Exploring development of social capital in a cMOOC through language and discourse

Citation for published version:
Joksimovic, S, Dowell, N, Poquet, O, Kovanovic, V, Gasevic, D, Dawson, S & Graesser, A 2017, 'Exploring development of social capital in a cMOOC through language and discourse' *Internet and Higher Education*, vol. 36, pp. 54-64. https://doi.org/10.1016/j.iheduc.2017.09.004

Digital Object Identifier (DOI):
10.1016/j.iheduc.2017.09.004

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Internet and Higher Education

EXPLORING DEVELOPMENT OF SOCIAL CAPITAL IN A CMOOC THROUGH LANGUAGE AND DISCOURSE

Srećko Joksimović[1], Nia Dowel[2], Oleksandra Poquet[1], Vitomir Kovanović[1], Dragan Gašević[3], Shane Dawson[1], Arthur C. Graesser[4]

[1]The University of South Australia, Adelaide, Australia
[2]The University of Michigan, Michigan, USA
[3]The University of Edinburgh, Edinburgh, United Kingdom
[4] The University of Memphis, Memphis, USA

ABSTRACT

Connectivist pedagogies are geared towards building a network of learners that actively employ technologies to establish interpersonal connections in open online settings. In this context, as course participants increasingly establish interpersonal relationships among peers they have greater opportunity to draw on and leverage the latent social capital that resides in such a distributed learning environment. However, to date there have been a limited number of studies exploring how learners build their social capital in open large-scale courses. To inform the facilitation of learner networks in open online settings and beyond, this study analyzed factors associated with how learners accumulate social capital in the form of learner connections over time. The study was conducted in two massive open online course offerings (Connectivism and Connective Knowledge) that were designed on the principles of connectivist pedagogy and that made use of data about social interaction from Twitter, blogs, and Facebook. For this purpose, linear mixed modelling was used to understand the associations between learner social capital, linguistic and discourse patterns, media used for interaction, as well as the time in the course when interaction took place. The results highlight the association between the language used by the learners and the creation of ties between them. Analyses on the accumulation of connections over time have implications for the pedagogical choices that would be expected to help learners leverage access to potential social capital in a networked context.

Keywords: MOOC, Social capital, Social network analysis, Linguistics, Discourse, Connectivism

EXPLORING DEVELOPMENT OF SOCIAL CAPITAL IN A CMOOC THROUGH LANGUAGE AND DISCOURSE

The importance of peer interactions for the learning process has been a consistent narrative in all forms of education. Research in the distance courses, online and blended courses, and more recently in open scaled courses in distributed environments have all stressed the need for developing peer to peer interactions to promote student learning and achievement of course goals (Bernard et al., 2009; Borokhovski, Tamim, Bernard, Abrami, & Sokolovskaya, 2012; Joksimović, Gašević, Loughin, Kovanović, & Hatala, 2015). As a new educational provision within online education, Massive Open Online Courses (MOOCs) have triggered heated media and academic discussions about a range of issues. For instance, there has been much debate over the validity of learning in such an open scaled environment as well as the challenges in establishing online interpersonal interactions at scale without losing a more socially oriented learning model (Gašević, Kovanović, Joksimović, & Siemens, 2014; Reich, 2015; Reich, Stewart, Mavon, & Tingley, 2016). The technical transition to learning at scale resulted in a need for existing pedagogical models to move beyond mere transmission of teacher-produced content. The capacity to deliver online course to the masses requires the ability to scale learner centric pedagogies in new ways that enable the production of social interactions among thousands of learners (Stewart, 2013).

The first MOOCs – today commonly known as connectivist MOOCs (cMOOCs) – emerged as an innovative solution to scaling learner interactions. They were designed as an alternative to the more conventional online education practices that delivered content via a single (centralized) platform. That is, conventional online education is, and remains, constrained in the number of opportunities readily available to learners to connect outside of teacher-controlled systems. In addressing this limitation, facilitators of the first cMOOCs scaled learner interactions by using diverse media for sharing, aggregating, and connecting information. In cMOOCs, learners were encouraged to interact with each other on the basis of personal goals and common interests (Mcauley, Stewart, Siemens, & Cormier, 2010). Establishing social ties with other learners mediated by technology was thought to be integral to the learning process (Anh, Butler, & Alam, 2013; Knox, 2014).

The connectivist model of learning (Siemens, 2005) assumes there is an untapped abundance of information that resides in distributed networks. The connectivist model perceives technology as distributed, courses less structured and without formal assessment, while the teaching is focused on instructional design and learner facilitation (Siemens, 2005). Knowledge was approached as distributed among the network of learners, whereas learning was viewed as the development and maintenance of networks of information, resources and contacts (Anderson & Dron, 2011). The main premise for learning in a connectivist setting is that learners form connections based on shared interests, at the same time learners are invited to explore various topics, to decide what to learn, and to choose communication media that are best suited to their needs (Mcauley et al., 2010).

Although online educators and researchers have explored and critiqued the theoretical grounds of connectivist courses (Bell, 2010), there remains a paucity of empirical research providing evidence of how such learning would unfold in the pedagogical context of connectivism. Empirical insights into learning in cMOOCs have been limited due to the technical difficulty of collecting cMOOC interactions distributed over the Internet. Consequently, the majority of cMOOC research has relied on self-report mechanisms, i.e. course evaluations, participant surveys and interviews (Fini, 2009; Kop, 2011; Kop, Sui, & Mak, 2011; Milligan, Littlejohn, & Margaryan, 2013). Observational evidence, however, should provide a more scalable approach in studying learning in connectivist settings.

In our prior work, we collected a dataset of two connectivist courses to gain insight into how learning unfolds in the pedagogical context of connectivism. For example, Skrypnyk, Joksimović, Kovanović, Gašević, and Dawson (2015) utilized observational data to capture the transition from course facilitation as primarily instructor-driven to a more learner-driven and self-organized model - the central pedagogical characteristic of cMOOCs (Siemens, 2010). The results demonstrated that as the number and density of students' connections in a network increased in the course there was an associated transition in power and control from facilitator to student. In essence, the growing network structure resulted in, some participants securing a network position that gave them "power and control" over the information flow in the course that was on par with the original course facilitators (teachers).

The current study further contributes to our understanding of learning in connectivist settings. It investigates factors associated with a successful learning experience from a connectivist perspective. Within the connectivist pedagogy, learning outcomes are not pre-defined by a facilitator. The creation of network links, or physically establishing connections from learner to learner, is considered learning in the sense that it enables faster access to new information and resources (Siemens, 2005). Connecting to another person opens access to different kinds of benefits, unavailable if the connection is not made. In this sense, a learner's position in the network represents the potential to learn from the network, due to their level of access to informational resources, personal support and/or professional opportunities that are embedded within the entire course network.

A learner's position in a social network is also reflective of the available social capital a learner can draw upon to support their learning endeavors (Haythornthwaite & De Laat, 2012). Individual social positioning at varying time points in a course can indicate the level of access to social capital and how this can influence successful participation in an open course. Such an approach is theoretically rooted within the network theory of social capital by Lin (Lin, Cook, & Burt, 2001). According to Lin, social capital is defined as a personal investment into building network connections (Lin, Cook, et al., 2001) that can be accessed to aid achievement of individual goals. Access to social capital is well captured and typically operationalized through the measures of network centrality as commonly used in social network analysis (Lin, Cook, et al., 2001; Lin, Fu, & Hsung, 2001) (SNA). Network measures incorporate both the number of connections made, and opportunities and limitations available to an individual due to the positions they occupy within a social network (Burt, 2000).

This study explored the factors related to the development of social capital of learners in the three main social media software (i.e., Twitter, Facebook, and blogs) used in two connectivist MOOCs (i.e., CCK11 and CCK12). Social capital was measured through centrality measures derived from social network analysis. We used linear mixed effects modeling to investigate whether the development of social capital is associated with how learners utilize language for communication, as measured through different linguistic and discourse features (Graesser, Mcnamara, & Kulikowich, 2011). To account for contextual factors that may mediate the association between learner discourse and social capital, linear mixed models included (a) the effects of social media through which interactions occurred, (b) the overall amount of learner activity and (c) the time in the course when interactions took place. The paper builds on the previous research presented in the Joksimović and colleagues (2015) study to offer a comprehensive analysis of factors that influence the development of social capital in online courses facilitated by social media.

THEORETICAL BACKGROUND

Social capital

Contemporary definitions of social capital can vary significantly. Despite the diversity of interpretations there is general agreement that social capital represents an investment in social relations for some future expected returns (Lin, 1999). Given the context of our research (i.e., studying learning in distributed online/networked settings), we adopted Lin's (2008) definition of social capital. Observed through the lens of three families of social concepts discussed by Paldam (2000), Lin's definition stems from the *network* family, implicitly building on the concept of *network payoff* that conceptualize social capital as being equal to the amount of benefits *one* can draw on his network. In essence, Lin's (2008) definition, interprets social capital from the perspective of individual network actors as they create new connections that enable them to access the resources embedded in the broader network structure. In contrast Bourdieu (1986) and Putnam (1993) for example, view social capital at a group-level (e.g., Bourdieu, 1986; Coleman, 1988; Putnam, 1993). This perspective privileges strong ties that are associated with *collective assets* (Williams & Durrance, 2008), such as solidarity, trust, reciprocity, and norms, to establish a longer term membership developed through network cohesion.

Social networking sites enable for the creation of both weak and strong ties. In his seminal work, Granovetter (1973) distinguished between *strong* (e.g., friends, family) and *weak* (e.g., acquaintances) social ties and showed evidence for the importance of weak social ties on the access to novel information resources. Early work on online communities hypothesized that the Internet, besides being used for maintaining strong social ties, also affords cost and time effective ways of maintaining weak social ties that can be potentially used for informational resources and/or access to opportunities (Liou, Chih, Hsu, & Huang, 2015; Yoo, Choi, Choi, & Rho, 2014). A recent review of evidence connecting social networking platforms (e.g. Twitter, Facebook, and various blogging platforms) with social capital concluded that social network sites are well suited for development, accumulation, and conversion of social capital, i.e., mobilization of social capital for a specific return (Ellison & Vitak, 2015). Furthermore, it has been suggested (Ellison, Wohn, Khan, & Fewins-Bliss, 2012) that social networking sites enable the creation of weak or strong ties from activated latent ties, i.e. the ties that are "technically possible but not activated socially"

(Haythornthwaite, 2005, p.137). In the context of cMOOCs and networks of learners, it is the activation of latent ties that affords an opportunity to leverage new information and resources in order to achieve desired learning gains evolving from the relationships with peers.

In building on Lin's definition, Gaag & Snijders (2003) proposed that measuring social capital should be limited to the access to resources, without accounting for the actual *use* of social ties. Gaag & Snijders (2003) argued that measuring social capital beyond structural access requires accounting for wider contexts beyond those that can be measured. By applying SNA at the level of network actors, the individual access to potential resources can be captured through SNA metrics (Borgatti, Jones, & Everett, 1998). Borgatti and colleagues reviewed network metrics and their hypothetical association with social capital. For example, an individual's degree, i.e. the number of connections, is theorized as positively related to social capital as individual gain; the more people an individual is connected to, the higher the likelihood that one of these connections will have potentially necessary information. In addition to degree centrality, in this study we adopted eigenvector, betweenness and closeness centrality. These measures are commonly used indicators that can provide a more in-depth, multi-dimensional assessment of the available social capital (Borgatti et al., 1998).

Contexts for social capital development

Contextual factors influence the way learners gain access to the available pool of social capital. For instance, students exercise different degrees of activity, convey information in different linguistic styles, and apply media that afford differing modes of interaction. Similarly, the time in the course when interactions take place is potentially important. All these contextual factors may be correlated with students developing and mobilizing their perceived social capital. These contextual factors are frequently observed across various educational courses. In this study, *learner activity*, *time of course*, *language* and chosen *social media* are the considered contextual factors in the analysis of how learners develop access to social capital in a network.

Language and discourse. Language is a primary means for expressing and exchanging content through a network. It is through language that participants are able to build connections and define social ties with other actors. With regard to analytical approaches, there has been extensive knowledge gleaned from manual content analyses of learners' discourse during educational interactions. For instance, the early research of Bernstein (1971) highlighted that individuals with more complex social networks tend to demonstrate more formal and elaborated speech forms than those with more simple and densely connected personal networks. Milroy and Margrain (1980) reported that the variety of language in use is dependent on the density of the social network and the multiplexity of the ties. According to Granovetter (1973), the intensity of ties established between actors affords an opportunity to track the linguistic phenomenon of code-switching, whereby speakers change conversational styles as they converse with interlocutors from the different parts of their sub-networks. These earlier studies illustrate the relationship between social ties and language. However, the manual content analysis methods used in those studies are no longer a viable option with the increasing scale of educational data. Consequently, researchers have been incorporating automated linguistic analysis that range from shallow level word counts to deeper level discourse analysis.

To extend analysis of learning-related phenomena beyond word count measures, one needs to conduct a deeper level discourse analysis with sophisticated natural language processing techniques, such as syntactic parsing and cohesion computation. For example, Dowell, Cade, Tausczik, Pennebaker, and Graesser (2014) explored the extent to which discourse features predicted student performance during computer-mediated collaborative learning interactions in groups of 4 students. Their results indicated that students who generated language with deeper cohesion and more complicated syntactic structures had higher performance scores on tests. Dowell and colleagues (2015) used a similar methodological design in their investigation of student performance in a MOOC. Specifically, they explored the extent to which characteristics of discourse diagnostically reveals leaners' performance and social position in a MOOC. Their results for performance mirrored the pattern that was observed for learning in the computer-mediated collaborative learning study (Dowell et al., 2015). Specifically, students who performed significantly better engaged in more expository style discourse, with higher referential and deep level cohesion, more abstract language, and more simple syntactic structures (Graesser, McNamara, & Kulikowich, 2011). However, linguistic profiles of the centrally positioned learners differed from the high performers. Learners with a more significant and central position in their social network generated a more narrative discourse style with less cohesion among ideas, as well as more simple syntactic structures and abstract words (Dowell et al., 2015). Based on these findings, the linguistic characteristics of learners may provide a promising approach for understanding the factors that lead to the formation of social ties among a group of learners.

In the current research we adopt a multilevel theoretical approach to the analysis of language and discourse. Psychological models of discourse comprehension and learning, such as the construction-integration, constructionist, and indexical-embodiment models, lend themselves nicely to the exploration of learning related phenomena in computer-mediated educational environments. These psychological frameworks have identified the representations, structures, strategies, and processes at multiple levels of discourse (Graesser & McNamara, 2011; Kintsch, 1998; Snow, 2002). Five levels have frequently been identified in these frameworks: (1) words, (2) syntax, (3) the explicit textbase, (4) the situation model (sometimes called the mental model), and (5) the discourse genre and rhetorical structure (the type of discourse and its composition). The computational linguistic facility used in the correct study, Coh-Metrix (described more in the methods), allows us to capture these main levels of discourse. In the learning context, learners can experience communication misalignments and comprehension breakdowns at different levels. Such breakdowns and misalignments have important implications for the learning process.

Social media. The social media (Twitter, Facebook, Blog) used by the learners in a course is also an important factor influencing interactions. Different social networking software have been known to impact the flow of information and community formation (Gruzd, Wellman, & Takhteyev, 2011). For example, Backstrom, Huttenlocher, Kleinberg, and Lan (2006) reported that community formation in large social networks depends on the structure of the underlying network. More precisely, the growth of communities does not depend on the relationships that an individual has within a network, but rather on the type and strength of these relationships. The use of media has also been shown to be related to the depth of ties connecting communicators (Haythornthwaite, 2002), where more weakly tied communicators rely on organizationally established means for exchanging information. Finally, Androutsopoulos (2006) has argued that the studies focusing on the diversity of language use in computer mediated communication, over time have shifted from “medium-related to user-related patterns of language use” (p.421). This suggests that different communication media (e.g., e-mail, blogs and chat) should be observed in terms of technological affordances that constrain discourse styles within the social media (Androutsopoulos, 2006).

Time. Previous studies on online learning have emphasized the relevance of the temporal dimension in the analysis of learning-related processes (Barbera & Reimann, 2014; Kovanović et al., 2015; Reimann, 2009). Integrating longitudinal data into statistical analyses can provide insights into micro-processes, developmental sequences, phases, and time scale durations (Chiu et al. in Barbera & Reimann, 2014). For example, the development of social presence in the community of inquiry framework has been connected with time (Akyol & Garrison, 2008), showing that, as the course progresses, students undergo a transitional phase from social presence to cognitive presence. This process is in line with the mainstream premise of small groups research that social structures evolve sequentially (Arrow, Poole, Henry, Wheelan, & Moreland, 2004). As another example, missing the early time for peer discussion may impact performance and drop-out, as demonstrated in face-to-face settings (Vaquero & Cebrian, 2013) as well as online interaction in MOOC research (Rosé et al., 2014). Due to these important implications, we measured the sequence of weeks in the courses under investigation.

Learner activity. The assumption that activeness of an individual reflects interest and motivation is often used in xMOOC studies, where trace data on course resources is correlated with student perseverance or academic achievement (DeBoer & Breslow, 2014). "Activeness" is also relevant to understanding how social capital is developed and accumulated (Skrypnyk et al., 2015). In their analysis of a network emerging from a cMOOC, Skrypnyk and colleagues (2015), identified a group of so-called prolific learners, characterized by their high out-degree. This group of learners' author text more frequently compared to their peers. Similarly, a group of participants, called super-posters (Huang, Dasgupta, Ghosh, Manning, & Sanders, 2014) have been identified through their extensive participation in xMOOC forums. In both cases, it is not necessarily the content of the messages, but the sheer volume and frequency of the contributions that make these learners more "visible". Moreover, in the context of the cMOOC, these prolific learners over time tend to attract more people to their discussions and are often instrumental to community formation. Therefore, this study measured the amount of learner contributions as one of the factors impacting the development of social capital.

RESEARCH QUESTIONS

The goal of the current research is to understand the influence of a broad suite of contextual factors in the development of social capital in a connectivist MOOC (cMOOC). Specifically, we investigate the role of language, media, time, and learners' activeness on centrality.

Communication is a primary means of exchanging information in emerging educational environments, like MOOCs, and as such it plays a critical and complex role (Dowell et al., 2015). The current study approaches the analysis of linguistic features used by MOOC participants and participants' overall engagement as a method to gain insights regarding the quality of ties formed between the learners. Additionally, because the relationship between learners occurs over time, it is difficult, if not impossible to consider learners' social position without time playing a role. Therefore, we explored temporal changes in learners' discourse and the position within the network as the course progresses. Finally, social media applications vary in their affordances for the use of language. Linguists do not approach Internet language as a fixed discourse register, despite its unique features (Crystal, 2001), but rather treat it as "resources that particular users might draw on in the construction of discourse styles in particular contexts" (Androutsopoulos, 2006, p.421). In other words, different types of media are seen as varying contexts for users to engage with. Different media types also influence the use of language and thereby help shape various discourse genres (Androutsopoulos, 2011).

Drawing on this theoretical and empirical background, we explored the following three research questions:

RQ1. How is the language used by cMOOC participants associated with the positions that define an individual's access to the social capital in the network of learners?

RQ2: What is the role of different social media on the development of the social capital?

RQ3. What are the temporal dynamics of social capital in a cMOOC?

METHOD

Data

This study examined blog, Twitter and Facebook posts from the 2011 and 2012 editions of the Connectivism and Connective Knowledge (CCK) course. These courses were designed as open online courses aiming to explore the ideas of connectivism and connective knowledge, and to examine the application of the connectivist framework in theories of teaching and learning. Both course offerings were facilitated over a 12-week period: CCK11 was delivered from January 17th, 2011 to April 11th, 2011, while CCK12 took place between January 23rd, 2012 and April 11th, 2012. Course resources were delivered using gRSShopper[1], while live sessions were carried out using Elluminate[2]. Given the specific (connectivist) nature of the course, students were not obliged to use any particular platform and/or media to interact with other students. However, course facilitators suggested students do share their insights and resources about the course content using

---

[1] http://cck11.mooc.ca/, and http://cck12.mooc.ca/
[2] https://sas.elluminate.com

technologies such as blogs, Facebook, Twitter or other discussion groups and social media. Finally, gRSShopper was used to provide students with a daily newsletter that aggregated content produced by the course participants on Twitter and their personal blogs. This method allowed automatic gathering of links to blog posts and copies of tweets. Facebook data were collected using Facebook API[3] in order to retrieve communication between course participants.

The data are publicly available from the respective course sites. Moreover, the collected data are available upon request, stored in the JSON format with the following information:

- **Twitter:** authors' name, date/time created, media attached (e.g., photo, video, web page), mentions, and hashtags;
- **Blogs:** authors' name, date created, title, URL, as well as posted comments with information about comment's author and date/time created;
- **Facebook:** besides basic information about authors' name and date/time created, Facebook posts contain all the information specified in API documentation.

To support the analysis of content created in multiple languages, messages posted in languages other than English were translated using Microsoft Translation API[4] (around 5% of messages were translated). The total numbers of posts produced in CCK11 (N_{post11}=5711, *M*=2.59, *SD*=4.47) and CCK12 (N_{post12}=2951, *M* =3.41, *SD*=9.06) differed, with CCK12 having fewer active students (N_{cck11}=997, N_{cck12}=429)[5]. However, despite a smaller cohort the participants demonstrated a higher average activity. The difference in activity can also be seen through the comparison of the volume of posts made on Facebook ($N_{post11f}$=1755, $N_{post12f}$=61) and blogs ($N_{post11b}$=1473, $N_{post12b}$=624) in both courses. Twitter-mediated communication sustained similar high levels of activity for both courses ($N_{post11t}$=2483, $N_{post12t}$=2266).

Analyses

In order to address the research questions, SNA was first conducted to calculate centrality measures defining the structural positions of individual learners in the networks for each course. Next, algorithms behind the Coh-Metrix principal components (described later) were applied to

---

[3] https://developers.facebook.com
[4] http://msdn.microsoft.com/en-us/library/dd576287.aspx
[5] Number of students for courses under study, represents the number of active students that participated in communication using three social media platforms analyzed.

calculate measures representing linguistic and discourse features of individual learners' interactions. All measures were calculated on a week-to-week basis in order to address the third research question. Finally, statistical analyses were performed to identify whether the linguistic features of learners' interactions, social media used, temporal dimension, and learners' activities were associated with their structural positions. A linear mixed effect model was conducted statistically assess the contributions of the alternative media, time, and learner activeness as well as the variance attributable to differences among individuals.

Social Network Analysis. Twitter, blogs and Facebook were the most widely used media for interacting in each course. Therefore, 72 undirected weighted graphs were constructed to represent interactions independently mediated by these three technologies for each week of each course. That is, each of the two courses included three networks that were formed from the different media types. These networks were constructed 12 times (one per week) for each of the three social media (i.e., Twitter, Facebook, and blogs) within the course. *Twitter graphs* included all authors and mentions as nodes of the network, whereas the edges between them were created if an author or an account were tagged within the tweet. For example, if a course participant @Learner1 mentioned @Learner2 and @Learner3 in a tweet, then the course Twitter network would contain @Learner1, @Learner2, and @Learner3 with the following edges: @Learner1 – @Learner2, and @Learner1 – @Learner3. Network graphs representing interactions in blogs and on Facebook included authors of the posts, i.e., blog owners or Facebook post initiators, as well as authors of comments to either of these. If a learner A1 created a blog or Facebook post, and then learners B1 and C1 added comments to that post, then the corresponding network would contain nodes A1, B1, and C1 with the following edges: A1-B1, and A1-C1. Graphs for each week included authors who posted and/or commented within the given week only.

Principles and methods of graph theory have been commonly used to assess the values of different network positions (Wasserman & Faust, 1994). Of particular importance is the notion of centrality that is commonly used to capture the importance of an individual node in the network (Wasserman & Faust, 1994). Therefore, the following well-established SNA measures (Freeman, 1978; Wasserman & Faust, 1994) were calculated for each learner in all network graphs:

- **Degree Centrality** – the number of edges a node has in a network;
- **Eigenvector Centrality** – the measure of influence of a given node;

- **Closeness Centrality** – the distance of an individual node in the network from all the other nodes;
- **Betweenness Centrality** – the number of shortest paths between any two nodes that pass via a given node.

The social network variables were analyzed using *igraph 0.7.1* (Csardi & Nepusz, 2006), a comprehensive R software package for complex social network analysis research.

Linguistic analysis. For linguistic analysis, the texts produced by individual learners via different media were parsed in weekly chunks. For example, all text produced by Learner 1 on Twitter in week 1 of CCK11 was treated as one unit, while all text produced by the same learner on Facebook in week 1 of CCK11 was treated as another unit. To analyze discourse patterns on multiple levels, we used Coh-Metrix, arguably the most comprehensive automated textual assessment tool currently available on the Web (Graesser et al., 2011; McNamara, Graesser, McCarthy, & Cai, 2014).

Coh-Metrix is a computational linguistics facility that analyzes higher-level features of language and discourse (Graesser et al., 2011; McNamara et al., 2014). Coh-Metrix has been used to analyze texts in K-12 for the Common Core standards and states throughout the U.S. (Arthur C Graesser et al., 2014; Nelson, Perfetti, Liben, & Liben, 2012). More than 50 published studies have demonstrated that Coh-Metrix indices can be used to detect subtle differences in text and discourse (McNamara et al., 2014). The Coh-Metrix website[6] provides over 100 measures at multiple levels, including genre, cohesion, syntax, words and other characteristics of language and discourse. Coh-Metrix also has measures of linguistic complexity, characteristics of words, and readability scores. There was a need to reduce the large number of measures provided by Coh-Metrix into a more manageable size. This was achieved in a study that examined 53 Coh-Metrix measures for 37,520 texts in the TASA (Touchstone Applied Science Association) corpus, which represents what typical high school students have read throughout their lifetime (Graesser et al., 2011). A principal components analysis was conducted on the corpus, yielding eight components that explained an impressive 67.3% of the variability among texts; the top five components explained over 50% of the variance. Importantly, the components aligned with the language-discourse levels previously

[6] www.cohmetrix.com

proposed in multilevel theoretical frameworks of cognition and comprehension (Graesser & McNamara, 2011; Kintsch, 1998; Perfetti, 1999; Snow, 2002) and thus are suitable for investigating trends in learning-oriented conversations.

In this study, the following five principal components of Coh-Metrix were calculated for each of the units (Table 1):

- **Narrativity.** The extent to which the text is in the narrative genre, which conveys a story, a procedure, or a sequence of episodes of actions and events with animate beings. At the other end of the continuum are more informational texts.
- **Deep Cohesion.** The extent to which the ideas in the text are cohesively connected at a deeper conceptual level that signifies causality or intentionality.
- **Referential Cohesion**. The extent to which explicit words and ideas in the text are connected with each other as the text unfolds.
- **Syntactic Simplicity**. Sentences with few words and simple, familiar syntactic structures. Polar opposite are structurally embedded sentences that require the reader to hold many words and ideas in their working memory.
- **Word Concreteness.** The extent to which content words are concrete, meaningful, and evoke mental images as opposed to abstract words.

Table 1. The summary statistics of the linguistic measures (z-scores) used in the study.

Name	Average	St. Dev.	Min	Max	Median
Narrativity	-0.920	1.672	-7.410	4.660	-0.580
Deep Cohesion	-0.099	1.394	-4.730	26.560	-0.180
Ref. Cohesion	-0.747	3.482	-17.100	10.100	-0.750
Syn. Simplicity	-0.230	3.068	-5.260	11.330	-0.870
Word Concreteness	-1.423	2.337	-7.600	14.580	-1.320

Statistical analysis. A mixed-effects modeling approach was adopted for all analyses due to the repeated measurements and nested structure of the data. Specifically, learners were nested within the courses in our analyses. Mixed-effects modeling is a recommended method for analyzing such datasets (Pinheiro & Bates, 2000). Mixed-effects models include a combination of fixed and random effects and can be used to assess the influence of the fixed effects on dependent variables after accounting for any extraneous random effects. Fixed effects correspond to the numerical or categorical variables that are of primary interest and represent fixed, repeatable levels among which comparisons are to be made. Random effects are categorical variables that represent variability among subjects, a random selection from a larger population to which the results can be extended.

A mixed-effects modeling approach yields a stringent test of the contributions of language, media, time, and learners' activeness on centrality by controlling for the variance associated with individual students and course differences. More specifically, this approach allows for testing our primary questions of interest, namely the correlation contributions of language characteristics, the media used, and time on social capital (measured via the four centrality measures) in an online educational environment. Therefore, four different linear mixed-effects models were constructed, one for each of the centrality measures. Within each model one centrality measure (i.e., degree, eigenvector, betweenness, and closeness) was considered as a dependent variable. The independent fixed effect variables included five Coh-Metrix *principal components*, *media* (Twitter, Facebook, and Blogs), and *week sequence* to assess any potential temporal influences on linguistic properties. The *count of posts* was incorporated to take into account the relative activeness of course participants. To address the impact of individual variance within a model, learners within a course and a course were treated as random effects.

Several steps were taken in relation to the choice of mixed effects regression models. For each of the dependent variables we constructed three models (Table 3): (a) a *null model* with the random effect only (*student within a course*), (b) a *fixed effects model* that included the random effect, as well as Coh-Metrix principal components, media (Twitter, Facebook, and Blogs), week, and post count as fixed effects, and (c) a *full model* that introduced *course random slope* to account for variability at the course level. A comparison of the null model with the *centrality* models determined whether language predicts social dynamics above and beyond the random effects. Intraclass Correlation Coefficient (ICC), (Raudenbush & Bryk, 2002), Second-order Akaike Information Criterion (AICc) and a likelihood ratio test (Hastie, Tibshirani, & Friedman, 2009) were used to decide on the best fitting and most parsimonious model. The ICC is commonly used in the model building process to determine the strength of the non-independence or the necessity of additional random variables. In the present study, we started with a simple random intercept model for student within course. The ICC was used to assess the value added by using a more complex model that allowed slopes to vary as well as intercepts. The ICC and AICc likelihood ratio tests indicated the more complex random intercept and slope significantly improved the degree and eigenvector models, but not the closeness or betweenness models (Table 2). We also estimated an effect size (R^2) for each model as goodness-of-fit measures, calculating the variance explained using the method suggested by Xu (2003).

Linear mixed-effects models were conducted using R v.3.0.1 software for statistical analysis with package lme4 (Bates, Mächler, Bolker, & Walker, 2015). The hypotheses specify the direction of the effect, however two-tailed tests were used for significance testing with an alpha level of .05. Model fit assessment and fixed effects for all models are discussed below and reported in Table 2 and Table 3, respectively.

RESULTS

Degree centrality

A likelihood ratio test indicated that the *full model* yielded a significantly better fit than the *null* and *fixed effects* model (see Table 2). The linear mixed-effects analysis revealed a significant main effect for *Narrativity,* $F(1, 3097.20) = 4.51$ $p = .034$, *Referential Cohesion,* $F(1, 2867.70) = 30.97$, $p < .001$, *Syntax Simplicity,* $F(1, 3089.20) = 4.32$, $p = .038$, *Week*, $F(1, 3089.30) = 24.69$, $p < .001$ and *Posts Count,* $F(1, 1733.80) = 1792.98$, $p < .001$, whereas *Deep Cohesion,* was marginally significant, $F(1, 3089.00) = 3.31$, $p = .069$. Specifically, individuals that acquired higher degree centrality expressed themselves using more conversational style discourse with less overlap between words and ideas (i.e. low referential cohesion), more complex syntactic structures, but more deep level cohesive integration (i.e. positive relationship with deep cohesion) (Table 3). Learners with higher activity levels (i.e., those who simply posted more) had higher degree centrality scores. Moreover, as the course progressed, learners tended to connect with their peers less often. We also observed a significant effect of media used, $F(2, 2833.10) = 84.00$, $p < .001$. The results indicated that course participants accumulated higher degree centrality scores within Facebook and Twitter social networks compared to the networks extracted from blogs (Table 3). The effect was probed further by exploring pairwise comparisons of least square means. There were significant differences in the accumulation of degree centrality between blogs and Facebook, $t(3031.20) = 10.42$, $p < .001$, 95% CI [0.40, 0.59], and blogs and Twitter, $t(2765.50) = 11.23$, $p < .001$, 95% CI [0.34, 0.48]. There was no significant difference between Facebook and Twitter, $t(2723.70) = -1.85$, $p = .060$, 95% CI [-0.18, 0.005].

Table 2. Inferential statistics for the model fit assessment

Degree						
	χ^2	**Df**	R^2	**AICc**	**ICC**	
					student	**course**
Null model			.42	8200.07	.20	
Fixed model	1534.51***	12	.58	6750.03	.11	
Full model	125.34***	14	.62	6629.10	.13	.22
Eigenvector						
	χ^2	**df**	R^2	**AICc**	**ICC**	
					student	**course**
Null model			.36	8379.50	.17	
Fixed model	667.10***	12	.39	7793.68	.08	
Full model	53.41***	14	.43	7744.36	.05	.19
Betweenness						
	χ^2	**df**	R^2	**AICc**	**ICC**	
					student	**course**
Null model			.30	8492.89	.12	
Fixed model	368.66***	12	.33	8204.03	.10	
Closeness						
	χ^2	**df**	R^2	**AICc**	**ICC**	
					student	**course**
Null model			.27	8663.12	.12	
Fixed model	162.39***	12	.26	8579.34	.10	

Note: χ^2 values show the differences between the model in the current row and the model in the previous row. Significance codes: *** p < .001

Eigenvector centrality

The likelihood ratio test between the null, fixed effects, and full model revealed a significantly better fit of the model that accounted for variation of students within different courses (Table 2). The model (see Table 3) showed a significant negative effect of Referential Cohesion, F(1, 2736.60) = 15.25, p < .001 and Week, F(1, 3081.30) = 6.88, p = .009, whereas the effect of Post Count, F(1, 2156.30) = 429.13, p < .001 was significant and positive. Similar to degree centrality, learners who exhibited lower scores of referential cohesion and created higher numbers of posts had higher eigenvector centrality values. Likewise, as the course progressed, eigenvector centrality tends to decrease. Finally, results also revealed a significant difference between media used (F(2, 2523.70) = 85.35, p < .001). Further analysis exploring pairwise comparisons of least square means showed significant differences between each pair of media: blogs vs. Facebook – t(2735.50) = 5.27, p < .001, 95% CI [0.18, 0.40], blogs vs. Twitter – *t*(2737.70) = -9.06, *p* < .001, 95% CI [-0.48, -0.31], and Facebook vs. Twitter – $t(2170.90) = -12.85$, $p < .001$, 95% CI [-0.80, -0.58].

Table 3. Analysis of the fixed effects for the models of the four measures of social centrality.

Parameters	Degree centrality β	Degree centrality SE	Degree centrality 95% CI (2.5% - 97.5%)	Eigenvector centrality β	Eigenvector centrality SE	Eigenvector centrality 95% CI (2.5% - 97.5%)
Narrativity	**0.036***	**0.011**	**0.001 - 0.044**	0.027	0.013	-0.008 - 0.043
Deep Cohesion	**0.032**	**0.008**	**0.001 - 0.031**	0.013	0.010	-0.014 - 0.025
Referential Cohesion	**-0.077*****	**0.005**	**-0.038 - 0.018**	**-0.066*****	**0.006**	**-0.036 - -0.012**
Syntax Simplicity	**-0.031***	**0.012**	**-0.048 - -0.009**	-0.009	0.014	-0.035 - 0.022
Word Concreteness	-0.006	0.004	-0.011 - 0.008	-0.012	0.005	-0.015 - 0.008
Facebook	**0.163*****	**0.048**	**0.403 – 0.594**	**0.096*****	**0.056**	**0.182 - 0.405**
Twitter	**0.197*****	**0.036**	**0.337 – 0.484**	**-0.190*****	**0.044**	**-0.484 - -0.309**
Post count	**0.604*****	**0.014**	**0.575 - 0.632**	**0.367*****	**0.017**	**0.332 - 0.403**
Week	**-0.063*****	**0.004**	**-0.026 - -0.011**	**-0.040****	**0.004**	**0.003 - 0.021**

Parameters	Closeness centrality β	Closeness centrality SE	Closeness centrality 95% CI (2.5% - 97.5%)	Betweenness centrality β	Betweenness centrality SE	Betweenness centrality 95% CI (2.5% - 97.5%)
Narrativity	-0.001	0.015	-0.030 - 0.029	0.015	0.014	-0.018 – 0.038
Deep Cohesion	0.012	0.011	-0.017 - 0.027	0.025	0.100	-0.009 – 0.032
Referential Cohesion	0.009	0.007	-0.010 - 0.017	**-0.041***	**0.006**	**-0.027 - -0.002**
Syntax Simplicity	-0.003	0.026	-0.035 - 0.030	**-0.044***	**0.015**	**-0.066 - -0.005**
Word Concreteness	-0.022	0.007	-0.020 - 0.006	-0.022	0.006	-0.020 – 0.005
Facebook	**0.242*****	**0.065**	**0.613 - 0.873**	-0.003	0.061	-0.123 – 0.121
Twitter	0.001	0.050	-0.101 - 0.099	**0.174*****	**0.047**	**0.268 – 0.457**
Post count	-0.023	0.019	-0.063 - 0.016	**0.323*****	**0.018**	**0.287 – 0.360**
Week	-0.016	0.005	-0.015 - 0.005	**-0.100*****	**0.005**	**-0.039 - -0.019**

Note: All variables are on a normal scale.

Betweenness and closeness centrality

The same models were conducted to investigate how linguistic features of computer-mediated communicative utterances predict **betweenness** and **closeness centrality**. Although in both cases a model with a random slope resulted with better overall goodness-of-fit measures (AICc, R^2, and ICC), the solution for random effects revealed a perfect negative correlation between random effects specified. This outcome indicates that the model overfit the data (Baayen, 2008). Therefore, models with random slope were discarded, and simpler models were used for analysis. Since the *closeness model* did not reveal any significant effect of linguistic properties measured (Table 3), it is not further reported in the paper.

For the betweenness model, the likelihood ratio test between the *null model* and *full model* indicated a better fit of the model that included fixed and random effects (Table 2). The fitted model revealed a significant negative effect of *Referential Cohesion*, $F(1, 3083.80) = 5.37$, $p = .020$, *Syntax Simplicity*, $F(1, 3100.60) = 5.31$, $p = .021$, and temporal factor (*Week*), $F(1, 3097.10) = 37.19$, $p < .001$, as well as a significant positive effect of the *Posts Count*, $F(1, 2482.00) = 311.47$, $p < .001$. Course participants who tended to use *simple linguistic constructs* with higher *referential cohesion* had lower betweenness centrality, while the increase in the *count of posts* was positively associated with the higher betweenness centrality (Table 3). It is important to note that *week* is also negatively associated with betweenness centrality. This might be due to the fact that students tended to engage less often with their peers towards the end of the course. The *media* used also yielded a significant effect on the values of betweenness centrality ($F(2, 2782.20)= 35.75$, $p<.001$) (Table 3). Further analysis using a pairwise comparison of least square means revealed significant differences between Twitter and blogs ($t(2847.40) = 7.69$, $p < .001$, 95% CI [0.27, 0.45]) and between Twitter and Facebook ($t(2652.70) = 6.09$, $p < .001$, 95% CI [0.25, 0.48]).

Time and Linguistic features

When we conducted an analysis of variance/co-variance matrix of fixed effects within the four models, we further observed the correlations among fixed effects. All models yielded low or zero correlations between linguistic features, such as Narrativity, Deep Cohesion, Referential Cohesion, Syntax Simplicity, Word Concreteness, and week of the course when they were measured. More precisely, correlation coefficients for the all the models varied from 0.003 to 0.130 (absolute values). The low correlations among the five Coh-Metrix components is compatible with the principal components analysis conducted on the normative TASA corpus which treated each principal component as orthogonal to the other components (Graesser et al., 2011). We are aware that there are other approaches for assessing the relationships among predictor variables in the analysis, but it was compatible with the claims on the orthogonality of the components and it also shows that linguistic properties did not change over time. On the other hand, it is interesting to note that the highest correlation was observed between the temporal factor and Referential Cohesion, – $r=-.13$, for all of the models. Therefore, a more sensitive statistical approach is needed to further assess the temporal changes in linguistic properties.

DISCUSSION

Interpretation of results with respect to research questions

The goal of the current research was to explore the influence of a broad suite of contextual factors in the development of social capital in a cMOOC. First, we adopted a computational linguistics methodology to identify the linguistic profiles associated with social capital. Further, we examined the temporal dynamics of social capital and whether social capital is influenced by any variations in communication media (i.e., Facebook, Twitter, and Blogs) as well as the amount of participant activity.

We observed that both the amount of activity (number of posts) and deep level linguistic characteristics play a role in learner interactions. This finding suggests there is a need for an analysis of the surface level characteristics and a more systematic and deeper analysis of the discourse in order to obtain a comprehensive understanding of the linguistic properties and learners' activities that are associated with the high volume of social connections. Clearly, a learners' level of activity is an important factor. As one might expect, more active learners are likely to grow their influence over the flow of information in a network, and eventually interact with other well-connected participants. This is reflected in the positive relationship between the number of posts and degree centrality, eigenvector, and betweenness centrality.

A deep linguistic analysis of the interactions also showed that language and discourse features of written messages in cMOOC environments also play an important role in the development of learners' social capital (RQ1). The results indicate that learners with more connections had a linguistic profile that is more narrative with lower referential cohesion and more complex syntax. However, deep cohesion and word concreteness were not consistently significant. Interestingly, discourse with higher narrativity, lower referential cohesion, and more complex syntax is characteristic of oral language and stories rather the academic language of expository text (Graesser et al., 2011; Graesser et al., 2014). Stated differently, the language and discourse used by learners' with more social capital has a more conversational style, which is suitable when speech participants have high common ground (Clark, 1996) and the material is easier to process.

Within the realm of social interaction, the "common ground" perspective is a widely accepted theoretical framework of communication (Knapp & Daly, 2002). Common ground refers to the knowledge and beliefs communicators assume each other shares. In the conversational context, this shared knowledge includes information that captures group membership, co-present experience, and previous shared interactions (Brennan & Clark, 1996; Clark & Brennan, 1991; Knapp & Daly, 2002). For example, individuals in an interaction are able to infer that they share several types of knowledge on the bases of being in a particular MOOC together, observing the same course content, or maintaining a record of what has been previously discussed. According to Clark and Brennan's framework, common ground plays a central role in determining many aspects of the interaction between individuals, including the communication style (Clark, 1992, 1996; Clark & Clark, 1977; Clark & Wilkes-Gibbs, 1986; Horton & Gerrig, 2005; Schober & Brennan, 2003).

The principal of least effort is one element of Brennan and Clark's communication framework that seems to have a particular relevance to learners' discourse in cMOOCs. The principal of least effort posits that achieving and maintaining common ground is an effortful activity for discourse participants, who have a propensity to minimize this effort. Specifically, the least effort principal maintains that individuals use the least amount of cognitive or linguistic effort needed to successfully communicate their message (Brennan & Clark, 1996; Clark & Krych, 2004; Clark & Wilkes-Gibbs, 1986). In these studies, effort is not an all-or-nothing process, but operates in different degrees. How much effort is needed to accomplish and maintain common ground in a given situation is defined by the grounding criterion (Brennan & Clark, 1996; Clark & Brennan, 1991), i.e., the degree of grounding shared by referents that is sufficient for the immediate purposes. For example, suppose two previously unacquainted individuals discuss their political views. The interaction likely demands more effort to be properly grounded, i.e., reconciled with the existing common ground. In contrast, it would be much easier and require fewer resources to convey the same information in a conversation between a 30-year married couple who have accumulated a considerable common ground.

There are interesting interpretations for the current study from the perspective of Clark and Brennan's Common Ground framework. In the context of this theoretical framework, the interaction between cMOOC participants is a form of collective action requiring participants to coordinate on content and on process (Brennan & Clark, 1996). Coordination on content requires that participants have or develop a shared understanding of what is the object of discussion. Learners that are more centrally located compared to less centrally located students, share more common ground with a larger proportion of other learners. Therefore, a centrally located social position reduces the grounding cost, i.e. the effort needed to build mutual understanding during communication. This would support our results showing learners with more social capital have a more conversational style, with less referential cohesion, but still maintain a deeper cohesive structure to their communication. At the other end of the spectrum, learners' with less social capital may need to compensate for the lack of common ground between their self and peers by using more cohesive, expository style discourse, which requires more effort.

Below, we provide an illustrative example, from the current dataset, of this relationship between the linguistic features of language and social centrality indicated by four SNA measures. One can compare the text produced two learners, L1 and L2, both participating in course discussions on Facebook.

L1

1. I was thinking about "originality" and Connectivism a bit (http://bit.ly) and found this rather challenging. I'd like to hear other people's views on what "originality" means in a connectivist world. What "uniqueness" does Connectivism allow?
2. Academics are like all other social groups, they tend to cluster around opinions (and counter-opinions). Trouble is to find the middle-ground where opinion cultures meet. This is where productive debate can happen. Compared to the "strong" opinionated camps (for or against) this middle-ground often appears as a rather small zone, with participants always walking the thin line.

L2

1. Great resource center… thank you, @L3
2. "A candle loses nothing by lighting another candle ..." ~ Mohammed Nabouss, Libyan journalist who was recently killed in Benghazi
3. Thank you for the post ... I had misfiled my url listing :-)

Both learners had the same level of activity, i.e. both made 4 posts. It is apparent that L1 uses a more oral narrative style and a lower referential cohesion, but there were longer sentences that afford more complex syntax. L1 was "better positioned" within the network of learners, indicated by higher degree (L1 – 8, L2 – 3), eigenvector (L1 – 0.75, L2 – 0.27), closeness (L1 – 0.01, L2 – 0.008), and betweenness centrality (L1 – 47.25, L2 – 14.67). In contrast, L2 had a more expository style with shorter sentences that pack in more factual content that is referentially connected.

The case of L1 and L2 also illustrates the mobilization of social capital for achieving a specific return (i.e., learning outcome). We observed how learners L1 and L2 were developing social capital over nine weeks of the course. As mentioned, L1 was "better positioned" within the network of learners, with the higher values of degree, eigenvector, betweenness, and closeness centrality. According to our assumptions, L1 had developed higher social capital throughout the course. The activation of their social capital was nicely shown in week 10, in which learner L1 received 13 replies and 2 "likes" on a post to the Facebook group. In contrast, L2 received no replies and only 1 "like". This happened, despite the fact that both posts have been seen by almost 100 peer learners, indicating a high number of latent ties, and yet, L1 was able to activate more connections.

We explored how differences in Twitter, Blogs and Facebook might mediate the development of network positions (RQ2). Although the analyses did not reveal a significant difference between Twitter and Facebook affordances, blogs did appear to cater to the development of connections within a narrower group of people. Such findings can be related to the differences in technological affordances for interactivity, and resonate with the studies on the use of language in different media. For example, Twitter is found to have a potential for *conversationalilty* (Purohit, Hampton, Shalin, & Amit, 2013), where communicative exchanges show cross-turn coherence online, and can be defined as sustained, topic-focused and person-to-person (Honey & Herring, 2009). This would suggest that the communicative affordances embedded in Twitter enables a higher number of simple, person-to-person conversations among unknown people.

Besides the obvious higher effort required to strike a casual conversation via somebody's blog, in contrast to Twitter, commenting on a blog post or creating a blog post implies more vulnerability and readiness for self-disclosure and indicates a higher degree of commitment and interest than tweets, which are limited to a maximum of 140 characters. However, it would be premature to discard blogs as an appropriate tool for connective courses due to their lower affordances for social capital. Further studies are needed to identify the strength of the interactions mediated through blogs, since blogs linked to each other, tend "to converse" more actively in the entries and comments, if they are on closely-related topics (Herring et al., 2005, p. 9). Such future studies may indicate that blogs are suitable for quality conversations with fewer and more familiar people (i.e., develop strong ties). Simply put, conversations around blogs will occur once social presence is established and the relationships between learners is based on a certain level of mutual trust (Garrison, Cleveland-Innes, & Fung, 2010).

Our findings also show that temporal dimension (RQ3) has a significant impact on the development of the social capital throughout the course. It seems reasonable to expect that social capital increases over time, along with the quantity and the strength of one's connections. However, our study showed that the most significant "contribution" to the development of the social capital is achieved within the first few weeks of the course, as indicated with the negative association between temporal factor and the four-centrality measures analyzed. This might be due to the decreased amount of student interaction as a course progresses. On the other hand, having more connections does not mean that all of them are equally influential. We also observed that learners tend to connect with less influential peers overtime. A possible interpretation might be that course participants are not able to identify peers with similar interests from the commencement of the course. Consequently, there is a tendency to initially connect with course facilitators and those highly influential others. As the course progresses and the interactions evolve participants become more familiar and therefore manage to activate some of their latent ties (Haythornthwaite, 2005), i.e. build connections with those course participants who may or may not have been prominent network participants, but are of relevance to specific individual learners. In order to enable learners to mobilize latent social ties and general knowledge in their networks, it is important to study different technological and pedagogical approaches that can assist in that process early in the course. Publishing user profiles, easily retrievable by others and making learners prior knowledge, skills, and goals is a promising venue for future research.

The measure of a learners' ability to broker information and shape the information flow had two distinct patterns. First, within the first half of the course, ability of course participants to broker information tended to increase. Second, throughout the second half of the course, these indicators decreased. Such patterns may be explained from the perspective of connectivism (Siemens, 2005) and the nature of interactions in online social networks (Kwak, Lee, Park, & Moon, 2010). It seems that in a "chaotic and ambiguous information climate created by networks"

(Siemens, 2010) at the very beginning of the course, there is a need for those who are able to share information, and frame the information flow. However, since creating connections through some social media is a low-effort activity, once learners have identified peers with similar interests, they form social groups around common topics, and the importance of central brokers tends to decrease.

Implications for Research and Practice

Our research suggests that linguistic analysis methodologies and monitors of learners' activity can be leveraged to determine a learner's position within a network and be used to help foster peer connections. It is no surprise that being an active participant of the learning process yields better outcomes, and in the case of cMOOCs, the skill of interacting with others more actively can predict an increase in learners' overall social capital. However, further investigations need to examine the "characteristics" of individual learners that not only increase the development of social capital but also the mobilization of social capital for a specific return. In this case, the mobilization of social capital is to facilitate the achievement of learning outcomes. For example, a system could provide learners in a MOOC or a regular online course with support on how to coherently construct their ideas and appropriately build on other learners' ideas. Adaptive assistance within learning environments would ultimately lead to better access to social capital – a concept that is well considered to influence student satisfaction, and perceived, and achieved learning outcomes in online settings (Kovanović, Joksimović, Gašević, & Hatala, 2014; Lu, Yang, & Yu, 2013).

It appears that some environments are more effective in facilitating the development of social capital than others. Specifically, Twitter and Facebook provided better opportunities for building connections with peer learners. However, Facebook and blogs were better options when it comes to reaching the more influential learners within the network. Our analyses confirm that Twitter is the social media platform that enables the best information outreach to all the participants quickly, which is of particular importance early in the course. Although the relationship between language and the temporal dimension requires a more robust analysis than undertaken in the study reported here, it would appear that learners do not change or improve their linguistic and communication skills throughout the course. Perhaps the language and communication skills are traits that are difficult to change. Such findings may indicate that only the students who already possess well-developed connection building skills benefit from activating social capital embedded in the network. If that is the case, the connectivist course design needs to also assist students in navigating networked learning.

Social media in higher education is becoming nearly ubiquitous in the era of digital learning (Bogdanov et al., 2012). Consequently, our investigation of different social media affordances and their potential to support various types of interaction are not limited to the context of MOOCs. The implications of our findings can be transferred to the broader online learning community. Several researchers (e.g., Blaschke, 2014; Corbeil & Corbeil, 2011) have observed that social media platforms are increasingly incorporated into traditional online classroom in order to foster student interaction and support students in developing self-regulated learning skills. However, one of the main conclusions derived from this literature is that cognitive and meta-cognitive development is only partially supported by technology, whereas the synergy of pedagogy and technological affordances should provide an optimal environment for student development. The majority of evidence on the impact of social media on learning has been derived from qualitative insights on studies with small sample sizes (Blaschke, 2014). Thus, our study provides additional insights into the usefulness of various social media in supporting learning in online settings.

Future research needs to investigate different instructional scaffolds and technological affordances that will guide students to develop necessary skills for learning in networked and highly distributed environments of cMOOCs. Those skills, identified as "new media literacies" (Dawson & Siemens, 2014), should enable learners to unlock opportunities afforded by media in such distributed learning contexts. Eventual changes in the linguistic features may also provide insight into an individual's progress in the development of these literacies. On the other hand, the relationship between language used and learning in networks found in this study indicates that discourse-centric learning analytics, using measures identified within the study presented, could have an important role in creating personalized feedback. Such feedback (timely, personalized and informative) would help course participants develop new media literacies and skills associated with them such as communication and information seeking.

Limitations

The study analyzed interactions between course participants within the three most commonly used social media platforms (i.e., blogs, Facebook, and Twitter). However, some limitations need to be acknowledged. For the automated data collection process, we relied on the gRSShopper as the source for collecting links to blog posts and copies of tweets. Unfortunately, most of the tweets were no longer available through the Twitter API at the time of our data collection (April-August 2014), so we were not able to analyze interactions that would include replies, retweets, and favorites features of the Twitter platform. However, the content (including mentions and hashtags) was preserved. Finally, the study analyzed the data from courses in a specific subject domain. Given that communication in different subject domains is sometimes associated with different communication patterns, it is important to analyze social interactions within courses from a different subject domain.

CONCLUSIONS

This study investigated the context on how learners leverage access to potential social capital in two connectivist MOOCs. The analysis was conducted through linear mixed effects modeling of the relationships between learners' network positions, linguistic and discourse features of the content they created and shared; social media through which the exchanges occurred; the overall amount of learner activity; and the time in course when interactions took place. Our findings indicate that both learner-contingent factors, such as linguistic and discourse features and amount of activity, as well as pedagogy-contingent factors, such as media in use or time in the course, impact an individual's development of social capital. The implications of the study are that facilitators of distributed courses should consider a broad array of responsibilities that include and extend simple network-formation beyond shaping and leveraging the information flows throughout the learning network. In this context, cMOOC facilitators need to assist learners in choosing specific media for facilitating interactions as a best–fit for an individual learner, as well as introducing instructional elements that enhance group and individual communication skills. The study also opens up further investigation of the relationship between social ties and language in use. The findings suggest that both shallow and deep level of analyses of text need to be considered as influencing factors on the development of social ties and network structures.

Beyond the micro-context of learning in a cMOOC, the study emphasizes the learning outcomes and positional goods acquired through scaled interactions by a student of a non-accredited distributed course (Marginson & others, 2004).

REFERENCES

Akyol, Z., & Garrison, D. R. (2008). The development of a community of inquiry over time in an online course: Understanding the progression and integration of social, cognitive and teaching presence. *Journal of Asynchronous Learning Networks*, *12*(3), 3–22.

Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *The International Review of Research in Open and Distance Learning*, *12*(3). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/890

Androutsopoulos, J. (2006). Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics*, *10*(4), 419–438. http://doi.org/10.1111/j.1467-9841.2006.00286.x

Androutsopoulos, J. (2011). Language change and digital media: a review of conceptions and evidence. *Standard Languages and Language Standards in a Changing Europe. Novus, Oslo*, 145–160.

Anh, J., Butler, B., & Alam, A. (2013). Learner Participation and Engagement in Open Online Courses: Insights from the Peer 2 Peer University. *MERLOT JOLT*, *9*(2).

Arrow, H., Poole, M. S., Henry, K. B., Wheelan, S., & Moreland, R. (2004). Time, Change, and Development: The Temporal Perspective on Groups. *Small Group Research: An International Journal of Theory, Investigation, and Application*, *35*(1), 73–105. http://doi.org/10.1177/1046496403259757

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press. Retrieved from http://books.google.ca/books?id=_-YongEACAAJ

Backstrom, L., Huttenlocher, D., Kleinberg, J., & Lan, X. (2006). Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 44–54). ACM.

Barbera, E., & Reimann, P. (2014). *Assessment and Evaluation of Time Factors in Online Teaching and Learning*. IGI Global.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Bell, F. (2010). Connectivism: Its place in theory-informed research and innovation in technology-enabled learning. *The International Review of Research in Open and Distance Learning*, *12*(3). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/902

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. http://doi.org/10.3102/0034654309333844

Bernstein, B. (1971). *Class, Code and Control, Volume 1: Theoretical Studies Towards a Sociology of Language*. Boston, MA: Routledge & Kegan Paul.

Blaschke, L. (2014). Using social media to engage and develop the online learner in self-determined learning. *Research in Learning Technology*, *22*(0). Retrieved from http://www.researchinlearningtechnology.net/index.php/rlt/article/view/21635

Bogdanov, E., Limpens, F., Li, N., Helou, S. E., Salzmann, C., & Gillet, D. (2012). A social media platform in higher education. In *Global Engineering Education Conference (EDUCON), 2012 IEEE* (pp. 1–8). http://doi.org/10.1109/EDUCON.2012.6201105

Borgatti, S. P., Jones, C., & Everett, M. G. (1998). Network measures of social capital. *Connections*, *21*(2), 27–36.

Borokhovski, E., Tamim, R., Bernard, R. M., Abrami, P. C., & Sokolovskaya, A. (2012). Are contextual and designed student–student interaction treatments equally effective in distance education? *Distance Education*, *33*(3), 311–329. http://doi.org/10.1080/01587919.2012.723162

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of {Theory} and {Research} for the {Sociology} of {Education}* (pp. 81–93). New York: Greenwood.

Brennan, S. E., & Clark, H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(6), 1482–1493.

Burt, R. S. (2000). The Network Structure of Social Capital. *Research in Organizational Behaviour*, *22*, 345–423.

Clark, H. (1992). *Arenas of Language Use*. University of Chicago Press.

Clark, H. (1996). *Using Language*. Cambridge University Press.

Clark, H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC, US: American Psychological Association.

Clark, H., & Clark, E. V. (1977). *Psychology and language: an introduction to psycholinguistics*. Harcourt Brace Jovanovich.

Clark, H., & Krych, M. A. (2004). Speaking while Monitoring Addressees for Understanding. *Journal of Memory and Language*, *50*(1), 62–81.

Clark, H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39.

Coleman, J. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, *94*(Supplement: Organizations and Institutions: Sociological and Economic Approahces to the Analysis of Social Structure).

Corbeil, J. R., & Corbeil, M. E. (2011). The birth of a social networking phenomenon. In *Cutting-edge Technologies in Higher Education* (pp. 13–32). Emerald.

Crystal, D. (2001). *Language and the Internet*. Cambridge University Press.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9.

Dawson, S., & Siemens, G. . (2014). Analytics to literacies: The development of a learning analytics framework for multiliteracies assessment. *International Review of Research in Open and Distance Learning*, *15*(4), 284–305.

DeBoer, J., & Breslow, L. (2014). Tracking progress: Predictors of Students' Weekly Achievement During a Circuits and Electronics MOOC (pp. 169–170). Atlanta, Georgia, USA. http://doi.org/10.1145/2556325.2567863

Dowell, N., Skrypnyk, O., Joksimović, S., Graesser, A. C., Dawson, S., Gašević, D., … Kovanović, V. (2015). Modeling Learners' Social Centrality and Performance through Language and Discourse. Presented at the In Proceedings of the 8th International Conference on Educational Data Mining, Madrid, Spain.

Ellison, N. B., & Vitak, J. (2015). Social Network Site Affordances and their Relationship to Social Capital Processes. In S. S. Sundar (Ed.), *The Handbook of the Psychology of Communication Technology* (pp. 205 – 227). Hoboken, NJ: Wiley-Blackwell.

Ellison, N. B., Wohn, D. Y., Khan, M. L., & Fewins-Bliss, R. (2012). Reshaping access: An overview of research on access to higher education, social media and social capital. White paper, retrieved from http://www-personal.umich.edu/~enicole/ReshapingAccessToDistrib.pdf.

Fini, A. (2009). The Technological Dimension of a Massive Open Online Course : The Case of the CCK08 Course Tools. *International Review of Research in Open and Distance Learning*, *10*(5).

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. http://doi.org/10.1016/0378-8733(78)90021-7

Gaag, M. V. D., & Snijders, T. (2003). Proposals for the Measurement of Individual Social Capital. In *Creation and returns of Social Capital." London : Routledge. (draft version downloadable from http://www.xs4all.nl/~gaag/work* (pp. 199–218). Springer.

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. http://doi.org/10.1016/j.iheduc.2009.10.002

Gašević, D., Kovanović, V., Joksimović, S., & Siemens, G. (2014). Where is research on Massive Open Online courses headed? A data analysis of the MOOC Research Initiative. *The International Review of Research in Open and Distributed Learning*, *15*(5). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1954

Graesser, A. C., & McNamara, D. S. (2011). Computational Analyses of Multilevel Discourse Comprehension. *Topics in Cognitive Science*, *3*(2), 371–398. http://doi.org/10.1111/j.1756-8765.2010.01081.x

Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal*, *115*(2), 210–229.

Graesser, A., Mcnamara, D., & Kulikowich, J. (2011). Coh-Metrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher*, *40*(5), 223–234. http://doi.org/10.3102/0013189X11413260

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Gruzd, A., Wellman, B., & Takhteyev, Y. (2011). Imagining Twitter as an Imagined Community. *American Behavioral Scientist*, *55*(10), 1294–1318. http://doi.org/10.1177/0002764211409378

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd editio). New York: Springer.

Haythornthwaite, C. (2002). Strong, Weak, and Latent Ties and the Impact of New Media. *The Information Society*, (18), 385–401.

Haythornthwaite, C. (2005). Social network methods and measures for examining e-learning. *Social Networks*. http://doi.org/10.1.1.135.6993

Haythornthwaite, C., & De Laat, M. (2012). Social Network Informed Design for Learning with Educational Technology. In *Informed Design of Educational Technologies in Higher Education: Enhanced Learning and Teaching* (pp. 352–374). IGI Global. Retrieved from http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-61350-080-4.ch018

Herring, S. C., Kouper, I., Paolillo, J. C., Scheidt, L. a., Tyworth, M., Welsch, P., & Wright, E. (2005). Conversations in the Blogosphere: An Analysis “From the Bottom Up.” *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 107b–107b. http://doi.org/10.1109/HICSS.2005.167

Honey, C., & Herring, S. C. (2009). Beyond Microblogging: Conversation and Collaboration via Twitter. In *HICSS ’09. 42nd Hawaii International Conference on System Sciences, 2009.* (pp. 1–10). http://doi.org/10.1109/HICSS.2009.89

Horton, W. S., & Gerrig, R. (2005). Conversational Common Ground and Memory Processes in Language Production. *Discourse Processes*, *40*(1), 1–35.

Huang, J., Dasgupta, A., Ghosh, A., Manning, J., & Sanders, M. (2014). Superposter behavior in MOOC forums. *Proceedings of the First ACM Conference on Learning @ Scale Conference - L@S ’14*, 117–126. http://doi.org/10.1145/2556325.2566249

Joksimović, S., Gašević, D., Loughin, T. M., Kovanović, V., & Hatala, M. (2015). Learning at distance: Effects of interaction traces on academic achievement. *Computers and Education*, *87*, 204–217. http://doi.org/10.1016/j.compedu.2015.07.002

Joksimović, S., Dowell, N., Skrypnyk, O., Kovanović, V., Gašević, D., Dawson, S., & Graesser, A. C. (2015). How do you connect? Analysis of Social Capital Accumulation in connectivist MOOCs (pp. 64–68). Presented at the Proceedings of the Fifth International Conference on Learning Analytics And Knowledge (LAK'15), New York, USA: ACM. http://doi.org/10.1145/2723576.2723604

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge University Press.

Knapp, M. L., & Daly, J. A. (2002). *Handbook of Interpersonal Communication*. SAGE.

Knox, J. (2014). Digital culture clash: "massive" education in the E-learning and Digital Cultures MOOC. *Distance Education*, *35*(2), 164–177. http://doi.org/10.1080/01587919.2014.917704

Kop, R. (2011). The Challenges to Connectivist Learning on Open Online Networks: Learning Experiences during a Massive Open Online Course. *The International Review of Research in Open and Distributed Learning*, *12*(3), 19–38.

Kop, R., Sui, J., & Mak, F. (2011). A Pedagogy of Abundance or a Pedagogy to Support Human Beings? Participant Support on Massive Open Online Courses. *The International Review of Research in Open and Distance Learning*, *12*(7), 1–10.

Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). Penetrating the Black Box of Time-on-task Estimation. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 184–193). New York, NY, USA: ACM. http://doi.org/10.1145/2723576.2723623

Kovanović, V., Joksimović, S., Gašević, D., & Hatala, M. (2014). What is the source of social capital? The association between social network position and social presence in communities of inquiry. *Proceedings of the First International Workshop on Graph-Based Educational Datamining (G-EDM), London, UK*. Retrieved from http://ceur-ws.org/Vol-1183/gedm_paper03.pdf

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a Social Network or a News Media? In *Proceedings of the 19th International Conference on World Wide Web* (pp. 591–600). New York, NY, USA: ACM. http://doi.org/10.1145/1772690.1772751

Lin, N. (1999). Building a network theory of social capital. *Connections*, *22*(1), 28–51.

Lin, N. (2008). A Network Theory of Social Capital. In D. Castiglione, J. W. Van Deth, & G. Wolleb (Eds.), *The Handbook of Social Capital* (1st ed., Vol. 1, pp. 50–69). Oxford University Press.

Lin, N., Cook, K. S., & Burt, R. S. (2001). *Social Capital: Theory and Research*. Transaction Publishers. Retrieved from https://books.google.co.uk/books?id=u_KTkBHY_kgC

Lin, N., Fu, Y., & Hsung, R.-M. (2001). Measurement techniques for investigations of social capital. *LIN, N.; GRUYTER, A. Social Capital: Theory and Research. New York*.

Liou, D.-K., Chih, W.-H., Hsu, L.-C., & Huang, C.-Y. (2015). Investigating information sharing behavior: the mediating roles of the desire to share information in virtual communities. *Information Systems and E-Business Management*, 1–30. http://doi.org/10.1007/s10257-015-0279-2

Lu, J., Yang, J., & Yu, C.-S. (2013). Is social capital effective for online learning? *Information & Management*, *50*(7), 507–522. http://doi.org/http://dx.doi.org/10.1016/j.im.2013.07.009

Marginson, S., & others. (2004). Competition and markets in higher education: a "glonacal"analysis. *Policy Futures in Education*, *2*(2), 175–244.

Mcauley, A. A., Stewart, B., Siemens, G., & Cormier, D. (2010). The MOOC Model for Digital Practice, 33.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix.* Cambridge, M.A.: Cambridge University Press.

Milligan, C., Littlejohn, A., & Margaryan, A. (2013). Patterns of Engagement in a Connectivist MOOC. *Journal of Online Learning and Teaching*, *9*(2).

Milroy, L., & Margrain, S. (1980). Vernacular Language Loyalty and Social Network. *Language in Society*, *9*(1), 43–70.

Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance* (Technical Report to the Gates Foundation) (pp. 1–58). Council of Chief State School Officers, Washington, DC2012. Retrieved from http://achievethecore.org/content/upload/nelson_perfetti_liben_measures_of_text_difficulty_research_ela.pdf

Paldam, M. (2000). Social Capital: One or Many? Definition and Measurement. *Journal of Economic Surveys*, *14*(5), 629–653. http://doi.org/10.1111/1467-6419.00127

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In *In C.M. Brown, & P. Hagoort (Eds.), The neurocognition of language* (pp. 167–208). Oxford University Press.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. Springer.

Purohit, H., Hampton, A., Shalin, V. L., & Amit, P. (2013). What Kind of # Conversation is Twitter ? Mining # Psycholinguistic Cues for Emergency Coordination, *5217*, 1–23.

Putnam, R. D. (1993). *Making Democracy Work*. Princeton, NJ: Princeton University Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. SAGE Publications.

Reich, J. (2015). Rebooting MOOC Research. *Science*, *347*(6217), 34–35.

Reich, J., Stewart, B., Mavon, K., & Tingley, D. (2016). The Civic Mission of MOOCs: Measuring Engagement Across Political Differences in Forums. In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale* (pp. 1–10). New York, NY, USA: ACM. http://doi.org/10.1145/2876034.2876045

Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, *4*(3), 239–257. http://doi.org/10.1007/s11412-009-9070-z

Rosé, C. P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P., & Sherer, J. (2014). Social factors that contribute to attrition in MOOCs. *Proceedings of the First ACM Conference on Learning @ Scale Conference - L@S '14*, 197–198. http://doi.org/10.1145/2556325.2567879

Schober, M. F., & Brennan, S. E. (2003). Processes of Interactive Spoken Discourse: The Role of the Partner. In *Handbook of Discourse Processes* (pp. 123–164). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from internal-pdf://Schober & Brennan (2003)-0688762881/Schober & Brennan (2003).pdf

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, *2*(1), 3–10.

Siemens, G. Teaching in Social and Technological Networks, Connectivism (2010).

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. *International Review of Research in Open and Distance Learning*, *16*(3), 188–217.

Snow, C. E. (2002). *Reading for Understanding: Toward a Research and Development Program in Reading Comprehension*. Rand Corporation.

Stewart, B. (2013). Massiveness + Openness = New Literacies of Participation? *MERLOT Journal of Online Learning and Teaching*, *9*(2).

Vaquero, L. M., & Cebrian, M. (2013). The rich club phenomenon in the classroom. *Scientific Reports*, *3*, 1174. http://doi.org/10.1038/srep01174

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

Williams, K., & Durrance, J. (2008). Social Networks and Social Capital: Rethinking Theory in Community Informatics. *The Journal of Community Informatics*, *4*(3). Retrieved from http://ci-journal.net/index.php/ciej/article/view/465

Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine*, *22*(22), 3527–3541. http://doi.org/10.1002/sim.1572

Yoo, J., Choi, S., Choi, M., & Rho, J. (2014). Why people use Twitter: social conformity and social value perspectives. *Online Information Review*, *38*(2), 265–283.