include repetitious punctuations, conspicuous capitalization, emoticons.	84.4%
	Self-disclosure (IS)	Presents details of life outside of class, or expresses vulnerability.	93.1%
	Use of humor (IH)	Teasing, cajoling, irony, understatements, sarcasm.	84.1%
Open communication	Continuing a thread (OCt)	Using reply feature of software, rather than starting a new thread.	98.9%
	Quoting from others' messages (OQ)	Using software features to quote others entire message or cutting and pasting selections of others' messages.	95.4%
	Referring explicitly to others' messages (OR)	Direct references to contents of others' posts.	92.7%
	Asking questions (OA)	Students ask questions of other students or the moderator.	89.4%
	Complimenting, expressing appreciation (OCa)	Complimenting others or contents of others' messages.	90.7%
	Expressing agreement (OE)	Expressing agreement with others or content of others' messages.	96.6%
Cohesive communication	Vocatives (CV)	Addressing of referring to participants by name.	91.8%
	Addresses or refers to the group using inclusive pronouns (CA)	Addresses the group as we, us, our, group.	88.8%
	Phatics, salutations (CS)	Communication that serves a purely social function; greetings, closures.	96.1%

* A measure of inter-rater reliability for each indicator of social presence.

Table 2. Mean (M) and standard deviation (SD) values for Grade Point Average (GPA), number of courses completed and days in the program for the control and treatment groups before enrolling in the course under study.

Group	GPA			# Courses			# Days in program		
	AVG	SD	p-value	AVG	SD	p-value	AVG	SD	p-value
Control	3.75	0.23	.73	5.26	2.50	.19	572.70	323.10	.06
Treatment	3.79	0.19		6.04	2.74		766.23	443.09	

Note: Due to a non-normal distribution of the variables, Mann-Whitney test was conducted to compare the difference between two groups with respect to GPA, number of courses and number of days being in program.

Table 3. Results of correlation analysis between final grade and each indicator of social presence.

Outcome variable	Indicators of social presence	Pearson's *r*	*p*-value
Final grade	Affective expressions	.18	.11
	Self-disclosure	**.21**	**.06**
	Use of humor	.18	.10
	Continuing a thread	**.34**	**.002**
	Quoting from others' messages	.10	.38
	Referring explicitly to others' messages	.07	.5
	Asking questions	**.34**	**.002**
	Complimenting, expressing appreciation	**.25**	**.02**
	Expressing agreement	.13	.25
	Vocatives	**.27**	**.02**
	Addresses or refers to the group using inclusive pronouns	.15	.18
	Phatics, salutations	.15	.16

Table 4. The results of the multiple regression analysis that tested the association between the indicators of social presence and final grades on the entire sample

Variable	R^2	B	*β*	VIF	*p*-value
Continuing a thread	0.16	0.09	.9	8.82	**.004**
Complimenting, expressing appreciation		-0.06	-.61	8.82	**.05**

R^2 - describes the overall model, i.e. how much of the variance is explained by fitting this model; B - indicates how much each predictor affected the final grades, having all other predictors constant; β - the number of standard deviations by which the final grade would change, as a result of one standard deviation change in the predictor; VIF (Variance Inflation Factor) - quantifies the severity of multicollinearity; the values less than 10 indicate that there is no cause for concern.

Table 5. Multiple regression analysis results, for the control and treatment group

Variable	R^2	B	*β*	VIF	*p*-value
Control group					
Continuing a thread	0.24	0.07	.51	4.92	.29
Complimenting, expressing appreciation		-0.12	-.89	4.92	**.08**
Treatment group					
Continuing a thread	0.25	0.09	.93	9.08	**.008**
Complimenting, expressing appreciation		-0.06	-.48	9.08	.16

Table 6. Mean and standard deviation values student grades distributed into student quartiles based on two indicators of social presence for the control and treatment group.

Control group		Treatment group	
Quartile	**Avg. grade (SD)**	**Quartile**	**Avg. grade (SD)**
Q1	3.8 (0.37)	Q1	2.63 (1.61)
Q2	3.88 (0.39)	Q2	3.54 (1.09)
Q3	3.82 (0.35)	Q3	3.84 (0.42)
Q4	2.86 (1.67)	Q4	3.69 (0.55)

Note: Quartiles for the control group are created based on the count of the *complimenting, expressing appreciation* indicator, while quartiles for the treatment group are based on the count of the *continuing a thread* indicator of social presence.

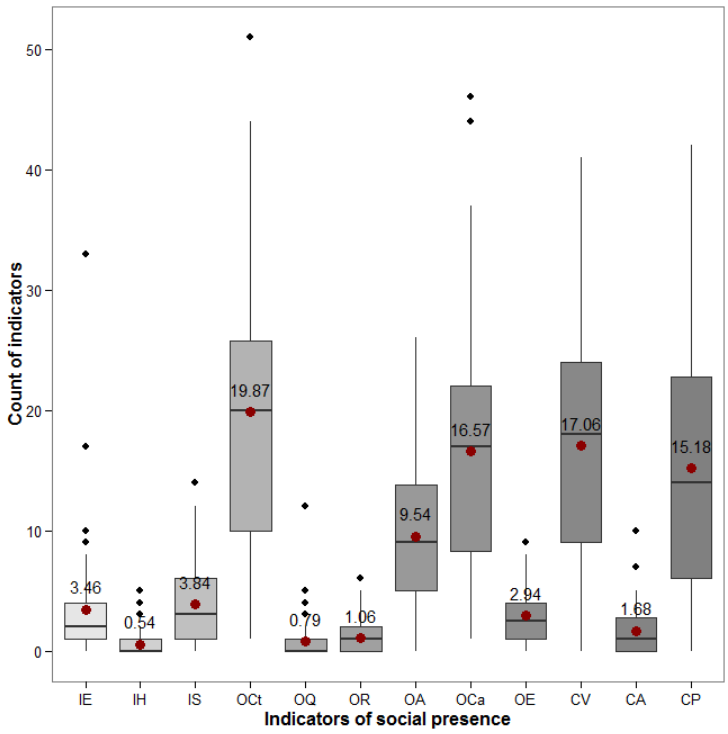

Figure 1. Descriptive statistics of the indicators of social presence (for the entire dataset – both control and treatment groups combined) with the mean values (read dot), confidence intervals, outlier values, and bar plots with 25th, 50th and 75th percentile (IE-Interpersonal, Affective; IS-Interpersonal, Self-disclosure; IH-Interpersonal, Use of humor; OCt-Open, Continuing a thread; OQ-Open, Quoting; OR-Open, Referring; OA-Open, Asking questions; OCa-Open, Complimenting; OE-Open, Expressing agreement; CV-Cohesive, Vocatives; CA-Cohesive, Addresses a group; CS-Cohesive, Phatics, Salutation).

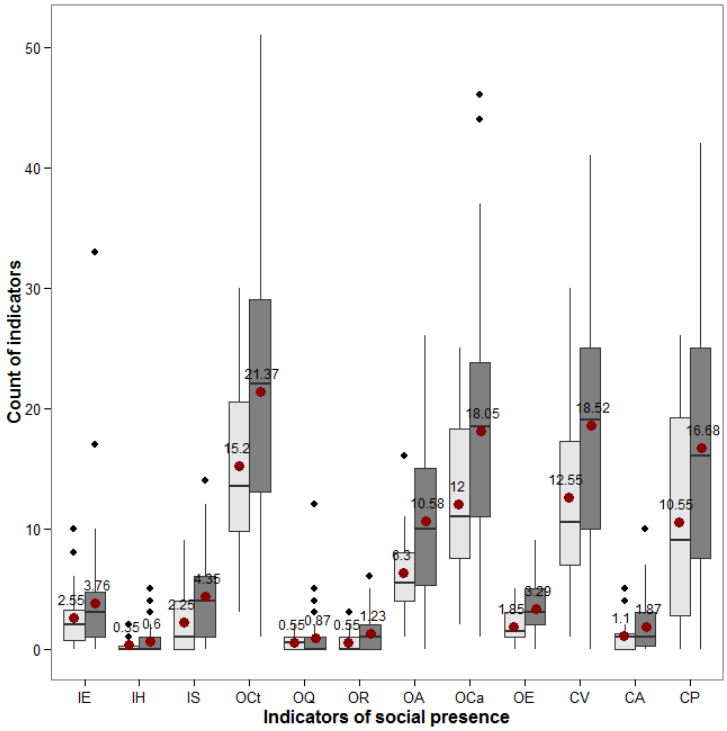

Figure 2. Descriptive statistics for indicators of social presence (for the dataset divided into the treatment and control groups), with mean values (read dot), confidence intervals, outlier values, and bar plots with 25th, 50th and 75th percentile (IE-Interpersonal, Affective; IS-Interpersonal, Self-disclosure; IH-Interpersonal, Use of humor; OCt-Open, Continuing a thread; OQ-Open, Quoting; OR-Open, Referring; OA-Open, Asking questions; OCa-Open, Complimenting; OE-Open, Expressing agreement; CV-Cohesive, Vocatives; CA-Cohesive, Addresses a group; CS-Cohesive, Phatics, Salutation).