JOURNAL OF LEARNING ANALYTICS

Volume 7(1), 59–71. http://dx.doi.org/10.18608/jla.2020.71.5

Investigating Collaboration as a Process with Theory-Driven Learning Analytics

Carmel Kent[1], Mutlu Cukurova[2]

Abstract

Although collaboration is considered a key 21st-century skill, oftentimes it is only assessed through the outcome measures of individual learners. In this paper, we draw upon collaborative cognitive load theory (CCLT) to explain the process of collaboration in learning from the point of view of the collective, rather than an individual learner. Using CCLT we suggest a new method of measuring the process of collaboration regardless of its outcome measures. Our approach — Collaborative Learning as a Process (CLaP) — uses social network analysis to evaluate the balance between interactivity gains and coordination costs of learner communities. Here, we demonstrate the approach using real-world data derived from the digital tracks of two online discussion communities. We argue that our conceptual approach can enable instructors and learners to unlock the black box of collaboration in learning. It has the potential to support the development of learner skills that go beyond cognition. We conclude the paper with the results of our investigation of the value of the approach to the online module instructor.

Notes for Practice

- Research on the assessment of collaborative learning is still ongoing, considering the advancement in online technologies and new theoretical considerations such as collaborative cognitive load theory (CCLT).
- Collaboration in learning is currently mainly assessed through summative evaluation of individuals.
- We draw upon CCLT to suggest a quantitative instrument for the assessment of the collaboration process in learning.
- We evaluate the collective rather than the individual, and the process rather than the outcome.
- The CLaP approach introduced here gives teachers a glimpse into the process of collaboration rather than its summative outcomes by viewing the changes occurring in the relationship between interactivity gains and coordination costs.

Keywords

Collaborative Learning as a Process (CLaP), assessment of collaboration in learning, cognitive load theory (CLT), collaborative cognitive load theory (CCLT), collective learning, social network analysis

Submitted: 05.04.2019 — **Accepted:** 17.12.2019 — **Published:** 03.04.2020

Corresponding author [1]Email: c.kent@ucl.ac.uk Address: Knowledge Lab, Institute of Education, University College London, 3–29 Emerald St., Holborn, London WC1N 3QS

[2] Email: m.cukurova@ucl.ac.uk Address: Knowledge Lab, Institute of Education, University College London, 3–29 Emerald St., Holborn, London WC1N 3QS

1. Background and Introduction

Every psychological process can be considered as cognitive in the sense of relying on information processing; however, "non-cognitive" has become the common term for "traits or skills not captured by assessments of cognitive ability and knowledge" (West et al., 2016, p. 1; Farrington et al., 2012), especially in the context of 21st century skills (Dede, 2010). The interest in measuring non-cognitive skills is increasing, as employers are challenged with graduates proficient at passing exams, but lacking the expected expertise in non-cognitive skills such as the ability to collaborate effectively (Dede, 2010; Casner-Lotto & Barrington, 2006; Murnane & Levy, 1996). The assessment of collaboration skills requires looking beyond particular learning outcomes such as grades, and beyond individual-level skills of learners (Farrington et al., 2012). In this paper, the learning goal is focused on how to collaborate, which requires an emphasis on the process of collaboration rather than domain-specific knowledge. We draw upon collaborative cognitive load theory (CCLT; Kirschner, Sweller, Kirschner, & Zambrano,

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

2018) to analyze and explain the process of collaboration in learning as a function of the collective working memory gains and coordination costs inherent in working together.

The assessment of non-cognitive skills is essential to provide teachers and learners with insights into their development. Summative, outcome-based methods of measuring non-cognitive skills (such as multi-choice end-of-term exams) are generally mature and technically simple to assess. However, their accuracy and external validity have been criticized (Knight, Buckingham Shum, & Littleton, 2013). For instance, can we really measure collaboration skills using individual tests? The challenge in assessing non-cognitive skills such as collaboration remains relevant today, particularly in terms of process-oriented and formative assessments (Gardner, 2011).

Kent, Laslo, and Rafaeli (2016) have shown a correlation between summative learning outcomes and processual interactivity patterns of behaviours in online collaborative learning discussions. However, it is not a strong correlation and there is no causal direction. This suggests that individual outcome evaluations alone might not be enough to judge the quality of collaboration in online learning settings. Instead, we suggest that the dynamic measurement of interaction networks should be considered. As the area of learning analytics has shown, analytical technologies can be instrumental in assisting teachers with the assessment itself (Strijbos, 2011). However, deductive approaches driven by learning theories in the learning analytics field, appear to be less frequent, and are indeed needed (Cukurova, 2018). Therefore, in this paper we offer a theory-driven learning analytics approach to the process evaluation of learner collaboration.

The paper is structured as follows: the introduction and related work sections present the literature on our chosen theoretical framework that drives the analytical framework. The methods section explains our methodology to assess collaboration in learning as a process. In the results section, we present two case studies to illustrate the potential of our approach using real-world data. Lastly, we discuss our conclusions.

2. Related Work

2.1. Collaboration and Learning Outcomes

Collaboration is about interacting with others to apply knowledge and to reach a shared understanding. Looking at the existing evidence, collaboration can reinforce knowledge and improve learner attainment, as well as improving their attitudes towards certain subject areas (Johnson & Johnson, 2002; Johnson, Johnson, Roseth, & Seob Shin, 2014). However, collaboration's positive outcomes for learning are not always consistent (Kester & Paas, 2005; Slavin, 2014). Most researchers would agree that learners should be appropriately guided while collaborating to achieve the expected learning outcomes (Kirschner, Sweller, & Clark, 2006; Cukurova, Bennett, & Abrahams, 2018). However, the peculiarities of the collaboration process and what makes it work or fail are still being questioned by our research community (Dillenbourg, 1999; Collazos, Padilla, Pozzi, Guerrero, & Gutierrez, 2014).

2.2. Theoretical Framework to Conceptualize the Process of Collaborative Learning

Kirschner et al. (2018) argued that existing theories on the human cognitive architecture underlying learning processes, particularly cognitive load theory (CLT), can provide valuable insights into collaborative learning. CLT proposes a transactional model to describe our cognitive resources used in individual learning (Sweller, 1988). Specifically, its emphasis is on the working memory (WM) resources required to carry out a learning task, which is used to assess the effort invested in individual learning. CLT is based on the notion that the WM, thought of as our main cognitive consumer of novel information, is known to be limited in capacity (Cowan, 2001) and duration (Peterson & Peterson, 1959). Long-term memory (LTM), on the other hand, thought of the cognitive storage of our existing information, does not impose any limits to learning in terms of capacity (Janssen, Kirschner, Erkens, Kirschner, & Paas, 2010).

According to CLT's environmental organizing and linking principle, external signals trigger the retrieval of information from the LTM to the WM (Janssen et al., 2010). The effective incorporation of new information elements into schemas in the LTM will make retrieval from the LTM back into the WM more effective. This interaction between the newly acquired information and already assimilated knowledge is at the heart of cognitive and constructivist theories of learning (Ausubel, 1968). In his assimilation theory, Ausubel (1968) made a distinction between meaningful and rote learning. To be meaningful, newly acquired information must be assimilated into existing cognitive structures. Thus, we use the concept of a "meaningful interaction" as the act of relating newly introduced information with already structured knowledge to serve as the core of our working definition for learning. This definition posits the importance of 1) learners' opportunities to be exposed to new concepts, information, and observations; 2) solid schemas of existing knowledge in the LTM; and 3) learning designs that reflect on the connections between new information and existing LTM schemas.

2.3. From Solo to Socio: Placing Group "Cognition" Within its Network of Interactions

In constructivist conceptualizations, learning is placed among learners before it takes place within the individuals. For example,

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Piaget's sociocognitive conflict theory saw the potential in learning within contradictions created between learners' existing understanding and their new experiences resulting from social interactions. Vygotsky's (1978) sociocultural view places individual learning as originating within social interactions, and human actions as mediated by knowledge co-construction semiotic tools (Palincsar, 1998). Similarly, other branches of cognitive sciences, such as Distributed Cognition (Hutchins, 1996) places cognition as distributed across objects, individuals, artefacts, and the environment. That is, when a group is performing a cognitive task, it manifests cognitive properties that differ from the cognitive properties of individual learning (Hutchins, 1996). From a CLT perspective, collaborative learning avails the group's cognitive resources to the disposal of the individual learner (Janssen et al., 2010). In these conceptualizations of collaborative learning, similar to the use of distributed computing or cloud computing, the effect of collaboration on the individual's learning is mainly achieved by affording more resources at the individual's disposal. In this sense, collaborative learning can be considered as a network of interactions between learners and artefacts that scaffold and set the context for individual learners by allowing them to be exposed to other people's relation making, and to other people's LTM. Based on these theoretical considerations, we define collaboration in learning as the process of establishing and growing a network of learners and knowledge artefacts (e.g., Siemens, 2005) that will serve as a shared cognitive resource (Kirschner et al., 2018) used by individual learners. This scaling — from interactions between newly introduced and already established knowledge at the individual level to the set of interactions between learners and concepts at the collective level — is instrumental to our conceptualization of group "cognition."

2.4. Analyzing the Process of Collaboration

The above conceptualization of collaboration in learning as a process is well-resonated in the literature. Roschelle and Teasley (1995) define collaboration as "a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem" (p. 70). However, the measurement of collaboration often focuses on individual gains in the learning process rather than the actual collaboration. For instance, Kirschner, Paas, and Kirschner (2011) and Kirschner, Kirschner, and Janssen (2014) discuss the concept of collective memory effect, suggesting that collaborative learning may be more effective than individual learning if the complexity of the learning task requires cognitive resources exceeding the limits of each individual learner's WM (Janssen et al., 2010). However, the collective memory effect is defined and evaluated from the perspective of the individual's learning gains, and it assumes the role of the group's cognitive resources as facilitating or enabling tools for individual learning, rather than as an entity by itself. Thus, in that research stream, the effectiveness of collaborative learning is measured through the individual lens (Johnson & Johnson, 2002), and the cognitive load is measured using the individual unit of analysis (Kirschner et al., 2018).

On the other hand, Ickes and Gonzalez (1996) argue that the study of the subjective phenomena occurring within a single cognitive system cannot be carried out using the same sterile approach when studying the intersubjective phenomena occurring between several cognitive systems. "Research on subjective social cognition derives from the paradoxical assumption that the best way to study social cognition is to first remove it from the social interaction context in which it naturally occurs" (Ickes, 2002, p. 7). Within the context of collaborative learning, though, it is not ideal to artificially consider each cognitive system as a separate entity without also evaluating the interdependent network effect.

Bridging such a gap requires us to question whether the cognitive ecosystem of the learning group can be modelled as a linear additive function of its individual cognitive ecosystems. Woolley, Chabris, Pentland, Hashmi, and Malone (2010) show that collective intelligence of a group, defined as the general performance of the group in a wide variety of tasks, is not correlated with an aggregative function of the intelligence of its participants, but rather on interactional traits of the group, such as social sensitivity, the equality in distribution of conversational turn-taking and the proportion of females. The cognitive systems of the group participants are intertwined in a network of interactions between LTMs and WMs of the agents (human and artefacts) involved in the collaboration. Similarly, Stahl (2010) claimed that the collaborative cognitive act could not be reduced to and discussed in terms of individual cognition. To Stahl, cognition does not necessarily need to be attributed to an individual, not even to a human: "Anything that can compute well enough to play chess or prove theorems can be a cognitive agent — whether they are a person, computer or collaborative small group" (p. 26). Therefore, as opposed to a linear addition of the individual cognitive systems, interactivity itself should be considered as a new resource at the disposal of each learner. This leads us to choose a network analysis approach to operationalize the collective entity of learner collaboration in learning.

3. Methodology: Calculating The Clap Evaluation Framework

Researchers have used a variety of approaches to analyze collaborative learning, such as content analysis methods (Veerman & Veldhuis-Diermanse, 2001), mixed-evaluation methodologies (Martinez, Dimitriadis, Rubia, Gomez, & De La Fuente, 2003), and social network analysis (SNA; De Laat, Lally, Lipponen, & Simons, 2007). Here, we use SNA as one potential way of realizing collaboration in learning as a process. SNA has been used extensively to analyze collaborative learning, especially to assess individual learner roles (Rabbany, Elatia, Takaffoli, & Zaïane, 2013), but also for assessing collaborative

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

patterns (Chounta, Hecking, Hoppe, & Avouris, 2014) and with specific focus on temporality (Claros, Cobos, & Collazos, 2016). Most of these investigations are inductive in nature, as the conclusions are drawn from applying common SNA measures, rather than deduced from educational theories. Although the use of SNA to interpret collaborative learning is not new to the field, to the best of our knowledge, this paper is the first attempt to use SNA to operationalize CCLT.

Based on CCLT, we suggest that analyses of collaboration in learning should consider at least the following two main values: 1) the overhead of coordination costs and 2) the added value of interactivity between the network participants. The coordination costs are the WM cognitive costs of general communication and coordination processes. They might be low due to being biologically primary, such as gestures, or high due to being biologically secondary, such as task-specific communication (Paas & Sweller, 2012). On the other hand, the interactivity gains are the cognitive contributions or aids resulting from interactions among co-learners.

Assessing the coordination costs (increasing the cognitive load) and interactivity gains (decreasing the cognitive load) using network analysis could provide an authentic estimation of collaboration processes in learning. We call this relationship between interactivity gains and coordination costs the Collaborative Learning as a Process (CLaP) index. In terms of a networked view, the group structural features that prevent an efficient, quick, effective flow of information would lead to high coordination costs. On the other hand, structural features that support a free, egalitarian, efficient flow of information, that expose learners to content and connect new information to existing knowledge, would lead to high interactivity gains. Given a specific collective learning process for a specific community within a finite duration, the higher the relation between the interactivity gains and coordination costs (i.e., the CLaP index), the higher the likelihood of effective collaboration among the participants.

3.1. Operationalizing the Concept of CLaP Analysis

In this section, we will suggest various operationalizations, or proxies, to calculate both the coordination costs and the interactivity gains using social network analysis. It is not our intention to provide an exhaustive list of proxies, but merely to suggest a potential operationalization for our theoretical arguments. Thus, we begin by suggesting a few guidelines for choosing a specific operationalization. First, the measure should analyze the process of collaboration rather than just its outcomes. Second, the unit of analysis should reflect on the whole group, aiming to reflect on the interdependencies between participants. Third, in order to measure the quality of the collaborative process, we aim to evaluate the relation between interactivity gains and coordination costs (CLaP index) rather than their independent absolute values. Our purpose is to provide an assessment that would operate regardless of the instructional design, and would not need to be tailored to the subject domain, both of which (among other factors) may affect the absolute values of coordination costs and interactivity gains. In the next section, we demonstrate these guidelines by using SNA.

3.2. Using SNA to Measure Interactivity Gains and Coordination Costs

Social network theory provides tools to understand how autonomous individuals combine to create sustainable societies (Borgatti, Mehra, Brass, & Labianca, 2009). One of its key axioms is that structure matters. For example, groups composed of the same set of member skills can perform differently depending on the patterns of interaction among them. For example, group cohesion (i.e., level of connectedness) may be accountable for the superiority of group performance over individual (e.g., Chiocchio & Essiembre, 2009).

A detailed explanation of SNA approaches is beyond the scope of this paper. However, we provide definitions of the SNA proxies referred to here in Appendix A (for details, see Wasserman & Faust, 1994; Easley & Kleinberg, 2010). Here, we suggest some structural properties that can be used to measure the dynamics of interactivity gains and coordination costs. These proxies are merely initial suggestions for operationalization. An extended, fine-tuned suggestion for proxies is also beyond the scope of this paper.

Before turning to proxies, it is important to mention a well-researched, non-structural property of learning communities: their size. As the number of interactions that learners maintain with their peers increases, their potential ability to learn increases, but at the same time there is a cognitive limitation to the number of ties humans can maintain (e.g., Dunbar, 1992). Many group dynamics studies have shown how smaller groups can more easily enable trust, commitment, and cohesion (Soboroff, 2012). However, since collaborative learning is such a complex construct, and since group size also relates to other structural properties of networks such as density and mean degree, we must investigate further which size and tie formation process result in the coordination costs overcoming the interactivity gains, leading to a lower CLaP.

3.2.1. Measures of Interactivity Gains

Interactivity gains are the cognitive wealth and benefit resulting from interactions with co-learners. Interactions, using social influence, are known to boost collaborative learning and collective performance (Bernstein, Shore, & Lazer, 2018). We suggest the three proxies below to measure interactivity gains.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Group cohesion is related to group members' motivation to advance the group's continuity and to group collaboration in general (Kwon Liu & Johnson, 2014; Tulin, Pollet, & Lehmann-Willenbrock, 2018). One of the network measures commonly used for cohesion is **density**. The denser the network, the more cohesive the group. Some rationales behind it are the shorter paths of information flow and the social "stickiness" that discourages group members from leaving (Tulin et al., 2018).

The clustering coefficient of a node essentially measures how well its direct neighbours are connected between themselves. Thus, a network with a higher **average of clustering coefficient** is likely to manifest more collaborative behaviours (Newman, 2001).

The **reciprocity** of interactions in a network indicates its level of mutuality in interactions. Mutuality and the members' level of participation in a two-way dialogue are indicators of collaborative learning (Damon & Phelps, 1989; Cukurova, Luckin, Millán, Mavrikis, & Spikol, 2017; Cukurova, Luckin, Millán, & Mavrikis, 2018). Thus, reciprocity is the third proxy we suggest for interactivity gains.

3.2.2. Measures of Coordination Costs

The coordination costs in our context are the cognitive resources required by group members to manage the interdependencies between themselves, their ideas, and interpretations, in order to complete a task collaboratively. The dynamics of interactions network evolvement is very crucial to the level of coordination costs. For the same resulting network structure, a different process of a network's growth may lead to a different amount of associated coordination costs (Goyal & Vega-Redondo, 2005). We suggest the six proxies below to measure the coordination costs: 1) variance in degree, 2) average closeness, 3) diameter, 4) average path length, 5) variance in betweenness, and 6) number of strongly connected components.

Variance in degree: The concept of network controllability is strongly related to coordination costs. According to control theory (Kalman, 1963), a dynamic network is controllable if there exists a specific choice of inputs for which it can evolve from an initial state to any desired final state within a finite time (Liu, Slotine, & Barabási, 2011). Thus, the more controllable a network, the fewer resources needed for coordination. Liu et al. (2011) have shown that sparse and heterogeneous networks (in terms of node degrees) are harder to control than dense and homogeneous ones. Thus, we suggest variance in the nodes' degree as a proxy for coordination costs.

Average closeness: Coordination has been operationalized by centralization, which was found to be negatively correlated with knowledge sharing (Tsai, 2002). Closeness centrality (Freeman, 1978) measures how close, and thus central, a node is in relation to all other nodes. The higher the closeness level of a node, the more independent its associated learner since they do not need to go through many others to access information. Thus, we use the average closeness in a network as a proxy for coordination. The lower the average closeness, the higher the coordination costs.

Diameter and average path length: Using the same argument, network distance measures such as diameter and average path length are also related to the inverse of closeness. Thus, coordination costs could also be explained by distance measures. The more distant nodes are from each other, the higher the coordination costs required to collaborate.

Variance in betweenness: A high betweenness value of a node means its associated learner is more likely to become a bottleneck in the network since the flow of information depends on them (Tulin et al., 2018). The higher the variance in betweenness, the more extreme values of betweenness exist, and thus more nodes with low betweenness might be dependent on more nodes with high betweenness. Thus, we suggest the variance in betweenness as a proxy for coordination costs. The higher the variance, the more costly coordination will be.

Number of strongly connected components (SCCs): The circulation of information and ideas is much faster within a strongly connected component than in a less connected one (Newman, 2001). Thus, assuming the coordination costs between SCCs is higher than within an SCC, lead us to suggest that the number of SCCs can also serve as a proxy for coordination costs. The larger the number of SCCs, the higher the coordination cost.

The SNA measures for both interactivity gains and coordination costs are summarized in Table 1. In section 4, we will demonstrate our approach using data extracted from online interactions. Although it is clear that online and face-to-face collaborative learning hold many differences, as both are based on a network of interactions, we hypothesize that this approach could be demonstrated for face-to-face collaborative learning as well.

4. TWO CASE STUDIES OF CLaP ANALYSIS

To demonstrate the operationalization of our approach, we use two case studies of two graduate classes in two different universities, both under a business faculty taught by the same instructor, using blended learning and the same discussion platform, as described in the next subsection.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 1. Summary of Suggested Network Measures and Interpretations

Network Measure	Operationalized Construct	Dimension	Interpretation
Density	Cohesion	Interactivity gains	High density indicates high interactivity gain.
Average clustering coefficient	Direct neighbours' connectedness	Interactivity gains	High average clustering coefficient indicates high interactivity gain.
Reciprocity	Mutuality in interactions	Interactivity gains	High reciprocity indicates high interactivity gain.
Variance in nodes degree	Controllability	Coordination costs	High variance in the node degree indicates high coordination costs.
Average closeness	Independence	Coordination costs	Low average closeness indicates high coordination costs.
Diameter	Longest distance in the network	Coordination costs	High diameter indicates high coordination costs.
Average path length	Average distance in the network	Coordination costs	High average path length indicates high coordination costs.
Variance in betweenness	Inequality	Coordination costs	High variance in betweenness indicates high coordination costs.
Number of strongly connected components	Speed of communication	Coordination costs	A high number of strongly connected components indicates high coordination costs.

4.1. Online Discussion Tool: Ligilo

The online discussion tool we use is called Ligilo[1], a hyperlinked discussion platform where each post is expressed as a node in a semantic network of posts. Using Ligilo, collaborative learning communities create collective concept maps while conducting online discussions. The relations among posts are semantically tagged by community members (e.g., "reminds me of" or "as opposed to"; see screenshots in Appendix B).

Ligilo was chosen since it is based on our working definition of learning as occurring meaningfully when new information is related to existing prior knowledge (Ausubel, 1968). Essentially, three types of learning interactions are implemented in Ligilo: 1) active contribution (e.g., creating, updating, or deleting posts); 2) content consumption (e.g., viewing posts, refreshing sub-posts, viewing a map of posts, viewing posts' links, viewing attachments, and searching); and 3) content organization (e.g., connecting non-connected posts, voting and un-voting for a post, following and unfollowing posts, and following and unfollowing learners). In our analysis, we adopted this separation of measuring the different three types of interactions based on the notion that the three sub-networks behave differently (which might result from a different learning design; Kent & Rechavi, 2018). Interactions of and with the platform, such as logins and profile updates, were excluded from the analysis.

4.2. CLaP Analysis of Two Communities

4.2.1. The Sample and Study Context of the Case Studies

We investigated two communities as part of our case study, pseudo-named Community 1 and Community 2. The descriptive details about these two communities' discussions are summarized in Table 2. As can be seen, although both communities were taught using blended learning, both within management faculties, using the same discussion tool by the same instructor, and with a similar moderation style, they present quite different descriptive features of collaboration such as the total number of interactions, posts, and cross-references.

The two communities were given the similar collective task of building a collective concept map, covering the module's concepts, including the mapping of articles and relevant examples to the various concepts. Community 1 was given a second collective task, starting the 5th week, of collectively creating a set of questions and answers for their final exam, also related to the concepts contributed at the first task. From a CLT perspective, Ligilo's interactions can be thought of as "footprinting" the collective schema as it is being generated. Thus, the effective incorporation of new information elements into the created map is designed to help the retrieval from the "collective LTM" back into the students individual WMs.

[1] www.go-ligilo.com

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 2: High-Level Description of the Two Learning Communities

Class name	Term	Total number of interactions	Total number of posts	Number of students	Total number of cross-references*
Community 1	Autumn 2015	7,600	253 (246 after deletions)	42	24
Community 2	Spring 2017	15,515	408 (399 after deletions)	32	222

* Cross-references are the actions of relating two existing (initially unrelated) posts.

In order to gather more contextual information about the two communities, we interviewed the course instructor. Community 1 was taught within a special postgraduate business program in which students arrive from out of the country for a one-year degree. Most students were thus not local and were very young (in their early twenties) relative to other MBA students, who often have substantial work experience. According to the course instructor, because of the students' young age, and perhaps due to their cultural heterogeneity, the interaction in class was not as the instructor would have expected. The instructor also argued that most of the students' motivation was not rooted in a genuine interest in the taught subjects. Community 2, on the other hand, was composed of local postgraduate students studying in a more common two-year program. Most of these students were experienced knowledge workers in their thirties and forties who chose to study the course out of professional interest. Thus, the instructor argues that due to their maturity, and perhaps cultural homogeneity, the interaction between the instructor, the students, and the learning materials was more efficient.

4.2.2. CLaP Analysis

Ligilo's bimodal network of posts and learners is a nearly bipartite graph. As a pre-analysis phase, we transformed this graph into a single modal graph in the following way: for each pair of learners *(j, i)*, if post x created by *i* with which *j* interacts, the network will possess a tie (edge) $j \rightarrow i$. The number of edges reflects on the number of absolute interactions. For example, if learner *i* views the same post (contributed by *j*) twice, the network will include two edges of type "view" between *i* and *j*.

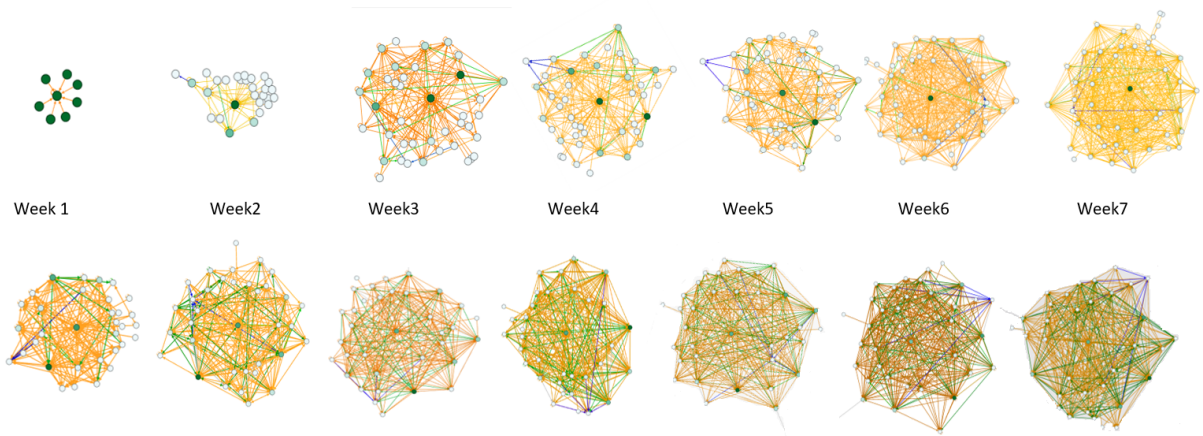

Figure 1. Evolving interaction networks of Community 1 (top) and Community 2 (bottom). Consumption type interactions are orange, contributions are blue, and organizational interactions are green.

Aligned with the instructor's comments, it becomes clear from Figure 1 that the two communities show different trends of interaction. First of all, Community 1 starts slowly with very few interactions and builds slowly. Whereas, Community 2 starts with a high number of interactions and continues to grow for the next seven weeks. It is also clear that Community 1 is composed mainly of consumption-type interactions and very few contributional and organizational interactions. On the other hand, Community 2 shows plenty of organizational interactions, particularly after the first two weeks, and more contribution interactions compared to consumption interactions.

Looking at basic metrics of network growth, as shown in Figure 2, the two communities' interaction networks show a similar rate of growth. Albeit Community 1 presenting much lower values and a much less variant network of interaction, the rate of growth is similar. However, using the CLaP index computed by interactivity gains and coordination costs provides us with a clearer picture of the collaboration process for both communities.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

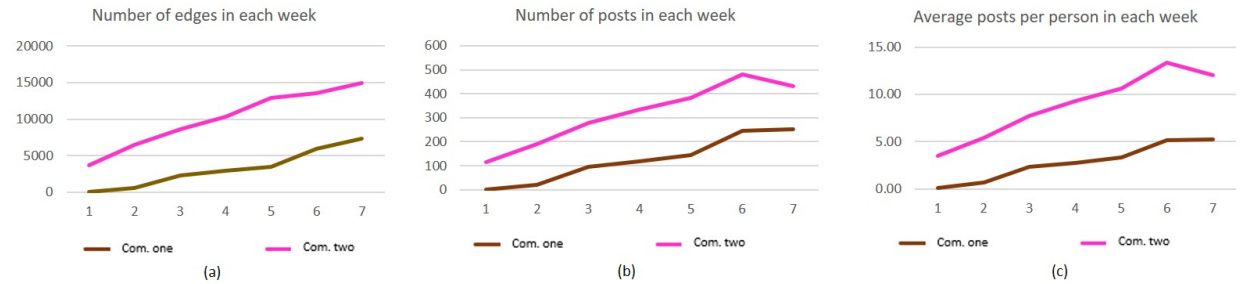

Figure 2. Interactions of Community 1 (brown) and Community 2 (pink): (a) the number of edges, (b) the number of posts per week, and (c) the average number of posts per person per week.

In Appendix C, we show the interactivity gains and coordination costs of the two communities as operationalized by all the metrics suggested in section 3.2.1. To demonstrate the use and validation of CLaP analysis, in the next section we describe a single CLaP operationalization in which interactivity gains are proxied as reciprocity and transaction costs are proxied as the variance in degree. These were chosen in accordance with the course instructor's two pedagogical goals for collaboration 1) learners should engage in mutual interactions and 2) learners should reach a shared understanding. More specifically, reciprocity was chosen to reflect how mutual the learner actions are; the variance in their degree of interactions was chosen as an aggregation of how similarly individual learners "see the world," and to what extent they reach a shared understanding. The more unbalanced they are, the more effort they would need to put into coordinating their understanding to reach a consensus.

4.2.3. Initial Validation of Clap as the Relationship Between Reciprocity and Degree Variance

One of the challenges associated with evaluating the process of collaboration in learning is that there is no commonly accepted standard validation measure (OECD, 2015). Individual summative assessments are not the right assessment for collaboration on the collective level (Kent, Rechavi, & Rafaeli, 2018). To address this challenge, we chose a mixed-methods approach in which we apply a quantitative analysis and then qualitatively validate the value of the CLaP approach with the module instructor (Chounta et al., 2014).

Woolley et al. (2010) address collective intelligence as the group's collective performance in a variety of tasks. In the context of our two communities, the collective tasks were carried by means of online interactions in a virtual shared space. From a CLT perspective, the individual efforts (measured by learner interactions) are done to avail the group's cognitive resources back to the disposal of the individuals (Janssen et al., 2010). Our hypothesis is that the CLaP index would significantly correlate with the number of the three types of interactions measured in Ligilo. Figures 3 and 4 show a strong positive correlation between the progress of the CLaP index (in purple), in both left images for Community 1 (Figure 3) and Community 2 (Figure 4), and the collective effort as measured by the number and the types of learner interactions on the right images respectively.

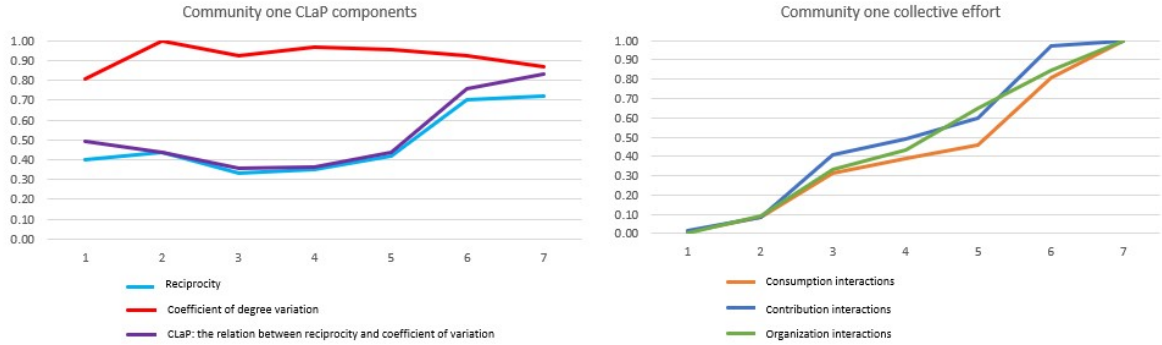

Figure 3. (Left) Community 1's CLaP components, normalized to a 0–1 scale; (right) Community 1's number of interactions, normalized to the size of the community and to a 0–1 scale.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

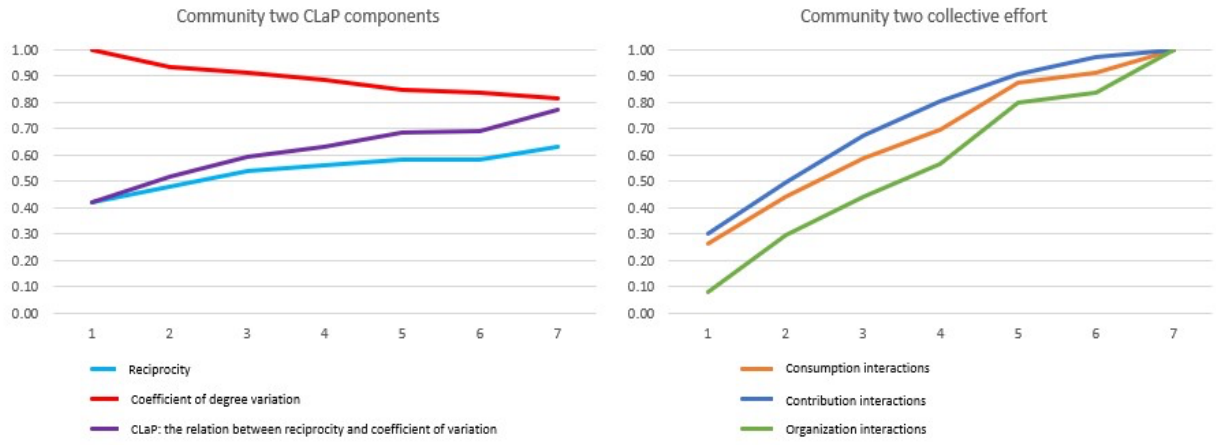

Figure 4. (Left) Community 2's CLaP components, normalized to a 0–1 scale; (right) Community 2's number of interactions, normalized to the size of the community and to a 0–1 scale.

Despite the obvious limitations of this kind of quantitative analysis (the small dataset of only seven data points for each community) significant correlation results are indicative of the CLaP index's validity to measure the collaboration process. In general, it is interesting to note that all three participation variables behave similarly, and that they all strongly correlate with the CLaP index (Pearson correlation between the three participation variables and the CLaP index was above 0.9). We also confirmed that CLaP is more highly correlated to the participation metrics than its two independent components (reciprocity and efficient of variation) individually did. This also indicates the higher meaning of their relationship as opposed to their absolute values for measuring collaboration.

To obtain external validity and rich insights into the potential of the index, we organized a feedback session with the course instructor. When presented with the charts above to reflect upon, the instructor argued that the CLaP index can be valuable in capturing the general collaboration quality of the class, which is especially problematic in large classrooms and distance learning settings such as MOOCs. When asked to interpret the graphs, the instructor talked us through some examples where CLaP raises issues that cannot be seen using simple participation metrics (such as those on the right side of Figures 3 and 4). For example, the stable correlation between the CLaP index and the participation charts begins from the first week of Community 2 and stays steady, but only begins to be stable by week 3 in Community 1. As a reflection on this irregularity over the first three weeks, the instructor pointed to tensions arising in Community 1. He explained how the learners had relatively low motivation to use online discussions and lacked appreciation for collaborative work in general. The instructor asserted that "*this made their collaboration process somewhat 'bumpy*.'" This might also explain why, in general, the collaboration process was limited in Community 1. Among Community 2 members, as argued by the instructor, "*the need to collaborate online was well understood and appreciated*." Reflecting upon the performance of Community 1, the instructor noted that "*students were not happy with the introduction of collaborative tasks and they were asking why they need to complete them. Previous cohorts of the same class didn't have the same tasks and students of the investigated cohort were uncomfortable for the first few weeks with the allocated collaborative tasks. This phenomenon is not captured in regular participation analysis, such as number of contributions, but can be easily seen with CLaP charts*."

Another example of CLaP's ability to capture the collaboration process, according to the instructor, was the indicated rise in the CLaP index in Community 1 during week 6. This was associated with the introduction of the second collaborative task, which was to collect questions and answers for the course's final exam. The instructor, reflecting upon this change in the CLaP index, emphasized that "*compared to the first task (a collective concept mapping), students more easily related the second task to their own motivations of doing well in their exam and observing an immediate impact on their individual outcomes. It is very interesting to see that the change of the collaborative task in week 6 led to a rise in collaborative behaviours of the cohort*." Such insights show that CLaP type of process evaluation is not only valuable for instructors to reflect on the class dynamics retrospectively, but also might help them design more relevant collaborative tasks in the future.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

5. DISCUSSION AND CONCLUSIONS

In this paper, we present an argument for investigating the collaboration process in learning based on CCLT. Our approach explains the collaboration process as a function of the relationship between interactivity gains and coordination costs.

5.1. Practical Implications

As demonstrated in section 4.2.3, CLaP analysis can be used by instructors to open-up the black box of the collaboration process taking place in their class. It can be used not just by presenting the trends in collective actions, but also by suggesting an explanation, by breaking it down to its components of coordination costs and interactivity gains. It is important to note here that we do not present CLaP as a single aggregate value of different proxies, which could be calculated to assess the level of collaboration of a community. This is mainly because the different measures of both interactivity gains and coordination costs are not interchangeable. Each measure adds its own dimension, reflecting the multidimensional nature of the process of collaboration in learning.

5.2. Theoretical Implications

Our theoretical contribution is manifested by turning the CLT flashlight onto collaborative learning from the perspective of the collective unit of analysis, rather than the individual learner. Collaboration in learning is a process rather than a result. It is not ideal to investigate collaboration in learning as a process by artificially considering each individual's cognitive system as a separate entity without evaluating the interdependent network effect. Hence, we perceive collaboration in learning as the process of establishing and growing an interactions network among learners and knowledge artefacts. We show that using social network analysis to calculate the relationship between coordination costs and interactivity gains can shed light on the collaboration processes of learning groups. To the best of our knowledge, this is the first attempt to operationalize CCLT within learning analytics.

5.3. Limitations and Future Studies

The CLaP framework is designed to be used cross-medium, cross-subject, and cross-instructional design, in the sense that it does not depend on content, or any other specifics of subject domain or design. On the other hand, it helps teachers and learners understand the impact of a change in these variables on the collaboration process. This is both a strength and a limitation. Thus, as shown in our case studies, CLaP analysis must be combined with other measures, such as instructor reflections and expert rating of the quality of collaboration (Chounta et al., 2014). This study is mainly conceptual, presenting an analysis framework for the collaboration process. Therefore, further work must be undertaken to empirically validate the approach using various data sources and other proxies. SNA literature offers numerous network formation properties that could potentially be used to measure coordination costs and interactivity gains. We name just a few proxies and demonstrate their use in section 4. However, we encourage future studies to experiment with other network measures in order to refine and validate the selection of best-fit measures for each learning context. In this paper, our intention is not to provide a conclusive, exhaustive list of proxies, but rather to demonstrate a potential operationalization of our theoretical arguments. Next, we plan to validate the CLaP approach in near real-time visualizations in a MOOC, validating its impact with more instructors and learner groups. We also plan to further develop CLaP in a co-design process with a group of instructors to produce some guidelines for possible recommended action plans resulting from CLaP reflections, and for developing guidelines as to which proxies of interactivity gains and coordination costs best suit different pedagogical goals.

DECLARATION OF CONFLICTING INTEREST

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The authors declared no financial support for the research, authorship, and/or publication of this article.

ACKNOWLEDGEMENTS

We wish to deeply thank Prof. Sheizaf Rafaeli for his contribution in refining this work

References

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York: Holt, Rinehart & Winston.

Bernstein, E., Shore, J., & Lazer, D. (2018). How intermittent breaks in interaction improve collective intelligence, PNAS: *Proceedings of the National Academy of Sciences of the United States of America, 115*(35), 8734–8739. https://dx.doi.org/10.1073/pnas.1802407115

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science, 323*(5916), 892–895. http://dx.doi.org/10.1126/science.1165821

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century US workforce*. Washington, DC: Partnership for 21st Century Skills.

Chiocchio, F., & Essiembre, H. (2009). Cohesion and performance: A meta-analytic review of disparities between project teams, production teams, and service teams. *Small Group Research, 40*(4), 382–420. http://dx.doi.org/10.1177/1046496409335103

Chounta, I. A., Hecking, T., Hoppe, H. U., & Avouris, N. (2014). *Two make a network: Using graphs to assess the quality of collaboration of dyads.* CYTED-RITOS International Workshop on Groupware, 53–66. Cham, Switzerland: Springer. https://dx.doi.org/10.1007/978-3-319-10166-8_5

Claros, I., Cobos, R., & Collazos, C. A. (2016). An approach based on social network analysis applied to a collaborative learning experience. *IEEE Transactions on Learning Technologies, 9*(2), 190–195. http://dx.doi.org/10.1109/TLT.2015.2453979

Collazos, C., Padilla, N., Pozzi, F., Guerrero, L., & Gutierrez, F. (2014). Design guidelines to foster cooperation in digital environments. *Technology, Pedagogy and Education, 23*(3), 375–396. https://dx.doi.org/10.1080/1475939X.2014.943277

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114. http://dx.do.org/10.1017/S0140525X01003922

Cukurova, M. (2018). A syllogism for designing collaborative learning technologies in the age of AI and multimodal data. Proceedings of the 13th European Conference on Technology Enhanced Learning (EC-TEL 2018), 3–5 September 2018, Leeds, UK (pp. 291–296). *Lecture Notes in Computer Science*, Springer. https://dx.doi.org/10.1007/978-3-319-98572-5_22

Cukurova, M., Bennett, J. M., & Abrahams, I. Z. (2018). Students' knowledge acquisition and ability to apply knowledge into different science contexts in two different independent learning settings. *Research in Science and Technological Education*, *36*(1), 17–34. http://dx.doi.org/0.1080/02635143.2017.1336709

Cukurova, M., Luckin, R., Millán, E., & Mavrikis, M. (2018). The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. *Computers & Education, 116,* 93–109. https://dx.doi.org/10.1016/j.compedu.2017.08.007

Cukurova, M., Luckin, R., Millán, E., Mavrikis, M., & Spikol, D. (2017). Diagnosing collaboration in practice-based learning: Equality and intra-individual variability of physical interactivity. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), Data Driven Approaches in Digital Education, pp. 30–42. EC-TEL 2017. *Lecture Notes in Computer Science*, *vol. 10474.* Springer, Cham. https://dx.doi.org/10.1007/978-3-319-66610-5_3

Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research, 13*(1), 9–19. https://dx.doi.org/10.1016/0883-0355(89)90013-X

Dede, C. (2010). Comparing frameworks for 21st century skills. *21st Century Skills: Rethinking How Students Learn,* 20, 51–76.

De Laat, M., Lally, V., Lipponen, L., & Simons, R. J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *International Journal of Computational Supported Collaborative Learning*, *2*(1), 87–103. https://dx.doi.org/10.1007/s11412-007-9006-4

Dillenbourg, P. (1999). What do you mean by "collaborative learning"? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–16). Amsterdam, Netherlands: Elsevier.

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*(6), 469–493. https://dx.doi.org/10.1016/0047-2484(92)90081-J

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge, UK: Cambridge University Press.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). Teaching adolescents to become learners. The role of noncognitive factors in shaping school performance: A critical literature review. Chicago, IL: University of Chicago Consortium on Chicago School Research.

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

Gardner, J. (2011). *Assessment and learning*. Thousand Oaks, CA: Sage Publications.

Goyal, S., & Vega-Redondo, F. (2005). Network formation and social coordination. *Games and Economic Behavior, 50*(2), 178–207.

Hutchins, E. (1996). *Cognition in the wild.* Cambridge, MA: MIT Press.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Ickes, W. (2002). Social cognition: Subjective and intersubjective paradigms. *New Review of Social Psychology, 1*(June), 112–121.

Ickes, W., & Gonzalez, R. (1996). "Social" cognition and social cognition: From the subjective to the intersubjective. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition?* Research on socially shared cognition in small groups (pp. 285–308). Thousand Oaks, CA: Sage Publications. https://dx.doi.org/10.4135/9781483327648.n12

Janssen, J., Kirschner, F., Erkens, G., Kirschner, P. A., & Paas, F. (2010). Making the black box of collaborative learning transparent: Combining process-oriented and cognitive load approaches. *Educational Psychology Review*, 22(2), 139–154. https://dx.doi.org/10.1007/s10648-010-9131-x

Johnson, D. W., & Johnson, R. T. (2002). Learning together and alone: Overview and meta-analysis. *Asia Pacific Journal of Education, 22*, 95–105. https://dx.doi.org/10.1080/0218879020220110

Johnson, D. W., Johnson, R. T., Roseth, C. J., & Seob Shin, T. (2014). The relationship between motivation and achievement in interdependent situations. *Journal of Applied Social Psychology, 44*(9), 622–633. https://dx.doi.org/10.1111/jasp.12280

Kalman, R. E. (1963). Mathematical description of linear dynamical systems. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control, 1(*2), 152–192. https://dx.doi.org/10.1016/S0076-5392(08)63251-8

Kent, C., Laslo, E., & Rafaeli, S. (2016). Interactivity in online discussions and learning outcomes. *Computers & Education, 97,* 116–128. https://dx.doi.org/10.1016/j.compedu.2016.03.002

Kent, C., & Rechavi, A. (2018). Deconstructing online social learning: Network analysis of the creation, consumption and organization types of interaction. International Journal of Research & Method in Education, 43(1), 1–22. http://dx.doi.org/10.1080/1743727X.2018.1524867

Kent, C., Rechavi, A., & Rafaeli, R. (2019). Networked learning analytics: A theoretically informed methodology for collaborative learning analytics. In Y. Kali, A. Baram-Tsabari, & A., Schejter (Eds.), *Learning in a networked society: Spontaneous and designed technology enhanced learning communities* (pp. 145–175). Springer. https://dx.doi.org/10.1007/978-3-030-14610-8_9

Kester, L., & Paas, F. (2005). Instructional interventions to enhance collaboration in powerful learning environments. *Computers in Human Behavior, 21*, 689–696. https://dx.doi.org/10.1016/j.chb.2004.11.008

Kirschner, P. A., Kirschner, F., & Janssen, J. (2014). The collaboration principle in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning, 2nd ed.* (pp. 547–575). New York: Cambridge University Press. https://dx.doi.org/10.1017/CBO9781139547369.027

Kirschner, F., Paas, F., & Kirschner, P. A. (2011). Task complexity as a driver for collaborative learning efficiency: The collective working-memory effect. *Applied Cognitive Psychology, 25*(4), 615–624. https://dx.doi.org/10.1002/acp.1730

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86. https://dx.doi.org/10.1207/s15326985ep4102_1

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, R. J. (2018). From cognitive load theory to collaborative cognitive load theory. International Journal of Computer-Supported Collaborative Learning, 13, 213–233. https://dx.doi.org/10.1007/s11412-018-9277-y

Knight, S., Buckingham Shum, S., & Littleton, K. (2013). Epistemology, pedagogy, assessment and learning analytics, *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13),* 8–12 April 2013, Leuven, Belgium (pp. 75–84). New York: ACM. https://dx.doi.org/10.1145/2460296.2460312

Kwon, K., Liu, Y. H., & Johnson, L. P. (2014). Group regulation and social-emotional interactions observed in computer supported collaborative learning: Comparison between good vs. poor collaborators. *Computers & Education, 78*, 185–200. https://dx.doi.org/10.1016/j.compedu.2014.06.004

Liu, Y. Y., Slotine, J. J., & Barabási, A. L. (2011). Controllability of complex networks. *Nature, 473,* 167–173. https://dx.doi.org/10.1038/nature10011

Martinez, A., Dimitriadis, Y., Rubia, B., Gomez, E., & De La Fuente, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions. *Computers & Education, 41*(4), 353–368. https://dx.doi.org/10.1016/j.compedu.2003.06.001

Murnane, R. J., & Levy, F. (1996). *Teaching the new basic skills: Principles for educating children to thrive in a changing economy.* New York: Free Press.

Newman, M. E. J. (2001). *The structure of scientific collaboration networks.* PNAS: Proceedings of the National Academy of Sciences of the United States of America, 98(2), 404–409. https://dx.doi.org/10.1073/pnas.021544898

OECD (Organisation for Economic Co-operation and Development). (2015). PISA 2015 collaborative problem solving

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

framework. Oxford, UK: OECD Publishing.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. Educational Psychology Review, 24(1), 27–45. https://dx.doi.org/10.1007/s10648-011-9179-2

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology, 49*(1), 345–375. https://dx.doi.org/10.1146/annurev.psych.49.1.345

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193–198. http://dx.doi.org/10.1037/h0049234

Rabbany, R., Elatia, S., Takaffoli, M., & Zaïane, O. R. (2013). Collaborative learning of students in online discussion forums: A social network analysis perspective. In A. Peña-Ayala (Ed.), *Educational data mining: Applications and trends* (pp. 441–466). Springer. https://dx.doi.org/1007/978-3-319-02738-8_16

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–197). Berlin: Springer-Verlag. https://dx.doi.org/10.1007/978-3-642-85098-1_5

Siemens, G. (2005). A learning theory for the digital age. *Instructional Technology and Distance Education, 2*(1), 3–10.

Slavin, R. E. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales de Psicología/Annals of Psychology, 30*, 785–791. https://dx.doi.org/10.6018/analesps.30.3.201201

Soboroff, S. D. (2012). Group size and the trust, cohesion, and commitment of group members [Unpublished doctoral dissertation]. University of Iowa.

Stahl, G. (2010). Group cognition as a foundation for the new science of learning. In M. S. Khine & I. M. Saleh (Eds.), *New science of learning: Cognition, computers and collaboration in education* (pp. 23–44). https://dx.doi.org/10.1007/978-1-4419-5716-0

Strijbos, J. W. (2011). Assessment of (computer-supported) collaborative learning. *IEEE Transactions on Learning Technologies, 4*(1), 59–73. https://dx.doi.org/10.1007/s40692-014-0009-7

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://dx.doi.org/10.1207/s15516709cog1202_4

Tsai, W. (2002). Social structure of "coopetition" within a multiunit organization: Coordination, competition, and intraorganizational knowledge sharing. *Organization Science, 13(*2), 179–190. https://dx.doi.org/10.1287/orsc.13.2.179.536

Tulin, M., Pollet, T. V., & Lehmann-Willenbrock, N. (2018). Perceived group cohesion versus actual social structure: A study using social network analysis of egocentric Facebook networks. *Social Science Research, 75*, 161–175. https://dx.doi.org/10.1016/j.ssresearch.2018.04.004

Veerman, A., & Veldhuis-Diermanse, E. (2001). Collaborative learning through computer-mediated communication in academic education. In *Proceedings of the 1st European Conference on Computer-Supported Collaborative Learning* (Euro-CSCL 2001), 22–24 March 2001, Maastricht, Netherlands (pp. 625–632). International Society of the Learning Sciences.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological process*. Cambridge, MA: Harvard University Press.

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

West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F., & Gabrieli, J. D. (2016). Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis, 38*(1), 148–170. https://dx.doi.org/10.3102/0162373715597298

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science, 330*(6004), 686–688. http://dx.doi./org/10.1126/science.1193147

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

APPENDIX A: SNA DEFINITIONS

- The average path length (APL) of a network is the average number of ties between all pairs of nodes in the network.
- The betweenness centrality of a given node (i) is the number of shortest paths between any other two nodes that pass through i. High betweenness means that a node is connecting many pairs of nodes that would not have been connected without it.
- The closeness centrality of node i is inverse of the sum of the length of shortest paths between i and all other nodes in the network. It is a measure of how central a node is in its network.
- The clustering coefficient of node i is the proportion of ties connecting i's neighbours to each other out of the number of ties that could possibly exist between them.
- The degree of a node i in a network is the number of ties connected to it. In a directed network, it is the number of ties outgoing from i to its neighbours plus the number of ties incoming into i from its neighbours.
- Density is the number of ties (or edges) in a network, as a proportion of the maximum possible number of ties that could have existed potentially within that network. In case there is more than one edge between two nodes, the density values can be larger than one.
- The diameter of a network is the longest path between any two nodes in the network.
- A strongly connected component is a sub-network in which each pair of nodes is connected by a path.
- Reciprocity is the proportion of ties that show two-way communication (also called reciprocal or mutual ties) out of the total number of existing ties.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

APPENDIX B: LIGILO SCREENSHOTS

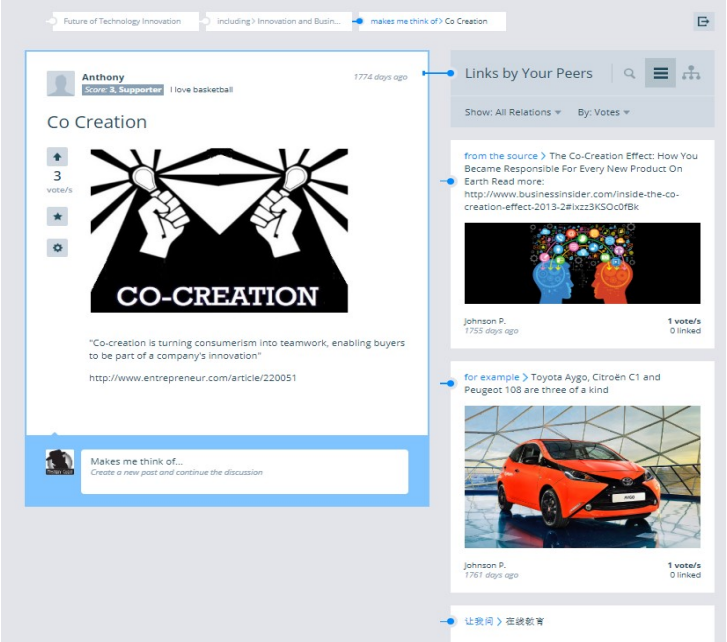

Figure 5. Ligilo's basic view. Posts to the left are connected by blue tags to posts on the right.

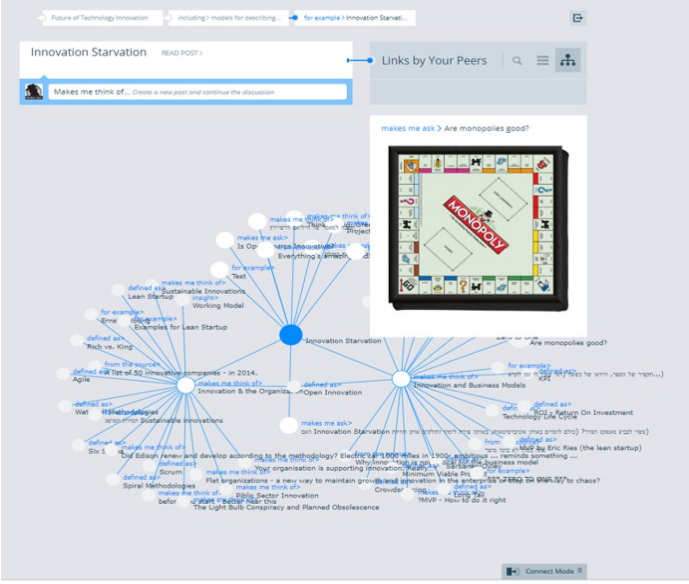

Figure 6. Bird's-eye view of posts in Ligilo.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

APPENDIX C: OPTIONAL INTERACTIVITY GAINS AND TRANSACTION COST METRICS FOR COMMUNITY 1 AND COMMUNITY 2

As discussed in section 3.2.1, we suggested reciprocity, density, and average clustering coefficient values as proxies to calculate interactivity gains. Figure 7 shows that the reciprocity, density, and average clustering coefficients of both communities increase with time, with similar trends, whilst Community 1 almost always shows higher values.

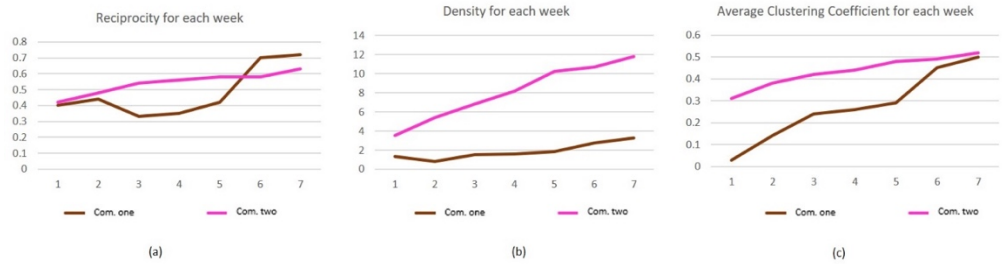

Figure 7. Values each week for Community 1 (brown) and Community 2 (pink): (a) reciprocity, (b) density, and (c) average clustering coefficient.

On the other hand, we use coefficients of variance in degree and betweenness, average closeness centrality, average path length, number of strongly connected components, and diameter in each week to evaluate the coordination costs. Please note that we have used the coefficient of variation instead of the variance in order to neutralize the difference of the average values between the communities. The coefficient of variation is the ratio between the standard deviation to the mean.

As can be seen from Figure 8, although both communities present similar trends in the investigated aspects of interactivity gains, their coordination costs are quite different. Community 2 shows a steady decrease in variance of betweenness and degree, while this is not the case in Community 1 which is generally showing higher coordination costs. On the other hand, in terms of distance metrics, whereas Community 2's distance decreases and closeness increases, Community 1's distance increases and closeness decreases. Similarly, the number of strongly connected components constantly decreases in Community 2, whereas it sharply increases initially in Community 1 and then plummets.

Figure 8. Weekly values for Community 1 (brown) and Community 2 (pink): (a) coefficients of variance in degree, (b) coefficient of variance in betweenness, (c) average closeness centrality, (d) average path length, (e) number of strongly connected components, and (f) diameter.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)