

Instructors' Perceptions of Networked Learning and Analytics

Perceptions des instructeurs quant à l'apprentissage et l'analyse en réseau

Scott Comber, Dalhousie University

Martine Durier-Copp, Dalhousie University

Anatoliy Gruzd, Ryerson University

Abstract

This study seeks to understand instructors' perceptions of social network analysis (SNA) and network visualizations as learning analytics (LA) tools for generating useful insights about student online interactions in their class. Qualitative and quantitative data were collected from three graduate courses taught at a Canadian university at the end of the academic term and came from two sources: (1) class-wide forum discussion messages, and (2) interviews with instructors regarding their perceptions of student networks and interactions. This study is unique as it focuses on instructors' self-assessments of online student interactions and compares this with the SNA visualization. The difference between instructors' perceptions of social network interactions and actual interactions underlines the potential that LA can provide for instructors. The results confirmed that SNA and network visualizations have the potential of making the "invisible" visible to instructors, thus enhancing their ability to engage students more effectively.

Résumé

Cette étude vise à comprendre les perceptions des instructeurs sur l'analyse des réseaux sociaux (ARS) et la visualisation de réseaux comme outils d'analyse de l'apprentissage (AA) produisant des perspectives utiles sur les interactions en ligne des étudiants de leur classe. Des données qualitatives et quantitatives ont été collectées dans trois cours des cycles supérieurs d'une université canadienne à la fin de la session scolaire. Ces données proviennent de deux sources : (1) les messages du forum de discussion de l'ensemble du groupe et (2) des entretiens avec les instructeurs au sujet de leurs perceptions sur les réseaux et interactions des étudiants. Cette étude est unique en ce qu'elle se concentre sur les auto-évaluations des instructeurs portant sur les interactions étudiantes en ligne, et les compare à la visualisation de l'ARS. La différence entre les perceptions qu'ont les instructeurs des interactions sur les réseaux sociaux et les interactions réelles souligne le potentiel que l'AA peut offrir aux instructeurs. Les résultats ont confirmé que l'ARS et les visualisations de réseaux ont le potentiel de rendre « l'invisible »

visible pour les instructeurs, améliorant ainsi leur capacité à motiver les étudiants plus efficacement.

Introduction

One of the most critical challenges faced by higher education institutions across Canada continues to be student retention, an issue that results in a tremendous cost and loss to both students and universities (Chiose, 2018). Research has revealed the complex, interdependent factors contributing to a student's decision about whether to continue their education, mainly involving how well they integrate academically and socially (Tinto, 1995; Zhang, Oussena, Clark, & Hyensook, 2010). It is well understood that early identification and early, continuous, and intensive intervention can improve student retention, yet higher education programs continue to struggle with identifying those most at risk and getting them the supports they need (Chiose, 2018; Siedman, 1996). Through the study reported in this article, we seek to inform better retention strategies, approaches, and tools for student retention in online environments.

The widespread integration of digital technology into higher education affords new teaching and engagement practices that can help instructors better understand and support student learning and, by extension, better support student retention. One such tool is learning analytics (LA). Through the tracking of collaborative learning, knowledge exchange, and engagement with course content, LA tools allow instructors to identify and potentially rectify potential challenges for students early on (IBM & Campus Technology, 2012). A growing body of research suggests that LA has the potential for new insights into learning processes by making hitherto invisible patterns in education data visible to researchers and end users (Greller & Drachsler, 2012). Yet little is known about instructors' need for, or perception of, these tools. This article addresses this gap in the published research by reporting on a recent study of instructors from three graduate courses at a Canadian university.

We are especially interested in instructors' ability to perceive social networks and student interactions in their classes, and how network visualizations of student interactions may facilitate an improved understanding of how student networks and interactions evolve. Are instructors seeing the same picture, so to speak, that LA is painting? This article first presents our theoretical approach and research questions. To establish a wider context for the study, we then provide a brief literature review on the applications, benefits, and potential risks of LA, and one particular LA tool, social network analysis (SNA). Next, we detail the methods used to uncover instructors' perceptions which are explored in the results and conclusions. We close with the study's limitations and recommendations for future research that will strengthen our capacity to provide student support.

Theoretical Approach and Research Questions

This study is viewed through a connectivist lens. Connectivism presents a theory of learning that acknowledges learning is no longer an internal, individual activity; technology, social networks, and structures have had a profound effect on learning in this digital age (Siemens, 2005). Connectivism is rooted in the understanding that learning and knowledge rests not with one person; it relies on diverse viewpoints. Learning is a process of seeing, making, and

nurturing connections between information sources and ideas. The capacity to know more is more important than what is currently known (Siemens, 2005).

Two core principles of connectivist learning are of particular interest in this study. Siemens suggests a core learning process is decision-making (choosing what to learn, and making meaning of the presented information). Decisions are made in an always-evolving environment; The ability to understand the new information requires a new decision that is critical to learning (Siemens, 2005). Further, Siemens states that the intent of networked learning activities is to provide accurate, up-to-date knowledge (Siemens, 2005). One of the fundamental aims of LA is to encourage decision and prompt action towards deeper learning (Van Barneveld, Arnold, & Campbell, 2012). Connectivist learning underpins this study as we examine the effect visualized interaction data can have on instructors' interpretation of, and response to, student learning.

This study was guided by three research questions:

1. How accurately do instructors perceive student interactions online?
2. How accurately do instructors perceive the class-wide interactivity level?
3. What factors affect instructors' perception of online interactions?

To explore these questions, we determined how three instructors perceived their students' interactions through a network drawing exercise (drawn during interviews). We then compared instructor-perceived social network visualizations to the actual visualizations captured through SNA and then solicited instructors' feedback on their perceived visualizations compared to the actual visualizations.

The study's objective was to discover differences and similarities between the instructor's perceptions of online interactions among students and the actual online interactions among students. Our pre-study assumption was that showing a network visualization to instructors is needed to reduce the gap between how instructors perceive students' online interactions versus what actual online interactions happened in their online classroom. We hypothesize that there is actionable knowledge revealed in this gap; Instructors could use more accurate representations of student interactions, made visible through LA tools, to inform interventions for student success.

Literature Review

Learning Analytics

Definition. Though there is no universally accepted definition, LA is well-understood as the “measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environment in which it occurs” (Long & Siemens, 2011, p. 34). LA has been principally driven by the growth of online learning and big data. Online learning generates a wealth of text that, currently, are presented to instructors and students primarily as sequences of messages, displayed in a linear, chronological fashion, one message after the other. In these records are data on student interaction and learning that are hidden because of the difficulty of managing the text produced. Evidence that is easily available in face-to-face classes – such as who sits with whom, who is paying attention in class,

and who always asks the questions – becomes invisible in the usual streams of text presented to instructors and students. LA has the potential for new insights into learning processes by making hitherto invisible patterns in education data visible to researchers and end users (Greller & Drachsler, 2012).

Application and benefits. Research has begun to reveal LA's potential for reconceptualizing and transforming how learning is designed, and how teaching is conducted, in the virtual classroom (Czerkawski, 2015; Gruzd, Paulin, Haythornthwaite, 2016; Pardo, Ellis, & Calvo, 2015; Romero et al., 2013). Much of this research discusses LA types that strengthen student learning and assessment. Xing et al.'s (2015) work developed automated ways to assess group-level learning by applying a clustering algorithm to log data from a group learning platform. Dascalu et al. (2015) use of natural language processing proposes a framework for assessing comprehension in collaborative learning environments based on conversational data from chats and forum discussion boards. Brooks, Thompson, and Teasley's (2015) research aims to develop a student success prediction model by applying an n-gram technique to analyze educational log data of students' interactions with online resources (such as web pages and video lectures). While the body of literature on LA is growing, much is still unknown about what, when, and why specific types of LA could be useful to instructors when assessing student and class performance.

Social Network Analysis

Definition. LA can rely on different types of data sources (e.g., course website usage data, class interaction data, test scores, and student profiles) and different computational techniques (e.g., statistical analysis, text analysis, and temporal analysis) for insights. This study focuses on one particular data source – class forum discussions – and one particular analysis method – SNA. Our highly connected environment has created new learning opportunities supported through practice and engagement with our individualized ecosystems of networked people, groups, services, and other entities (Downes, 2006). These ecosystems, while beneficial for learners, are difficult for researchers' and educators' efforts to understand, assess, and design effective learning experiences (Gruzd, Paulin, & Haythornthwaite, 2016). SNA maps, measures, and visually represents those networked relationships to help us understand who is connected and why, and what enables or encumbers information and knowledge sharing (Gruzd, Paulin, & Haythornthwaite, 2016; Serrat, 2017).

Application. Researchers have used SNA to study how students collaborate and learn in online environments. SNA is particularly useful in revealing how students interact with both class content and one another in online classes where discussion boards are the primary communication platform. For instance, Aviv, Erlich, Ravid, & Geva (2003) used SNA to examine how students interacted with one another in formal (closed) discussion boards versus informal (open) ones. They found that formal discussions better facilitated knowledge construction processes and allowed for the formation of more cohesive cliques in the class than less formal discussions in open fora.

SNA has been used to demonstrate the roles of communication style and position within social networks in successful online collaboration. In a 2007 study, Cho, Gay, Davidson & Ingraffea identified a high willingness to communicate and a central position within a social

network as key factors that influence collaborative learning. One especially interesting finding from their study is that those learners who were at the periphery of the collaborative network were more likely to form new and multiple connections within the class, whereas those with pre-existing social ties tended to restrict their connections to their pre-determined network. Furthermore, the researchers were able to establish that students who were more central in the network (in other words, those who collaborated with more people) generally got higher final grades. Likewise, Gašević, Zouaq, & Janzen (2013) found a positive correlation between students with high social capital (those with greater access to others as well as information) and students with higher GPA.

Researchers have demonstrated how SNA can inform the design of online discussions and collaborative practices within courses. In their 2013 study, Thormann, Gable, Fidalgo, and Blakeslee examined the influence of the number of student moderators in an online class on the overall participation and development of critical thinking. The researchers studied three iterations of the same class in which online discussions were facilitated by one, two, or three student moderators. They found that overall participation and critical thinking measures were higher for the class with three moderators. Collins and Gruzd (2017) examined co-citation networks based on assignments that required students to locate relevant references in online classes. The study's broad goal was to develop SNA-driven metrics to assess who among the learners cited relevant but common sources, who discovered unique references, and who just missed the objectives of the assignment completely. The authors also showed how SNA can help instructors design and classify assignments from more to less structured.

Network Visualization. Network visualization is an important element of SNA and arguably one of the most effective parts of the method. Once a network is discovered, it can be visualized as a graph consisting of nodes that represent people (such as class participants) and lines between the nodes that represent various relations among them. In a network visualization based on online discussions, for example, a connection could represent who is interacting with (or replying to) whom. Haythornthwaite (2002, 2006, 2008) along with Haythornthwaite and Laat (2010) provide a robust, in-depth discussion of different types of learning networks and relations. SNA renders visible the configuration of a learning network where information and knowledge is variably dispersed across learners..

Once the network is displayed, the resulting visualization can reveal the general interactivity and cohesion of online discussions and begin to answer several questions. How does social learning take place? Who are the most and least connected students in the class? Is there a single student who is dominating discussions? What is the structural influence and importance of those who dominate as discussion initiators or responders? Does culture or language influence interactions? Are conversations clustered around a single group or many different groups of students? (Gruzd, Paulin, & Haythornthwaite, 2016). Network visualizations can also help to determine if there is a general tendency of certain types of students (e.g., those who are friends, who are former classmates, who have the same gender, or who are group or team members) to only talk to each other.

Potential Risks with Learning Analytics and Social Network Analysis Application

Researchers and practitioners have pointed out that SNA and other LA tools should be used with caution. They have raised concern about the risk of strict interpretation, lack of parity in access to data, danger of engaging with only what can be measured, and possible negative effects on learner's independent, self-regulated learning (Duval, 2011; Wise, Zhao, & Hausknecht, 2013). The construction of knowledge via networked nodes of learners in online discussion boards requires flexible, open, and accessible dialogue and interaction (Couros, 2009; Wise, Zhao, & Hausknecht, 2013). Therefore, the transparency of SNA data is a critical consideration (Dringus, 2012; Wise, Zhao, & Hausknecht, 2013).

Another concern is that LA and SNA models may display an incomplete picture or fail to include enough context for accurate interpretation. SNA can often neglect to indicate or include the contribution of so-called "lurkers", those who read a lot but post little, or more introverted types of thinkers who are slower or more reluctant to post (Dringus, 2012; Hernández-García et al., 2015; Wise, Zhao, & Hausknecht, 2013). Knight and Littleton (2016) argue that with regards to all LA techniques applied to conversation, understanding the "significance of any given utterance involves understanding the context in which it is made" (p. 115). LA models which can discern these nuances are desirable (Knight & Littleton, 2016). Gasevic et al. (2016) caution against the development and application of generalized LA models for student success, suggesting that LA will not be of practical value unless it is designed in companionship to the course it will serve, keeping sight of situated teaching and learning practice.

Addressing the Challenges

Several researchers have provided recommendations for pedagogical design to address these challenges in LA and SNA application. First and foremost, LA data need to be integrated with the goals of the learning activity or course. The activity's purpose, the instructor's expectation and understanding of productive interaction, and how learning analytics will measure and represent these must all be made explicit to the learner (Wise, Zhao, & Hausknecht, 2013). There have been successes of integration within some online academic settings. Purdue University implemented a system called Course Signals that allows instructors to provide feedback in real time to students, partly based on data collected from students' participation levels in the course (Arnold & Pistilli, 2012). Northern Arizona has a similar learning analytics project called Grade Performance Status (Picciano, 2012). This study indicated higher retention among students who received real-time intervention.

Other recommended practices include making time and space for critical reflection on LA models and introducing parity between instructor and student in terms of access to LA data (Wise, Zhao, & Hausknecht, 2013). This is consistent with best practices for digital pedagogy in general: open, social, and supported by interoperable ecosystems (Couros, 2009; Siemens, 2005). Lockyer, Heathcote, and Dawson (2013) describe a framework for aligning insights drawn from LA data with three tenets of learning design (set of resources, tasks for learners to carry out with resources, and support mechanisms). For example, instructors can ask themselves, how will SNA models affect "the sequence of learning tasks, resources, and supports that a teacher constructs for students over part of, or the entire, academic semester" (Lockyer et al., 2013, p. 1442)

Network discovery and visualization tools like SNAPP (Bakharia & Dawson, 2011), NodeXL (Hansen, Shneiderman, & Smith, 2010), and Netlytic (Gruzd, 2011) are making it easier to capture and visualize students' learning networks. LA dashboards are additional tools particularly useful for increasing parity of access to meaningful learning data. The dashboards display visual or graphical representations of historical and current learning to support decision making by providing instructors and students with a deeper understanding of learning behaviours, patterns, and engagement (Corrin & de Barba, 2015). Studies have shown promising results in terms of LA dashboards' capacity to help students interpret and respond to real-time feedback on performance (Ali et al., 2012; Corrin & de Barba, 2015; Duval, 2011; Vassileva & Sun, 2007). Researchers are emphasizing the need for the feedback provided through these dashboards to be "actionable" through building awareness and developing the sense-making of both students and instructors (Verbert, Duval, Govaerts, & Santos, 2013).

Despite the proliferation of LA and SNA tools and applications and an emerging understanding of their benefits, risks, and ways to mitigate the concerns, questions remain about instructor's perceptions of them. Are network visualizations useful to instructors and if yes, for what purposes, and what features of such networks should be paid attention to, and why? The current article builds on previous research to examine the potential use of network visualizations as an LA tool for instructors to achieve greater understanding of the social knowledge building process in their online classrooms and better position them to provide effective student support.

Method

To answer our research questions, we collected data from three graduate courses taught at a Canadian university. The classes were selected based on two criteria: the class had more than ten students, to ensure a higher volume of discussion messages; and the class relied on discussion boards available in a learning management system (LMS) for required or optional conversation and dialogue. Courses having group work were exempted because this would potentially skew the data towards inner-group connections. These three graduate classes were chosen for ease of access to LMS data, as well as their relative similarity in composition.

Prior to the beginning of the study, the authors obtained the institutional Research Ethics Board approval along with the instructors' agreement to enroll their classes in the study. The three classes enrolled in the study were in graduate level programs (Business Administration, Public Administration, and Information Management). The students in these programs had many unique characteristics. Specifically, all three programs admit only students who are considered mid-career professionals; students must have at least five years of professional experience. Most students were working full-time during their program of study. One course was taught by a full-time faculty member and the other two were taught by part-time faculty.

All three courses utilized Blackboard Learn as the primary LMS. The courses were online/blended courses of 12–14 weeks duration (online component), with a 2–4-day face-to-face session at the end of the term. Of the three, one had no grades attributed to the discussions, one had participation points worth 10%, a portion of which was attributed to the discussions, and one had points per discussion posting, for a total of 20% of the final grade. Online teaching activities comprised of managing online discussion fora, running synchronous and asynchronous online lectures, responding to student inquiries through various communication channels (email,

discussion boards, and office hours), grading online student assignments, and running the final intensive (the face-to-face portion of each class).

Once classes commenced, a description of the research project was posted on the three respective LMS, and students were invited to take part in the project. Those who did not opt to participate had their Forum data removed from the data that were collected for analysis. There were 64 potential student participants (22, 16, and 26 respectively from each of the three courses); Of those students, 49 opted to participate in the study.

The data were collected at the end of the academic term from the following sources: class-wide messages posted by students as part of their online discussions (forum data), and interviews with instructors regarding their perceptions of student networks and interactions (interview data). Table 1 contains information about the number of students, the number of study participants, and the number of messages generated by each class.

Table 1

Description of Datasets

Network	#Students	#Participants	#Messages
Class A	22	17	139
Class B	16	12	410
Class C	26	20	206

Once the data were collected, we used SNA to discover and examine social interactions and collaborative learning processes in online classes. As noted above, SNA has proven to be an effective method for studying shared knowledge construction, influence, and support behaviour among learners (Aviv, Erlich, Ravid, & Geva, 2003; Cho, Gay, Davidson, & Ingraffea, 2007; Marreiros, Santos, & Ramos, 2010; Paulin, Gilbert, Haythornthwaite, Gruzd, & Absar, 2015). The following briefly discusses how the two datasets (Forum and Interview data) were used to build interaction networks among class members.

The first step was to build interaction networks depicting *who communicated with whom* in class, based on observed data from threaded discussions. The observed networks were then used to evaluate how accurately instructors recalled and perceived in-class interactions that happened online. To represent the Forum data as a network, we used a Name Network approach, which conducts a text analysis of retrieved messages to find all mentions of personal names, and then builds an interaction network by connecting people who directly referred to (mentioned) each other by name in their forum messages. This approach of discovering interaction networks has been shown to reveal explicit and implicit connections among class participants, and to provide on average 40% more information about social ties in a group as compared to an alternative approach of connecting participants based on *who replies to whose messages*, known as a Chain Network approach (Gruzd, 2009). In the online learning context, Name networks have also been shown to provide a better reflection of perceived social ties than the Chain network; they also detect social relations that are crucial in shared knowledge construction and community building such as learning, collaboration, and requests for help. The software program Netlytic (Gruzd, 2011) was used to translate the Forum data into student interaction networks. To import

data into Netlytic, we first compiled the entirety of the discussion boards into a single text file, one file per class. Those data were then reformatted for importing into an Excel file with five fields: thread title, posted date, posted time, author, and post body. Once the data were properly formatted in an Excel document, they were imported into the Netlytic software, which identified the various fields and displayed *who communicated with whom* in each class. Figure 1 shows the resulting networks.

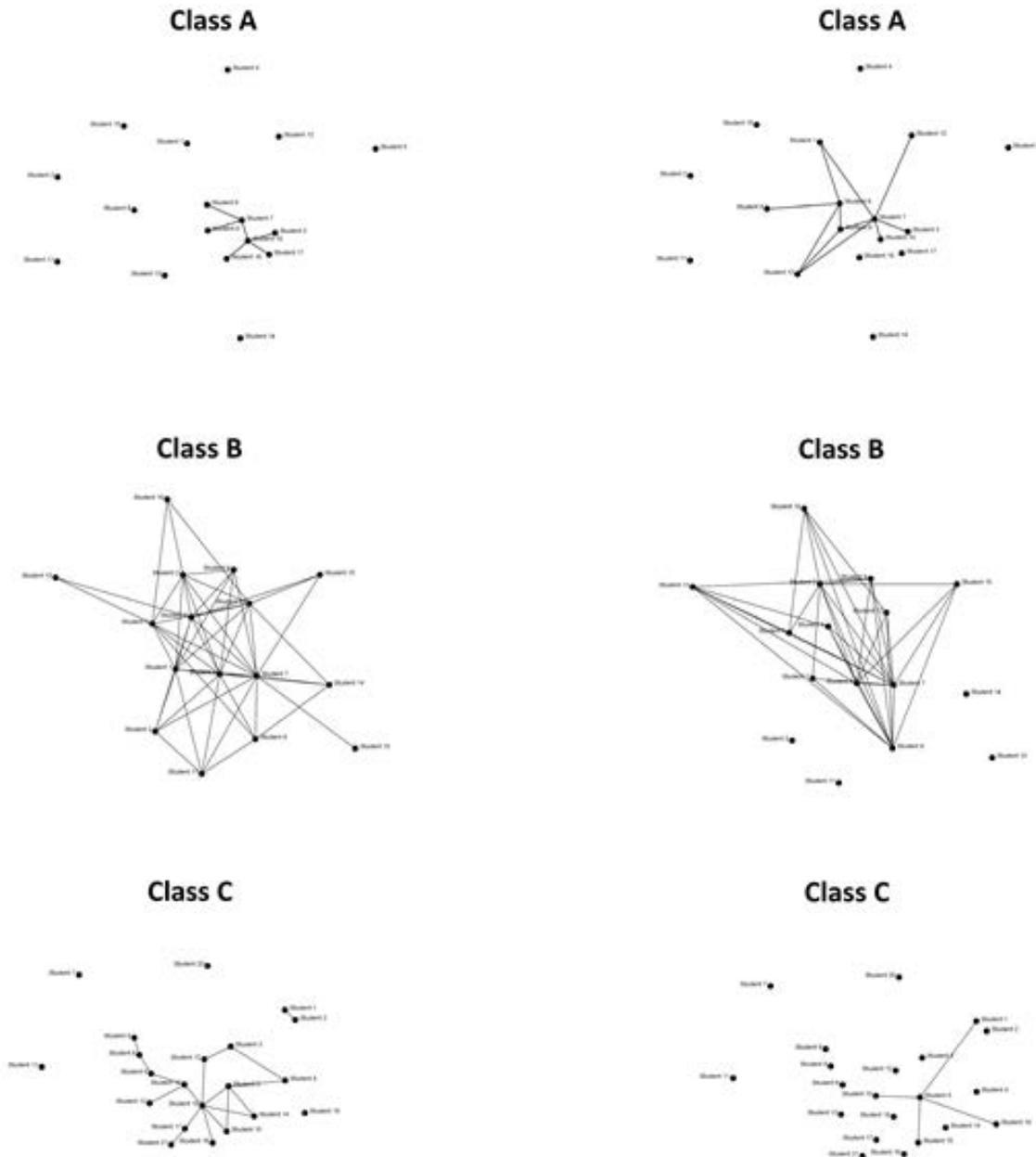


Figure 1: Forum interaction networks, based on the Name Network algorithm (left), and instructors' perceived networks (right).

The second type of network was based on the instructor’s recollection of who talked to whom in online fora. One-hour, semi-structured interviews were held with the three faculty members at the end of the academic term being studied. The interview field notes and audio recordings were compiled using Camtasia Relay software.

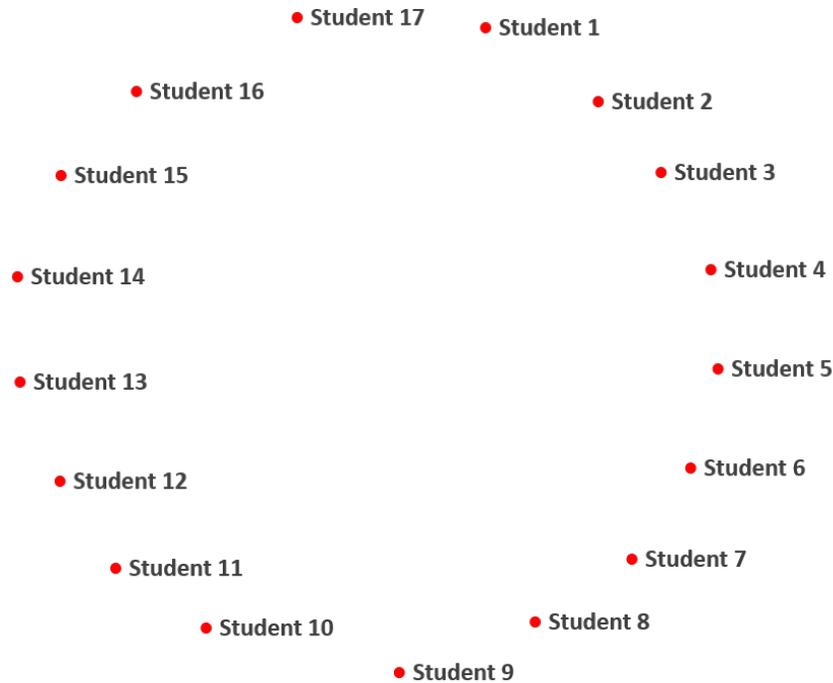


Figure 2. Representation of student participants. Full participant names have been removed and replaced with student numbers for demonstration purposes.

Representation of student participants

During the interviews, each instructor was shown a graphical representation of students in their class on their computer screen. Each student name was depicted as a circle (node) and each node was randomly placed in a circle (see Figure 2). Instructors were then asked questions regarding students’ online interactions with other students. Starting with the student at the top of the circle and moving clockwise on the circle of student-named nodes, instructors were asked to tell the researchers with whom they believed this person had online interactions. When an interaction with another student was named, the researcher drew a line from the initial student to the student named by the instructor. Each time a line was drawn by the researcher the instructor would see it depicted on their screen. This process continued until all online interactions were drawn for each student in the circle. Figure 1a shows the resulting instructor-reported networks.

Once the drawing was completed, instructors were asked if they could remember which students had the most interactions, how they would compare the online discussion to other online courses taught, whether they noticed any particular clusters of online interactions, and if they, after looking at the completed drawing, noticed anything that surprised them.

Instructors were then presented with the graphical representation of online student interactions built in Netlytic based on the Name Network approach. For consistency, the nodes were placed in the same circular format and in the same order used in Figure 2. These graphically depicted interactions showed only those interactions where the name of another student was mentioned in the interaction. Thus, the network shown to instructors represented *who mentioned whom* in the discussion forum.

Instructors compared their drawn networks to those created by Netlytic and were asked to describe the Netlytic-discovered interactions. Other questions were asked, such as if the Netlytic diagram highlighted something that had been missed in the instructor-created diagram. Finally, instructors were asked if they saw a link between a student's course performance and the level of the student's online interactions.

The final step was to compare the two types of networks (those based on the forum data and those based on the interview data/instructor's recall) to determine to what extent they were similar or different. Using SNA software called ORA, we performed a Quadratic Assignment Procedure (QAP) correlation analysis on our network data. This allowed us to conduct a pair-wise comparison between the networks in order to test for statistical significance. Because the ties in the instructor-reported networks did not have any weights or direction assigned to them, we also binarized and symmetrized the Forum networks prior to conducting a QAP analysis. This was done to ensure that ties in both networks were binary (1-present or 0-absent) and undirected (if A is connected to B, then B is connected to A). We also removed all nodes representing the instructors and teaching assistants from the Forum networks.

Results

How Accurately did Instructors Perceive Student Interactions Online?

Table 2 shows the resulting QAP correlation analysis. To ensure that this type of analysis was adequate for this study, we also built *who replies to whom networks*, based on information on who posted after whom to a particular thread. Of the three classes, only Class B had a moderate correlation between the instructor's reported network and the forum network based on the students' actual interactions. The correlation coefficient is equal to 0.26 (statistically significant, $p < 0.05$). This suggests that the Class B instructor had an accurate perception of in-class interactions overall.

Connections reported by the other two instructors did not accurately correlate with the observed forum data. This may be due to a relatively lower interaction level in forum discussions by the instructors in Classes A and C. To check this supposition, we examined the network position of each instructor in the forum network before they were removed from the network to conduct a QAP correlation analysis. This allowed us to evaluate a potential relationship between

the instructor’s position in the network (their connectivity) and how accurately they perceived the network structure.

Table 2

QAP Correlation Between the Instructor-Reported and Forum Networks

Network	Correlation	Significance
Class A (Name vs Interview Networks)	0.20	0.087
Class A (Chain vs Interview Networks)	0.14	0.151
Class B (Name vs Interview Networks)	0.26	0.027*
Class B (Chain vs Interview Networks)	0.19	0.121
Class C (Name vs Interview Networks)	0.08	0.305
Class C (Chain vs Interview Networks)	-0.08	0.238
Class A (Name vs Interview Networks)	0.20	0.087

* Statistically significant

Class A’s instructor posted 15 (~11%) of the 139 comments posted to the class forum. Most of these 15 posts were general announcements or weekly questions to the students. Notably, except for one post that mentioned both of the teaching assistants, the instructor did not mention any student’s name directly; as a result, the instructor’s out-degree centrality value was equal to 2 (two connections). And since no one mentioned the instructor directly, the in-degree centrality value was 0. At the same time, the instructor did reply to some of the student posts directly, but mostly focused on providing feedback based on the content of students’ posts without mentioning their names. As a result, Class A’s instructor was not the most connected person in the Name network (the two teaching assistants were).

Class C’s instructor posted 41 (20%) of 206 messages. In the Name network, the instructor mentioned two students directly (out-degree centrality=2) and five students mentioned the instructor by name or by a more formal “professor” (in-degree centrality=5). Most of the messages from students to the instructor were related to class logistics and not directly related to the class content. It appears that Class C’s instructor’s engagement level was relatively higher than Class A’s, but not nearly as high as Class B’s.

Finally, Class B’s instructor posted 54 (13%) of 410 messages. In the Name network, the instructor mentioned eight students (out-degree centrality=8) and two students directly mentioned the instructor (in-degree centrality=2). So Class B’s instructor posted more messages over the course of the semester (although it was fewer percentage-wise than Class C’s) and connected to more students (higher out-degree centrality) than either of the two other instructors.

How Accurately did Instructors Perceive the Class-Wide Interactivity Level?

In addition to examining why instructors over-reported or under-reported individual students’ connections, it is also useful to examine how accurately instructors perceived the

overall interactivity level in the class. We used an SNA measure called *density* as a proxy to the class-wide interactivity level. Density is a ratio of the number of existing ties to the number of all possible ties in the network. It can range from 0 to 1. Higher values of this measure represent highly interactive classes where many students talk to many others.

Table 3

Network Densities

Class	Density (Name Network)	Density (Interview Network)
Class A	0.04	0.08 (over-estimated)
Class B	0.42	0.30 (under-estimated)
Class C	0.09	0.02 (under-estimated)

Table 3 shows the density measure values for the networks in the study. The class that generated the highest number of messages (Class B) also had the highest density value (0.42), which means that about 40% of ties exist relative to the total number of possible connections (when density is equal to 1 or 100%, everyone is talking to everyone). To determine whether an instructor over- or under-estimated the overall class interactivity, we need to compare the density values of the two networks: the forum network versus interview network. Since Class B’s density of the forum network is higher than that of the interview network, we can conclude that there are fewer connections in the interview network, which means that the instructor under-estimated the class-wide interactivity level. In other words, there were more interactions among students than were perceived by the instructor. Class C’s instructor also under-estimated the interactivity (reported fewer connections than were actually observed in the Forum data). But Class A’s instructor over-estimated it (reported more connections than were actually present in the Forum data). It is important to note that a straightforward comparison between two density values does not necessarily tell us whether or not a difference is statistically significant, but it can indicate a general trend.

What Factors Affect Instructors’ Perception of Online Interactions?

During the interviews, instructors mentioned reasons why their recollection of student-to-student interactions on the class fora might have been obscured. Online discussions were only one of many communication and collaboration tools available to, and used by, students. Instructors admitted that some student interactions might have happened through other channels such as emails or chat discussions during live sessions. As a result, instructors reported that it was difficult for them to differentiate online discussions that happened on the class-wide fora from those that they observed in different media. These also include offline interactions happening during face-to-face classroom-based sessions at the end of the semester. For example, one of the students did not participate online as much as his/her instructor stated. This happened because the instructor attributed that student’s active participation during a face-to-face intensive when reporting their online connections.

Discussion

Based on the QAP correlation analysis, we hypothesize that the instructor's level of engagement in online discussions may be a good predictor of how accurately an instructor perceives student online interactions. Even though we cannot formally test this hypothesis with the current data because we have only three cases, the result offers a foundation for future hypothesis testing and theory development.

The result of the SNA density measurements suggests that the two instructors in classes with the highest volume of messages and interactivity (Classes B and C) under-estimated the overall level of interactivity in class. This could be because it is harder for instructors to keep up with class discussions in classes with more messages; As a result, an instructor might miss some interactions. As for Class A, the instructor over-estimated the level of interactivity, possibly accounting for interactions that happened outside the discussion boards such as those during classroom-based sessions or through other media. This was also the only class where online discussions were not graded.

Instructors mentioned other factors that may explain some differences in the reported networks (especially over-reporting of connections). If students performed better during the face-to-face intensive or better in class in general, then the instructor would likely perceive them to be more connected. Instructors also noted that it was easier to recall the most active online participants as opposed to those who participated less frequently. This observation is especially important as it highlights the need for LA techniques to better enable instructors to assess the participation and engagement of less frequent posters.

Based on the interview data, the most connected students in the network could be either high achievers or students who reported to their instructor that they learned the most from the class. There were also instances when academically stronger students did not actively participate in forum discussions as much and thus appeared to be isolated or weakly connected in the observed network. This suggests that the connectivity in the forum networks may hint at how much students put in to, and got out of, the class as opposed to the actual grade they received (or would receive). This insight is also important for future development of LA based on forum data as it helps to more accurately interpret network positions of students in a forum network.

Conclusion

This study contributes to the research that suggests SNA and network visualizations can be valuable LA tools to enable the detection and understanding of patterns in learning behaviours and interactions, creating an opportunity for early tailored student support (Chow, 2013). Clearly, SNA can assist instructors in more accurately perceiving their students' interactions. This is particularly important given the instructors' perceived connection between students' stronger engagement in the online discussions with students' improved learning experience, as well as the connection between frequency of interactions and enjoyment of, and satisfaction with, the course. With the data generated through LA tools, instructors and course designers can adjust learning outcomes, enhance course curricula, and make informed, intelligent, responsive, and strategic teaching decisions based on those learning patterns and projects. By making the

invisible visible, LA can better position instructors to engage students more effectively to achieve an enhanced learning experience and potentially increase student retention overall.

Study Limitations and Future Research

Only three courses were enrolled in the study, and hence three instructors were interviewed. As a result of the small sample, there were some contradictory comments made by the instructors. For example, one believed that more experienced students interacted more, while another claimed the opposite. A larger sample would allow for more generalizable results.

Interviews brought to light that more visualization details should be sought, including directionality of the contacts, frequency, and intensity. These elements could be built in to future analytics design software and their effectiveness evaluated through research. This future research could also integrate principles of networked learning design and sensemaking (Goodyear & Carvalho, 2013)

Lateness of capturing instructors' data via interviews may have affected their memory and how they perceived student interactions. Interviews with instructors should commence immediately after the end of the course but before on-site intensives, which could further affect instructors' recollections.

Finally, of the 64 potential students, only 49 were involved in the research study. The removal of students from the network structure necessarily means the visualizations was not a true representation of all interactions as we do not know the position or influence of the students who opted not to participate. To mitigate the potential impact of missing ties on the study results, we treated all ties derived from the forum data as bi-directional.

Considering the discrepancies between instructors' understanding of student networks in their courses compared to the realities presented by SNA, instructors and students would both benefit from further research into how to make these visualizations more clear, integrated, and accessible. Increasing instructor capacity to know more than they currently do about their students' collaborative behaviours is a priority. Questions about instructor social capital in student networks would also be an appropriate avenue for further research, particularly with regards to instructor ability to maintain and nurture connections in all nodes in the network, whether directly or indirectly.

Acknowledgements:

The authors would like to thank Derek Tay and Morven Fitzgerald whose collaborations made this research/paper possible.

References

- Absar, R., Gruz, A., Haythornthwaite, C., & Paulin, D. (2015). Media multiplexity in connectivist MOOCs. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 424-425). New York, NY: ACM. doi:10.1145/27235762723654

- Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education*, 58(1), 470-489. doi:10.1016/j.compedu.2011.08.030
- Arnold, K. E., & Pistilli, M. D. (2012). Course signals at Purdue: Using learning analytics to increase student success. In *Proceedings of the second international conference on learning analytics and knowledge* (pp. 267-270). New York, NY: ACM. doi:10.1145/2330601.2330666
- Aviv, R., Erlich, Z., Ravid, G., & Geva, A. (2003). Network analysis of knowledge construction in asynchronous learning networks. *Journal of Asynchronous Learning Networks*, 7(3), 1-23. Retrieved from <https://cebf360b-a-62cb3a1a-s-sites.googlegroups.com/site/reuvenaviv/Home/my-papers/JALNPaperSept2003.pdf>
- Bakharia, A., & Dawson, S. (2011). SNAPP: A bird's-eye view of temporal participant interaction. In *Proceedings of the first international conference on learning analytics and knowledge* (pp. 168-173). New York, NY: ACM. doi:10.1145/2090116.2090144
- Brooks, C., Thompson, C., & Teasley, S. (2015). A time series interaction analysis method for building predictive models of learners using log data. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 126–135). New York, NY: ACM. doi:10.1145/2723576.2723581
- Chios, S. (2018, April 3). The big data revolution: Will it help university students graduate? *The Globe and Mail*. Retrieved from <https://www.theglobeandmail.com/news/national/education/can-big-data-analysis-stop-students-from-dropping-out-of-university/article31939870/>
- Cho, H., Gay, G., Davidson, B., & Ingraffea, A. (2007). Social networks, communication styles, and learning performance in a CSDL community. *Computers & Education*, 49(2), 309-329. doi:10.1016/j.compedu.2005.07.003
- Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), 683-695. doi:10.1080/13562517.2013.827653
- Collins, R., & Gruzd, A. (2017). Learning within digital media: Investigating the relationships between student citation networks, assignment structures, and learning outcomes. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. Maui, HI. Retrieved from <https://scholarspace.manoa.hawaii.edu/bitstream/10125/41407/1/paper0258.pdf>
- Corrin, L., & de Barba, P. (2015). How do students interpret feedback delivered via dashboards? In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 430–431). New York, NY: ACM. doi:10.1145/2723576.2723662
- Couros, A. (2009). Open, connected, social – implications for educational design. *Campus-Wide Information Systems*, 26(3), 232-239. doi:10.1108/10650740910967393

- Czerkawski, B. C. (2015). When learning analytics meets e-learning. *Online Journal of Distance Learning Administration*, 18(2). Retrieved from <https://pdfs.semanticscholar.org/8eec/a2e1a71501bfe8183654a8d6121a1ff9265e.pdf>
- Dascalu, M., Stavarche, L. L., Trausan-Matu, S., Dessus, P., Bianco, M., & McNamara, D. S. (2015). ReaderBench: An integrated tool supporting both individual and collaborative learning. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 436-437). New York, NY: ACM. doi:10.1145/2723576.2723647
- Downes, S. (2010). Learning networks and connective knowledge. In *Collective intelligence and E-Learning 2.0: Implications of web-based communities and networking* (pp. 1-26). IGI Global. Retrieved from <https://philpapers.org/archive/DOWLNA>
- Dringus, L. P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Networks*, 16(3), 87-100. Retrieved from <https://files.eric.ed.gov/fulltext/EJ982677.pdf>
- Duval, E. (2011). Attention please!: Learning analytics for visualization and recommendation. In *Proceedings of the first International Conference on Learning Analytics and Knowledge* (pp. 9-17). New York, NY: ACM. doi:10.1145/2090116.2090118
- Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, 28, 68-84. doi:10.1016/j.iheduc.2015.10.002
- Gašević, D., Zouaq, A., & Janzen, R. (2013). "Choose your classmates, your GPA is at stake!" The association of cross-class social ties and academic performance. *American Behavioral Scientist*, 57(10), 1460-1479. doi:10.1177/0002764213479362
- Goodyear, P., & Carvalho, L. (2013). The analysis of complex learning environments. In H. Beetham & R. Sharpe (Eds.), *Rethinking pedagogy for a digital age: Designing for 21st century learning*, 49-63. Abingdon, UK: Routledge.
- Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, 15(42), 42-57. Retrieved from <https://dspace.ou.nl/bitstream/1820/4506/1/Translating%20Learning%20into%20Numbers%20-%20LA%20framework.pdf>
- Gruzd, A. (2009). Studying collaborative learning using name networks. *Journal of Education for Library & Information Science*, 50(4), 237-247. Retrieved from <https://www.jstor.org/stable/40732586>
- Gruzd, A. (2011). Exploring virtual communities with the internet community text analyzer (ICTA). In B. K. Daniel (Ed.), *Handbook of research on methods and techniques for studying virtual communities: Paradigms and phenomena*. Hershey, PA: IGI Global.

- Gruzd, A., Paulin, D., & Haythornthwaite, C. (2016). Analyzing social media and learning through content and social network analysis: A faceted methodological approach. *Journal of Learning Analytics*, 3(3), 46-71. doi:10.18608/jla.2016.33.4
- Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Burlington, MA: Morgan Kaufmann.
- Haythornthwaite, C. (2002). Strong, weak, and latent ties and the impact of new media. *Information Society*, 18(5), 385-401. doi:10.1080/01972240290108195
- Haythornthwaite, C. (2006). Facilitating collaboration in online learning. *Journal of Asynchronous Learning Networks*, 10(1), 7-24. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.104.2245&rep=rep1&type=pdf>
- Haythornthwaite, C. (2008). Learning relations and networks in web-based communities. *International Journal of Web-Based Communities*, 4(2), 140-158. Retrieved from <http://hdl.handle.net/2142/35047>
- Haythornthwaite, C., & Laat, M. D. (2010). Social networks and learning networks: Using social network perspectives to understand social learning. In *Proceedings of the seventh international conference on networked learning*, pp. 183-190. Retrieved from https://www.researchgate.net/profile/Maarten_Laat/publication/228960053_Social_networks_and_learning_networks_Using_social_network_perspectives_to_understand_social_learning/links/0046353536fa0e6d56000000.pdf
- Hernández-García, Á., González-González, I., Jiménez-Zarco, A. I., & Chaparro-Peláez, J. (2015). Applying social learning analytics to message boards in online distance learning: A case study. *Computers in Human Behavior*, 47, 68-80. doi:10.1016/j.chb.2014.10.038
- IBM & Campus Technology (2012). Building a smarter campus: How analytics is changing the academic landscape. Retrieved from ftp://ftp.software.ibm.com/la/documents/gb/mx/Building_a_Smarter_Campus.pdf
- Knight, S., & Littleton, K. (2016). Dialogue as data in learning analytics for productive educational dialogue. *Journal of Learning Analytics*, 2(3), 111-143. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1127064.pdf>
- Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist*, 57(10), 1439-1459. doi:10.1177/0002764213479367
- Marreiros, G., Santos, R., & Ramos, C. (2010). GLSS – Group Learning in Shared Spaces considering aspects like emotion and personality. In J. C. Augusto, J. M. Corchado, P. Novais, & C. Analide (Eds.), *Ambient intelligence and future trends-international symposium on ambient intelligence (ISAmI 2010)*, vol. 72 (pp. 37-44). New York NY: Springer.

- Pardo, A., Ellis, R. A., & Calvo, R. A. (2015). Combining observational and experiential data to inform the redesign of learning activities. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 305-309). New York, NY: ACM. doi:10.1145/2723576.2723625
- Paulin, D., Gilbert, S., Haythornthwaite, C., Gruzd, A., & Absar, R. (2015). Beyond the backchannel: Leveraging Twitter to enact learning processes. In *Proceedings of the 2015 iConference*. Urbana, IL. Retrieved from <http://hdl.handle.net/2142/73779>
- Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks*, 16(3), 9-20. Retrieved from <https://files.eric.ed.gov/fulltext/EJ982669.pdf>
- 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. doi:10.1016/j.compedu.2013.06.009
- Serrat O. (2017) Building a Learning Organization. In: *Knowledge Solutions*. Springer, Singapore. Retrieved from <https://digitalcommons.ilr.cornell.edu/intl/119/>
- Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1), 3-10. doi:10.3109/0142159X.2016.1173661
- Thormann, J., Gable, S., Fidalgo, P. S., & Blakeslee, G. (2013). Interaction, critical thinking, and social network analysis (SNA) in online courses. *International Review of Research in Open and Distributed Learning*, 14(3). Retrieved from <http://www.irrodl.org/index.php/irrodl/article/viewFile/1306/2606>
- Tinto, V. (2000). Taking student retention seriously: rethinking the first year of college. *NACADA Journal*, 19(2), 5-10. doi:10.12930/0271-9517-19.2.5
- Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. *EDUCAUSE Learning Initiative*, 1(1), 1-11. Retrieved from https://www.researchgate.net/profile/Angela_Van_Barneveld/publication/265582972_Analytics_in_Higher_Education_Establishing_a_Common_Language/links/575f12e108ae9a9c955fade7/Analytics-in-Higher-Education-Establishing-a-Common-Language.pdf
- Vassileva, J., & Sun, L. (2007). Using community visualization to stimulate participation in online communities. *E-Service Journal*, 6(1), 3-39. doi:10.19173/irrodl.v14i3.1306
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500-1509. doi:10.1177/0002764213479363

- Wise, A. F., Zhao, Y., & Hausknecht, S. N. (2013, April). Learning analytics for online discussions: A pedagogical model for intervention with embedded and extracted analytics. In *Proceedings of the third international conference on learning analytics and knowledge*, 48-56. New York, NY: ACM. doi:10.1177/0002764213479363
- Xing, W., Wadholm, R., Petakovic, E., & Goggins, S. (2015). Group learning assessment: Developing a theory-informed analytics. *Journal of Educational Technology & Society*, 18(2), 110–128. Retrieved from <https://pdfs.semanticscholar.org/583a/ef543418b0b52c507d028c9e698cb3a6b785.pdf>
- Zhang, Y., Oussena, S., Clark, T., & Hyensook, K. (2010). Using data mining to improve student retention in HE: a case study. In *Proceedings of 12th International Conference on Enterprise Information Systems*, Portugal. Retrieved from <http://eprints.mdx.ac.uk/5808/>

Authors

Scott Comber is a University Teaching Fellow at the Rowe School of Business, Faculty of Management, Dalhousie University. Scott's primary area of research is the scholarship of teaching and learning. He studies innovative teaching practices that improve the way we engage with, and educate, university students. E-mail: scott.comber@dal.ca

Martine Durier-Copp is Director of the Centre for Graduate and Executive Education within the Faculty of Management, which manages all online/blended and executive education masters programs, and is also Adjunct Professor in the School of Public Administration. She is currently Dalhousie University Managing Director of the CELNeT Lab (Collaborative e-learning Network Lab). Martine is a Social Sciences and Humanities Research Council supported researcher on e-learning, e-leadership, and managing virtual teams. E-mail: martine.durier-copp@dal.ca

Anatoliy Gruzd is a Canada Research Chair in Social Media Data Stewardship and an associate professor and research director of the Social Media Lab at Ryerson University's Ted Rogers School of Management, Toronto, Canada. Anatoliy studies how social media use is changing the ways in which people and organizations communicate, connect, and how these changes impact our society. E-mail: gruzd@ryerson.ca



This work is licensed under a Creative Commons Attribution-NonCommercial CC-BY-NC 4.0 International license.