2016

Analyzing Social Construction of Knowledge Online by Employing Interaction Analysis, Learning Analytics, and Social Network Analysis

Charlotte Nirmalani Gunawardena
University of New Mexico, lani@unm.edu

Nick V. Flor
University of New Mexico

David Gomez
University of New Mexico

Damien Sanchez
University of New Mexico

Recommended Citation

Gunawardena, C. N., Flor, N. V., Gomez, D., & Sanchez, D. (2016). Analyzing social construction of knowledge Online by Employing Interaction Analysis, Learning Analytics, and Social Network Analysis. The Quarterly Review of Distance Education, 17(3), 35-60.

Citation:
Gunawardena, C. N., Flor, N. V., Gomez, D., & Sanchez, D. (2016). Analyzing social construction of knowledge Online by Employing Interaction Analysis, Learning Analytics, and Social Network Analysis. *The Quarterly Review of Distance Education,* 17(3), 35-60.

Analyzing Social Construction of Knowledge Online by Employing Interaction Analysis, Learning Analytics, and Social Network Analysis

Charlotte N. Gunawardena [a], Nick V. Flor [b], David Gómez [a], and Damien Sánchez [a]

[a] University of New Mexico, Organization, Information & Learning Sciences, College of University Libraries and Learning Sciences,
MSC 05 3020, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA

[b] University of New Mexico, Anderson School of Management, MSC 05 3090, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA

Abstract

This article examines research methods for analyzing social construction of knowledge in online discussion forums. We begin with an examination of the Interaction Analysis Model (Gunawardena, Lowe, & Anderson, 1997) and its applicability to analyzing social construction of knowledge. Next, employing a dataset from an online discussion, we demonstrate how interaction analysis can be supplemented by employing other research techniques such as learning analytics and Social Network Analysis that shed light on the social dynamics that support knowledge construction. Learning analytics is the application of quantitative techniques for analyzing large volumes of distributed data ("big data") in order to discover the factors that contribute to learning (Long & Siemens, 2011, p. 34). Social Network Analysis characterizes the information infrastructure that supports the construction of knowledge in social contexts (Scott, 2012). By combining interaction analysis with learning analytics and Social Network Analysis, we were able to conceptualize the process by which knowledge construction takes place in online platforms.

Introduction

Asynchronous discussion forums have become the main vehicle through which the teaching learning process in online courses is facilitated. Even Massive Open Online Courses (MOOCs) are incorporating discussion forums to enhance interaction in otherwise teacher-directed and -designed online courses. Yet the question remains as to how learning within these discussion forums can be deciphered and understood. New methodologies such as learning analytics have provided means to analyze large sets of data from online courses. Learning analytics is the application of quantitative techniques for analyzing large volumes of distributed data ("big data") in order to discover the factors that contribute to learning. We explore how the process of learning, especially the process of knowledge construction in online asynchronous discussions, can be mapped, analyzed, and understood using approaches such as interaction analysis and learning analytics. We discuss these methods of analyses from our own personal perspectives and experiences analyzing transcripts of a computer discussion where students engaged in collaborative learning.

Interaction Analysis and Learning Analytics

Jordan and Henderson (1995) observe that interaction analysis approaches view learning as a distributed, ongoing social process, in which evidence that learning is occurring or has occurred must be found in understanding the ways in which people collaboratively engage in learning. Interaction analysis looks at "the interaction of human beings with each other and with objects in their environment. It investigates human activities, such as talk, nonverbal interaction, and the use of artifacts and technologies" (Jordan & Henderson, 1995, p. 39). Thus, interaction analysis considers interaction as a function of the reciprocal influence among human beings, objects, and their environment.

ANALYZING SOCIAL CONSTRUCTION OF KNOWLEDGE ONLINE

The Interaction Analysis Model (IAM) was developed by Gunawardena, Lowe, and Anderson (1997) to qualitatively examine these interactions (or discussions) during the phases of knowledge construction. The IAM was employed to examine the interaction that occurred in an online global debate to determine whether knowledge was constructed within the group through talk and dialogue, and whether participants changed their understanding or developed new knowledge as a result of group interaction. Based on social constructivist theory (Vygotsky, 1978), the model describes five Phases of knowledge co-construction: *sharing and comparing* constitute Phase I; *dissonance* is the focus of Phase II; *negotiation and co-construction* comprise Phase III, *testing tentative constructions* is incorporated in Phase IV, and *statements and application of newly constructed knowledge* are at the heart of Phase V. See Figure 1 for a detailed description of the IAM, including numerous operations for each of its five phases. The model itself serves as a framework that defines social construction of knowledge as a function of interaction, which is understood as reciprocal cognitive influence among individuals.

PHASE I: SHARING/COMPARING OF INFORMATION. Stage one operations include:	
A. A statement of observation or opinion	[PhI/A]
B. A statement of agreement from one or more other participants	[PhI/B]
C. Corroborating examples provided by one or more participants	[PhI/C]
D. Asking and answering questions to clarify details of statements	[PhI/D]
E. Definition, description, or identification of a problem	[PhI/E]
PHASE II: THE DISCOVERY AND EXPLORATION OF DISSONANCE OR INCONSISTENCY AMONG IDEAS, CONCEPTS OR STATEMENTS. (This is the operation at the group level of what Festinger [20] calls cognitive dissonance, defined as an inconsistency between a new observation and the learner's existing framework of knowledge and thinking skills.) Operations which occur at this stage include:	
A. Identifying and stating areas of disagreement	[PhII/A]
B. Asking and answering questions to clarify the source and extent of disagreement	[PhII/B]
C. Restating the participant's position, and possibly advancing arguments or considerations in its support by references to the participant's experience, literature, formal data collected, or proposal of relevant metaphor or analogy to illustrate point of view	[PhII/C]
PHASE III: NEGOTIATION OF MEANING/CO-CONSTRUCTION OF KNOWLEDGE	
A. Negotiation or clarification of the meaning of terms	[PhIII/A]
B. Negotiation of the relative weight to be assigned to types of argument	[PhIII/B]
C. Identification of areas of agreement or overlap among conflicting concepts	[PhIII/C]
D. Proposal and negotiation of new statements embodying compromise, co-construction	[PhIII/D]
E. Proposal of integrating or accommodating metaphors or analogies	[PhIII/E]
PHASE IV: TESTING AND MODIFICATION OF PROPOSED SYNTHESIS OR CO-CONSTRUCTION	
A. Testing the proposed synthesis against "received fact" as shared by the participants and/or their culture	[PhIV/A]
B. Testing against existing cognitive schema	[PhIV/B]
C. Testing against personal experience	[PhIV/C]
D. Testing against formal data collected	[PhIV/D]
E. Testing against contradictory testimony in the literature	[PhIV/E]
PHASE V: AGREEMENT STATEMENT(S)/APPLICATIONS OF NEWLY-CONSTRUCTED MEANING	
A. Summarization of agreement(s)	[PhV/A]
B. Applications of new knowledge	[PhV/B]
C. Metacognitive statements by the participants illustrating their understanding that their knowledge or ways of thinking (cognitive schema) have changed as a result of the conference interaction	[PhV/C]

*Figure 1.*The Interaction Analysis Model (IAM) developed by Gunawardena, Lowe, and Anderson, 1997.

The function of interaction in the process of knowledge construction can be further explored by newer methods such as learning analytics. There are many definitions of learning analytics (Long & Siemens, 2011, p. 34), but common to them all is the application of tools and methods for extracting, analyzing, and visualizing learner data. The scientific goal is to understand the factors, structures, and processes of learning. The applied goal is to leverage this understanding to design cognitive artifacts that transform the learning process in a beneficial way. The field of learning analytics has grown with the need to analyze large sets of data on learning generated by MOOCs, informal learning environments, and open social media platforms like Twitter, where people can interact with thousands of other people distributed globally.

Literature Review

Previous Research Using the Interaction Analysis Model (IAM)

The IAM has been used globally to analyze online discussions. It has helped several researchers to determine occurrences of social construction of knowledge in online discussions (e.g., Buraphadeja & Dawson, 2008; Chai & Tan, 2009; De Wever, Van Keer, Schellens, & Valcke, 2007; De Wever, Van Winckel, & Valcke, 2008; De Wever, Van Keer, Schellens, & Valcke, 2009; De Wever, Van Keer, Schellens, & Valcke, 2010; Heo, Lim, & Kim, 2010; Hou, Chang, & Sung, 2008; Hou, Chang, & Sung, 2009; Islas, 2004; Paulus, 2007; Sing & Khine, 2009; Schellens, Van Keer, De Wever, & Valcke, 2007; and, Tan, Ching, & Hong, 2008). While similar to qualitative content analysis, the IAM differs from content analysis (Krippendorff, 2004), which uses mutually exclusive categorical variables that measure objectively the existence and/or frequency of data from separate units of analysis without considering their relationship. The IAM is focused on examining the relationship between interactions.

Like Heo, Lim, and Kim (2010), as well as Paulus (2007) relied on the IAM to document how online communication contributed to the social maintenance of a group of students. Similarly, Schellens, Van Keer, De Wever, and Valcke (2007), as well as De Wever, Van Keer, Schellens, and Valcke (2009, 2010), analyzed the impact of assigning and rotating roles among group members during the social negotiation phases in discussion forums. "It is worth noting that IAM explicitly attributes the success of the asynchronous discussion-based online learning and critical thinking to social constructivism" (Buraphadeja & Dawson, 2008, p. 139). To reiterate, Chai and Tan, (2009) pointed out "the IAM was selected because it is premised on a social constructivist theoretical foundation" that was consistent with their study.

In a more recent review of studies that employed the IAM, Lucas, Gunawardena, and Moreira, 2014 observed that despite the existence of different models, the IAM remains one of the most used in the study of online interaction. Results reinforced the adequacy of the model for analyzing knowledge construction in different types of communication tools, but also suggested the need to look at how learning is orchestrated and the importance of redefining some aspects of the model. For example, Lucas and Moreira (2010) observed that despite focusing on interaction as the vehicle for knowledge construction, the IAM lacks the capability to demonstrate the social and interaction dynamics that go beyond the categorization proposed for the knowledge construction phases. Furthermore, they state that the IAM does not provide an accurate picture of the progress and development of students' knowledge which can be complemented by procedures such as Social Network Analysis, which enables the study of individual interactions in relation to the group, and provides a better understanding and accurate visualization of the contribution of such interactions to the group's collaborative knowledge construction process. Therefore, this study seeks to explore how learning analytics can shed more light on the social dynamics that accompany knowledge construction.

Applying Learning Analytics to Textual Data

To date, much of learning analytics has focused on applying statistical methods to large learner datasets. The methods fall into different categories depending on whether the researcher's intent is to describe the data (descriptive analytics), forecast from the data (predictive analytics), or recommend a course of action (prescriptive analytics). Descriptive methods include: traditional descriptive statistics, social media metrics, clustering, principal components analysis, sentiment analysis, and Social Network Analysis. Predictive methods include: multiple regressions, logistic regressions, and non-linear predictive models such as neural networks. Prescriptive methods combine both descriptive and predictive methods, to recommend courses of action. See Konstan and Riedl (2012) and Herlocker, Konstan, and Riedl (2000) for descriptions of the methods used by Amazon and Netflix automated recommenders.

Qualitative researchers can expand their analysis of online discourse by employing analytics. However, the kinds of quantitative methods used in modern analytics are complementary to qualitative methods because they fundamentally lead to different kinds of theories, namely, *variation theories* versus *process theories* (Mohr, 1982). Quantitative methods, with their emphases on clusterings, correlations, factors, and regressions are categorically inadequate for describing the rich processes through which knowledge is co-constructed in a socially-distributed cognitive system (see Flor, Coulson, and Maglio, 2006 for an example of such a process), and therefore, employing both qualitative interaction analysis and quantitative learning analytic methods can expand our understanding of online social construction of knowledge.

Despite the fundamental difference in the kinds of theories generated, analytics can assist qualitative researchers by highlighting areas of interest in their data and suggesting future areas of research and exploration. We believe one of the exciting challenges with learning analytics lies in the application of techniques like data scraping, statistics, programming, and visualization to qualitative data, particularly when guided by models such as the IAM. Thus, the IAM and analytics are complimentary methods that can produce more robust findings.

Data Scraping

One very common analytics operation is data scraping, which is a colloquial and more general term for what used to be called web scraping using a web spider (Eichmann, 1995). This operation involves the automated extraction of data from an end-user source into a computer program, where various kinds of analyses can then be performed more easily. An example is a "bot" that extracts social media data into a relational database. Once in the relational database, an analyst can run database queries to filter and analyze the content. In this analysis we demonstrate how the basic scraping of words from a discussion forum into an Excel sheet can provide insights for an IAM analyst about the kinds of topics around which knowledge was socially constructed.

Sentiment Analysis

Sentiment analysis is a technique that also takes postings as input, but instead of reporting explicitly represented, factual information, it attempts to discover and report implicit, subjective information about a posting (for reviews, see Pang and Lee, 2008; Liu and Zhang, 2012). The classic example of sentiment analysis is the mining of product reviews on sites like Amazon.com for positive or negative opinions (Feldman, 2013). The most basic technique for sentiment analysis is the counting of positive opinion words ("great", "amazing") and negative opinion words ("bad", "awful") in a posting, and then calculating an overall score (Ding, Liu, & Yu, 2008). More advanced techniques use natural language processing and neural networks to map postings to sentiment (Socher, Perelygin, Wu, Chuang, Manning, Ng, & Potts, 2013).

Social Presence Analysis

Taking a more theory-driven approach to the mining and analysis of textual data requires counting only those words and phrases associated with a specific theoretical construct and testing if there is a relationship between that construct and the phases of the IAM. Since one of the critiques of the IAM is its inability to account for the social processes that accompany knowledge construction, we selected the construct “social presence,” (the degree to which a person is perceived as a ‘real person’ in mediated communication), which has been shown to be a strong predictor of learner satisfaction (Gunawardena & Zittle, 1997); perhaps a factor in student achievement and retention (Rourke, Anderson, Garrison, & Archer, 1999); and an important aspect of the online sociocultural environment (Tu, 2001; Whiteside & Garrett Dikkers, 2012; Yildiz, 2009). We call this a *social presence analysis*, and it is similar to a sentiment analysis, but focusing on the social presence score of a posting. We believe investigating social presence through learning analytics methods can shed light on whether social presence is related to knowledge construction.

Co-Word Analysis

The rise of social media networks like Twitter, and the widespread availability of software for automatically scraping and generating social network diagrams, have led to a renewed interest in Social Network Analysis (Bastian, Heymann, & Jacomy, 2009; Java, Song, Finin, & Tseng, 2007). Social Network Analysis (SNA) is a set of methods for analyzing the relational aspects of social structures (Scott, 2012; Wasserman & Faust, 1994) and it is typically used to study the information exchanged between people in groups and in communities.

However, nothing prevents a researcher who studies knowledge construction from applying SNA to words instead of people. The rationale is that both individual knowledge construction and social construction of knowledge are processes of negotiation of meaning, that is, word sequences not only create the mental spaces needed for negotiation of meaning, but they also place elements and schemata into spaces, and they blend these elements into new schemata. Thus, researchers can identify words that are central to knowledge co-construction by analyzing and graphing the relationship between them, as opposed to analyzing and graphing the relationship between people who are central to knowledge co-construction. Therefore, we will conduct what we refer to as *co-word analysis,* according to the central tenets of SNA.

Social Network Analysis

In the context of online higher education, there are three themes in the literature about student-to-student interaction in online discussion forums, namely, (a) studies focused on knowledge construction such as the ones mentioned in previous research using the IAM, (b) studies focused on social networks (e.g., Haythornthwaite & De Laat, 2010; Firdausiah & Yusof, 2013; and Toikkanen & Lipponen, 2011), and (c) studies that are a combination of both. For instance, researchers typically approach knowledge construction by applying qualitative methods that rely heavily on analysis of text, while they typically approach social networks by applying SNA, which relies on centrality (importance) measures and diagrams of interaction.

From a SNA perspective, actors (also known as nodes) and their actions are viewed as interdependent rather than independent autonomous units. The actors in this analysis are students. Relational ties (linkages, also known as arcs or edges) between students are interaction channels for transfer, or "flow", of information through postings in online discussion forums. Social network diagrams can depict student roles as lasting patterns of interactions among students. Thus, each student has interactions with other students, each of whom interacts with a few, some, or many others, and so on. Therefore, the construct “social network” refers to the finite set of students and the interactions among them in a discussion forum.

In addition, from the point of view of SNA, density is a holistic measure of a social network. According to Faust (2006), network density (*d*) is defined as the number of ties (interactions) in the network divided by the possible by number of ties (interactions), as illustrated by the following formula:

d= actual ties (interactions)/maximum possible ties (interactions)

Generally speaking, SNA reveals interaction patterns and produces diagrams as a way of "x-raying" or mapping social networks.

Purpose

Lucas and Moreira (2010) observed that the IAM lacks the capability to demonstrate the social and interaction dynamics that go beyond the categorization proposed for the knowledge construction phases, which can be complemented by procedures such as SNA. Applying learning analytics enables the study of individual interactions in relation to the group, and provides a better understanding and accurate visualization of the contribution of such interactions to the group's collaborative knowledge construction process. Therefore the purpose of this article is to extend our understanding of social construction of knowledge that emerges from interaction analysis conducted with the IAM with learning analytics and SNA.

Method

This section outlines how we used interaction analysis based on the IAM and learning analytics methods to examine social construction of knowledge in online discussions. Our data set contained 42 postings generated by 15 students who participated in an online discussion on the topic of "culture", which was part of a graduate class on e-learning. This same data set was used for all analyses reported here. First we conducted interaction analysis based on the IAM using qualitative content analysis methods, and then we analyzed the same discussion using the five aforementioned learning analytics methods to shed further light on the IAM analyses in order to expand our understanding of the social dynamics of knowledge construction.

Interaction analysis was conducted using qualitative coding for meaning using the IAM and assigning the IAM Phases and associated operations, as explained in Figure 1, to each discussion posting. Note this method generated descriptive quantitative data on the number of occurrences of each phase in the discussion. We examined the entire transcript and then took as input each posting, which was the unit of analysis, and mapped discourse to the IAM Phases and specific operations within the phases. We also reviewed the links to prior postings to determine the process of interaction. For example, this was executed to determine if the posting built on a prior posting in order to assign the appropriate IAM Phase to which the message belongs.

The following example illustrates the application of the IAM. The discussion transcript we analyzed tasked students with discussing definitions of culture and e-learning and the relationship between the two. Using the IAM as a framework, we looked for evidence of operations within postings, which correspond to the different IAM Phases. The analysis was done by highlighting the operations in the transcript in addition to noting the number of occurrences of these operations. This analysis was performed for every posting in the discussion. Note that postings can provide evidence of knowledge construction at multiple levels. In the following example, JG provides evidence of two operations within Phase I (coded in bold).

Thread	Initial	PhI/A	Ph/B	Ph/C	Ph/D	Ph/E	PhII/A	PhII/B	PhII/C	PhIII/A
Jack Griffith: Definition of Culture and eLearning My own definition of culture is: A group of people who share and follow a set of spoken and in some cases unspoken beliefs, ideas, traditions, values, and knowledge both tacit and explicit. My own definition of eLearning is: Learning and teaching that utilizes technology and electronic media formats over traditional resources. **The three pillars that determine the success or failure of e-learning programs are the interconnectedness among (1) person, (2) behavior, and (3) environment.** These are the three major areas that interventions should target. 1.E-learners' cognitive skills: E-learners must have the prerequisite knowledge and skills necessary to participate in e-learning. Computer competency through training, and practice, and time management skills are essential. 2.Environment: Organizations must support e-learning by offering a supportive culture, incentives, models, resources, and fostering e-learning self-efficacy. 3. Belief and behavior: E-learners' must have high e-learning self-efficacy and the appropriate behavioral skills such as taking responsibility for learning (Mungania, 2003). From this example in the research article "The Seven eLearning Barriers Facing Employees," I begin to see where eLearning and culture can have some commonalities. I believe that like these 3 pillars that determine the success or failure of eLearning programs, these 3 pillars can create or remove barriers within a culture. I notice how culture is associated with traits such as cognitive skills of the culture, the environment in which the culture thrives, and the beliefs and behaviors of the culture. **As we implement eLearning strategies into different cultures, we must look at these 3 pillars to create and maintain a successful eLearning program for the target culture.** ELearning, like culture is not one size fits all and requires some different approaches to how it is employed in order to suit the target culture. I look forward to your definitions and to this discussion, have a great week everyone!	JG	2								

Figure 2. Example of the Excel IAM transcript coding sheet.

Data scraping was conducted using two types of software (a) Microsoft Excel for its built-in descriptive-statistics functions and (b) a custom parser of unique words and unique word-pairs, which filtered out punctuation, then displayed the number of occurrences of both words and word-pairs. The data from the discussion forum was first imported into Microsoft Excel. Each posting was put in a cell, in a row with other related information such as the time of the posting, and the ID of the student who contributed the posting. The custom parser was run on all 42 rows, and it produced a unique word-list/frequency list followed by a unique word-pair/frequency list. Excel's built in functions were then used to generate the descriptive statistics and to sort the lists in order of descending frequency.

Sentiment analysis was conducted using Microsoft Excel and a custom sentiment analyzer that took as input the discussion forum, and returned the number of positive words and negative words per posting, as well as the difference between positive and negative words. The words were taken from Liu, Hu, and Cheng's (2005) lexicon of positive and negative words and matched with the words in the transcript. The data from the discussion forum was first imported into Microsoft Excel and each posting was put in its own cell. Then the custom sentiment analyzer was run on all 42 rows.

Social presence analysis was conducted using (a) Microsoft Excel for both its built-in descriptive-statistic functions and its built-in data analysis macros and (b) a custom social presence analyzer that took as input the discussion forum, and returned the number of words or phrases that matched items in the social presence lexicon created for this study. The number of matches denoted the social presence score.

The words and phrases in the social presence lexicon were chosen from the transcript according to the initial IAM coding. We picked out the specific words that established social presence. This task was informed by the work of Whiteside and Garrett Dikkers (2012) whose five element Social Presence Model provided guidance to identify which words within previously identified IAM Phase postings were most influential. Groupings of both positive and negative words were compiled in a Microsoft Excel sheet. To conduct the social presence analysis, the data from the discussion forum was first imported into Microsoft Excel, with each posting placed in its own cell. Next the positive and negative words from the lexicon were input into the social presence analyzer. The custom social presence analyzer was then run on all 42 rows, which output a social presence score for each row corresponding to a single post. Each social presence score was then placed in one of three columns, corresponding to the IAM Phase (note: there were no Phase IV or Phase V postings). Finally, a single-factor (one-way) ANOVA was run on the three columns to determine if there was a difference in social presence.

Co-word analysis was conducted by using (a) NodeXL, a social networking analysis plug-in for Microsoft Excel (Hansen, Shneiderman, & Smith, 2010) and (b) a custom parser of unique words and unique word-pairs. This is the same parser used in the data scraping analysis, but for this examination we used the word-pairs to generate a word network. We first imported the data from the discussion forum into Microsoft Excel with the NodeXL plug-in installed. Then we ran the custom parser on all 42 postings, which output the unique words, the unique word pairs, and the frequencies of both words and pairs. We entered the words as nodes in NodeXL, with the node labels being the words themselves, and word frequency as the node size. Finally we entered the word-pairs as edges into NodeXL, and the word-pair frequency as the edge size.

SNA was conducted by using NodeXL. As stated earlier, SNA is a set of methods for analyzing the relational aspects of social structures and it is typically used to study the information exchanged between people in groups and in communities, so in this example we applied SNA to an individual as the unit of analysis to approach knowledge construction based on the idea that both individual knowledge construction and social construction of knowledge are a process of negotiation of meaning. Thus, we were able to identify individuals that are central to knowledge co-construction by analyzing and graphing the relationship between them.

In SNA, the first step to generate a discussion forum diagram is to obtain the centrality measures of each student that published a posting. Centrality measures are defined as the number of postings, in-degree, and out-degree account for student overall degree of centrality, which results from the number of interactions that each student has in a social network (Otte & Rousseau, 2002). *Betweenness* centrality is a supplemental measure of student centrality equal to the number of shortest paths from all students to all others that pass through that student. "A betweenness measure commonly reflects an individual's potential access to information as it flows through the network" (Dawson, Macfadyen, Lockyer, & Mazzochi-Jones, 2011, p. 20).

Our first step in the SNA was to calculate the centrality measures by manually inputting relevant data from the discussion forum such as the number of times someone referred to another in a post (mentions) into Microsoft Excel with the NodeXL plug-in installed. Note that typically SNA is conducted using replies to various threads to calculate centrality measures. However, the discussion forum we used contained many messages posted to a single thread so instead we opted to use mentions. For example, if person A mentions person B in a posting, there is an edge (depicted as an arrow) from A to B. Next edges were added to the spreadsheet with labels containing posting sequence number as well as the IAM Phase of the posting from the interaction analysis. NodeXL was then used to calculate the centrality measures and produce a social network diagram of the interactions.

Results and Discussion

We present the findings for each type of analysis and discuss these findings in each of the following sections. We first discuss the interaction analysis procedure and results and then discuss the procedures and results from data scraping, sentiment analysis, social presence analysis, co-word analysis, and SNA.

Findings from Interaction Analysis Using the Interaction Analysis Model (IAM)

Coding for knowledge construction regarding the definitions of culture and e-learning identified that the discourse contained a high degree of peer-to-peer interaction with students sharing many personal anecdotes concerning their individual definition of culture and e-learning. Students built off of one another's ideas as they clarified and added to points made by other students. These interactions led some students to modify their previously provided definitions. However, knowledge construction in this transcript did not exceed the IAM Phase III, the beginning phase of knowledge construction. The code counts assigned to the IAM Phases in this transcript are as follows:

I – 116

II – 35

III – 14

IV – 0

V – 0

While knowledge was constructed during the discussion, no tangible actions were taken to test the knowledge or apply it to a new context as evidenced by the lack of Phase IV and V codes. The most frequently occurring trends in this transcript were statements of agreement in Phase I and restating or advancing an argument by providing personal anecdotes or cited materials in Phase II. Given that the task assigned was definitions of culture and e-learning, Phase III was mostly limited to the personal negotiation of the meaning of terms. During the discussion, students oscillated between Phases I and II and individuals expressed their new found knowledge by modifying their personal definitions of culture and e-learning as shown in Figure 3.

The IAM was found to be appropriate to analyze computer transcripts for knowledge construction because it successfully identified the foundations for knowledge creation in the lower phases. For example, knowledge was created in the discussion forum mostly by participants sharing personal experiences with one another. Some questions were posed within Phase I postings which could have led to social construction of knowledge. However, there was not an identifiable difference between participants referring back to the questions in other students' postings and simply picking out interesting elements in other postings. This suggests that direct questions are not needed to sustain dialogue in a discussion forum with the exception of the initial question that focuses the group discussion. From a qualitative perspective, we found the IAM easy to apply to the postings. Next we discuss the results of the learning analytics methods as applied to the transcript.

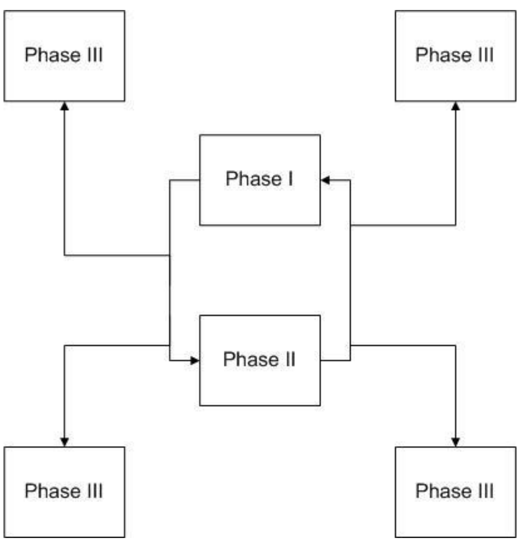


Figure 3. Social construction of knowledge based on the IAM coding.

Findings from Data Scraping

The discussion forum contained a total of 9282 words, of which 1961 were unique. The average word repetition was 4.73 (μ=4.73; σ=18.35). The 25th, 50th, and 75th percentiles were represented by word repetitions of one, one, and three, respectively, suggesting a highly-skewed left distribution, with a small number of words repeated frequently. Table 1 summarizes these statistics.

Table 1

Word Summary

Total unique words	1961.00
Total words in the forum	9282.00
Average word occurrence	4.73
Median	1.00
Standard deviation	18.35
Maximum Occurrence	401.00

Table 2 summarizes the top 25 words in terms of repetitions.

Table 2

Top-25 Word Occurrences

Rank	**Word**	**#**	**Rank**	**Word**	**#**	**Rank**	**Word**	**#**	**Rank**	**Word**	**#**	**Rank**	**Word**	**#**
1	the	401	6	A	223	11	it	104	16	As	72	21	You	65
2	and	261	7	In	187	12	for	102	17	Are	67	22	with	61
3	of	260	8	that	162	13	my	85	18	Or	66	23	have	59
4	to	248	9	Is	143	14	this	74	19	We	65	24	An	58
5	I	229	10	**culture**	107	15	**e-**	73	20	Be	65	25	**learning**	54

							learning							

Table *3* below summarizes the grammatical types of the top-25 words.

Table 3

Word Types

Determiner	the, a, an
Conjunctions	and, or
Prepositions	of, to, in, for, as, with
Pronoun	I, that, it, my, this, we, you
Verb	is, are, be, have
Adverb	As
Noun	culture, e-learning, learning

Focusing just on the top nouns, one finds that “culture”, “e-learning”, and “learning” are among the top 25 words. As mentioned, the topic of the forum was indeed “culture”. Therefore, the data suggests that it may be possible to use a simple ranking procedure (picking out the top nouns) to know the topic of any online discussion. Such a simple procedure is invaluable for informal learning on social networks, like Twitter and Facebook where there is no formal title or topic heading.

One advantage of data scraping is the quick computer generated ability to find key phrases such as “I think” or “I disagree,” that may signal key phases in the process of social construction of knowledge as described in the IAM. Also noteworthy is the frequency of prepositions, adverbs, and conjunctions. The construction of meaning and new knowledge requires blending concepts in mental spaces (Fauconnier & Turner, 2008). Prepositions and adverbs are elements of prepositional phrases and connectives that construct the mental spaces necessary for blending concepts (Fauconnier, 1994, p. 17). Determiners like “the”, “a”, and “an” introduce or specify elements in mental spaces (Fauconnier, 1994, p. 20). Their high occurrence in the data, suggest that it may be possible to find correlations between space builder words and the five IAM Phases.

Findings from Sentiment Analysis

The words in the discussion forum matched 248 words in the positive lexicon, and 107 words in the negative lexicon. The average number of positive words per posting was 5.90 (μ=5.90; σ=4.16), and the average number of negative words per posting was 2.55 (μ=2.55; σ=2.48). The 248 positive words represented 2.67% of the 9282 words; the negative words represented 1.15% of which 1961 were unique. The average word repetition was 4.73 (μ=4.73; σ=18.35). Table 4 summarizes these statistics.

Table 4

Discussion Forum Sentiment Summary

Total positive words	248.00 (2.67%)
Total negative words	107.00 (1.15%)
Mean positive words per posting	5.90 (σ=4.16)
Mean negative words per posting	2.55 (σ=2.48)

Table 5 summarizes the specific posting sentiments. The Roman numeral in parentheses, next to the posting number, is the maximum IAM Phase we coded.

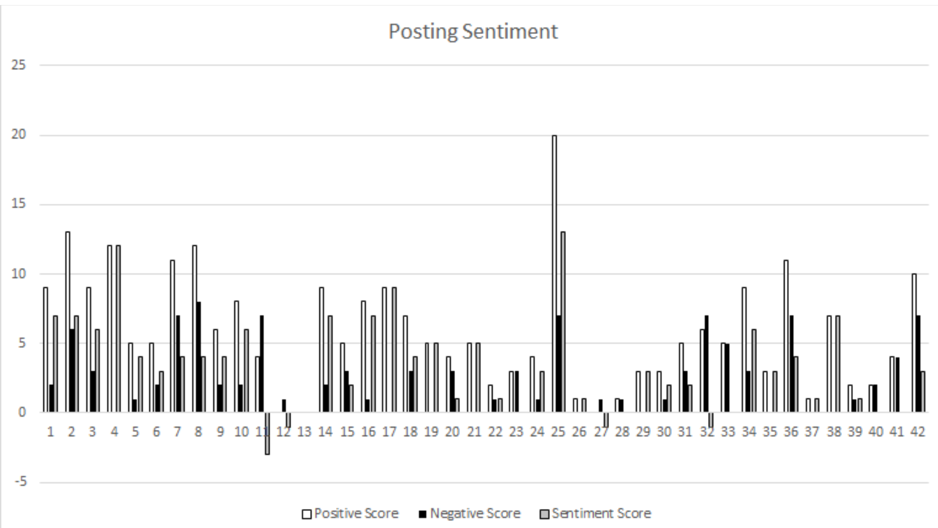
Table 5

Posting Sentiment

Posting # (max Phase)	Positive Words	Negative words	Difference	Posting # (max Phase)	Positive Words	Negative Words	Difference
1 (II)	9	2	7.00	22 (I)	2	1	1.00
2 (II)	13	6	7.00	23 (I)	3	3	-
3 (II)	9	3	6.00	24 (II)	4	1	3.00
4 (II)	12	0	12.00	25 (III)	20	7	13.00
5 (II)	5	1	4.00	26 (III)	1	0	1.00
6 (II)	5	2	3.00	27 (II)	0	1	(1.00)
7 (II)	11	7	4.00	28 (III)	1	1	-
8 (I)	12	8	4.00	29 (II)	3	0	3.00
9 (I)	6	2	4.00	30 (III)	3	1	2.00
10 (II)	8	2	6.00	31 (I)	5	3	2.00
11(I)	4	7	(3.00)	32 (II)	6	7	(1.00)
12 (I)	0	1	(1.00)	33 (II)	5	5	-
13 (I)	0	0	-	34 (II)	9	3	6.00
14 (III)	9	2	7.00	35 (III)	3	0	3.00
15 (III)	5	3	2.00	36 (II)	11	7	4.00
16 (III)	8	1	7.00	37 (I)	1	0	1.00

17 (III)	9	0	9.00	38 (II)	7	0	7.00
18 (II)	7	3	4.00	39 (II)	2	1	1.00
19 (I)	5	0	5.00	40 (II)	2	2	-
20 (I)	4 (V)	3	1.00	41 (II)	4	4	-
21 (III)	5 (V)	0	5.00	42 (II)	10	7	3.00

Figure 4 depicts the change in sentiment over time.



Note: The horizontal axis is the posting number and the vertical axis represents the scores.

Figure 4. Positive, negative, and sentiment scores for the discussion forum on culture and e-learning.

In this particular forum, early postings that were overall negative sentiment or neutral sentiment (11-13), corresponded to lower levels of knowledge co-construction (IAM Phase I). There is declining positive sentiment from postings 14-24, which suddenly peaks in positivity at posting 25. During this decline, knowledge co-construction ranged from Phase I to the early operations of Phase III. The spike in positive sentiment at posting 25, scored as a Phase III co-construction was followed by consistently high levels of knowledge co-construction to the end of the forum (Phases II and III), with only postings 31 and 37 being the exceptions (Phase I). Sentiment analysis was thus able to point out this important transition in knowledge construction. Posting 25 not only acknowledged the contributions of others to the discussion of the definitions but also moved the group to consider a variety of issues in developing a definition. Posting 25 is provided below:

1 09/02/12 11:41:00 I have read through the threads in this section, and I am very impressed with the level of thoughtfulness and the extent of teaching experience embedded in this group. I feel fortunate to be part of this class. I do not come from a similar background as most of you. Betty is perhaps the closest classmate with whom I can identify a cultural similarity. I have never been an "official" teacher other than a lab instructor for freshmen physics students. In an unofficial capacity, I have taught employees and military personnel in my charge specific skills that would allow them to advance in their careers. What follows is my definitions: Culture: To be able to define this term, it is necessary to determine its context. For a biologist, it is something to be found in a Petri dish. For a health food enthusiast, it's yogurt; and for a Brooklynite, pronouncing the word with the empathesis on the second syllable, it's what Upper West Side Manhattanites use for amusement. I suspect that for our purposes, we had better not stray too far from a socio-political or even an anthropological context. *I believe it to be an all-*

encompassing term for what takes place whenever two or more people interact. We are not to limit ourselves to the physical aspects of this interaction, but must examine the psychological, emotional and spiritual components of such interactions if we are to grasp the true meaning of the concept. I suppose that under these conditions or criteria, culture could be nothing short of human life itself. Indeed, what would human life be if not for its interpersonal interactions. For how do we build anything - whether that be families, nations or bridges - without human interaction? So culture is that quality of human existence that is defined by the interaction of its participants. e-learning: This is a loaded term. At first glance, this simple and innocent-looking combination of two words, more accurately a letter and a word, seems a necessary consequence of our modern world. On the surface, the word carries the meaning of combining leaning and electronic. But it is safe to say that in today's world, the meaning of electronic has subtly shifted to digital in this context. With this shift, the meaning has incorporated the concept of a computer and the use of it as a tool to gain access to the virtual world of information. Now we enter a speculative area whereby we can envision a brave new world where learning is enhanced by these electronic access tools. This is no longer a simple combining of meanings, but a paradigm shift in perspective. This is accomplished by the unavoidable comparison between the human brain and computers, and an even more

> profound comparison between the human nervous system and the WWW/Internet. What are we to do? This new world is ours to explore with all its euphoric epiphanies and hidden dangers. May we have the wisdom to know the difference?

In particular, the findings suggest that while a specific posting's positive or negative sentiment does not appear to be tied to a specific phase, downward and upward-sloping sentiment *trends* may correspond to specific phases. Finally, a spike in positive sentiment may signal a shift in the discussion to higher phases of knowledge co-construction. We must emphasize that the results reported are specific to the forum analyzed, and future research is needed to determine the generality of the findings.

Findings from Social Presence Analysis

The discussion forum matched 135 unique words or phrases in the social presence lexicon. The average number of social presence words or phrases per posting was 14.69 (μ=14.69; σ=8.76).

Table 6 summarizes the specific posting sentiments.

Table 6

Posting Social Presence Scores

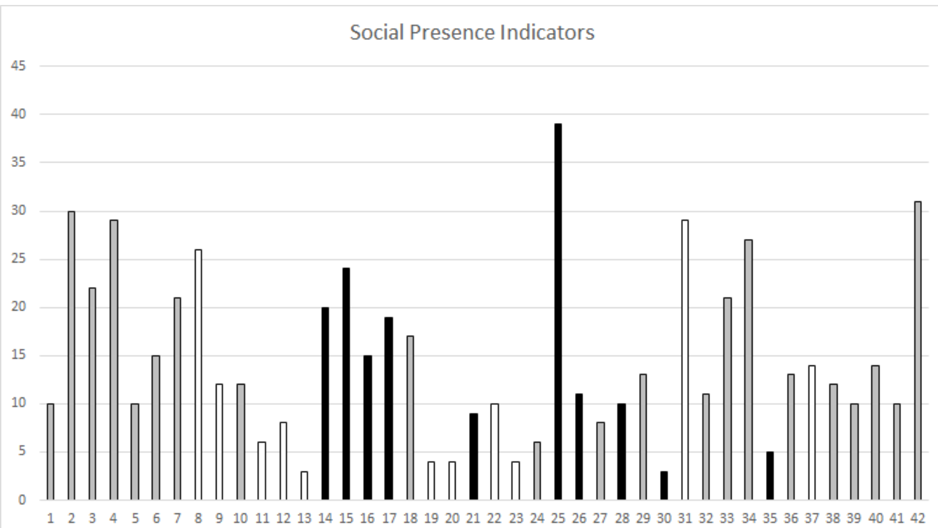
Posting #	Social presence score	Max IAM Phase	Posting #	Social presence score	Max IAM Phase
1	10	II	22	10	I
2	30	II	23	4	I
3	22	II	24	6	II

4	29	II	25	39	III
5	10	II	26	11	III
6	15	II	27	8	II
7	21	II	28	10	III
8	26	I	29	13	II
9	12	I	30	3	III
10	12	II	31	29	I
11	6	I	32	11	II
12	8	I	33	21	II
13	3	I	34	27	II
14	20	III	35	5	III
15	24	III	36	13	II
16	15	III	37	14	I
17	19	III	38	12	II
18	17	II	39	10	II
19	4	I	40	14	II
20	4	I	41	10	II
21	9	III	42	31	II

Note: The horizontal axis is the posting number and the vertical axis represents the scores.

Figure 4. Positive, negative, and sentiment scores for the discussion forum on culture and e-learning.

5 depicts the change in social presence over the 42 postings.



Note: The horizontal axis is the posting number and the vertical axis represents the social presence score. The IAM Phases I, II, and III are shaded in white, gray, and black, respectively.

Figure 5. Social presence scores for the discussion forum on culture and e-learning.

The IAM Phase I postings had an average social presence score of 10.91 ($N=11$; $\mu=10.91$; $\sigma=8.55$). The IAM Phase II postings, which were half of the discussion forum ($N=21$) had an average social presence score of 16.29 ($\mu=16.29$; $\sigma=7.51$). Finally, the IAM Phase III postings had an average social presence score of 15.50 ($N=10$; $\mu=15.50$; σ=10.08). Based just on the means and a visual inspection of Note: The horizontal axis is the posting number and the vertical axis represents the scores.

Figure 4. Positive, negative, and sentiment scores for the discussion forum on culture and e-learning.

5, it would appear that higher-levels of social presence are associated with higher phases of the IAM. However, a single-factor ANOVA revealed no statistically significant difference between group means ($F(2,39)=1.409$, $p=.25$).

Our hypothesis that higher levels of social presence in a posting are associated with higher IAM Phases was not confirmed. Although our hypothesis was not confirmed, social presence scores remain useful for IAM analysts. It simply means that there are probably other factors that contributed to the IAM Phase of a posting, of which social presence is only one. Moreover, the discussion forum we used as a dataset did not contain the full range of the IAM Phases (Phase III was the maximum).

Social presence scores are still useful because they can be associated with postings that *may* be at high or low IAM Phases. For example, 12 postings have a social presence score of 20 or greater. Of those, only two are at the IAM Phase I, the remaining 10 are at the IAM Phase II and III, the higher levels of knowledge construction. Similarly, five postings have a social presence score under five, and of those only one is in IAM Phase III; the rest are in IAM Phase I. Future research should further investigate the association between social presence and the higher levels of knowledge construction according to the IAM.

Findings from Co-Word Analysis

There were a total of 9240 word pairs of which 6840 were unique. The average word-pair repetition was 1.35 ($\mu=1.35$; $\sigma=1.37$). The 25^{th}, 50^{th}, and 75^{th} percentiles were represented by word-pair repetitions of one, one, one, respectively, denoting a highly-left skewed distribution. Table summarizes the word-pair descriptive statistics.

Table 7

Descriptive Statistics: Word-Pairs

Total unique word-pairs	6840.00
Total word-pairs	9240
Average word-pair occurrence	1.35
Standard deviation	1.37
Maximum word-pair	48

Table summarizes the top word-pairs, and their possible roles in Conceptual Blending. — a basic mental operation for creating meaning via a dynamic process that integrates different elements of cognitive structure (Fauconnier & Turner, 2002; p.37). Six of the top-10 word pairs are mental space builders, and the other four connect elements between spaces.

Table 8

Top-10 Word-Pairs: Sorted by # of Occurrences

Rank	**Word-Pair**		**#**	**Possible Role**
1	of	the	48	Space builder
2	in	the	27	Space builder
3	I	think	22	Space builder
4	is	a	20	Connector
5	in	a	20	Space builder
6	to	the	20	Space builder
7	it	is	18	Connector

8	of	culture	16	Space builder
9	culture	is	16	Connector
10	to	be	16	Connector

Network descriptive statistics include betweenness centrality, in-degree centrality, and out-degree centrality. Betweenness centrality is a measure of how many times a node (word) appears on the shortest path between two other nodes (Wasserman & Faust, 1994, p.188-192). In terms of words *qua* nodes, in-degree centrality measures how many times a given word was preceded by another word, and out degree centrality measures how many times a given word was followed by another word (Wasserman & Faust, 1994, p. 178-180). Table summarizes the top nodes sorted by betweenness centrality.

Table 9

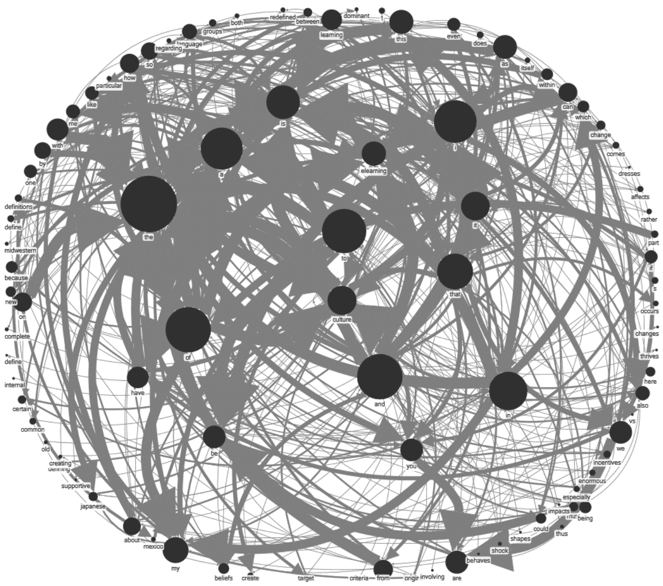
Word Centrality Measures

Node	**Betweenness-Centrality**	**In-Degree**	**Out-Degree**
And	681576.431	210	183
The	672787.996	151	242
To	511314.577	187	142
a	397443.671	95	151
In	314605.348	149	82
Of	294478.177	141	104
I	264636.190	146	91
That	230879.062	129	93
For	184954.829	88	58

Is	169056.118	61	85
It	137825.011	75	65
Or	125082.526	53	52
My	120488.837	50	51
culture	92877.021	54	50
This	83610.273	55	54
As	77715.580	51	38
With	77179.949	46	36
Are	71513.706	36	56
We	69086.292	44	47
On	68012.974	39	29
An	67696.195	33	35
You	65417.546	44	42
Not	63183.394	31	37
Be	59164.045	20	50
Its	58299.466	33	31
e-learning	56691.268	39	45
By	54266.979	29	22
Have	54096.903	30	37

Culture is the noun with the highest centrality. However, it is preceded by 13 more central words, which function as mental space builders, and as connectors between mental spaces.

Figure 6 summarizes the subnetwork of word neighbors for "culture", which was the topic of discussion. Node size represents the frequency of the word in the discussion forum, and the size of the edge between two pairs of words represents the frequency of that word pair.



Note: Big nodes in the center represent words that organize the concept's development, while smaller nodes in the periphery represent the words that reify the concept.

Figure 6. Subnetwork of immediate word neighbors for "culture".

The bigger nodes include the topic (culture) and subtopic of the forum (e-learning), as well as words that we have already identified as part of the Conceptual Blending process (e.g., prepositions and conjunctions). What is useful about this subnetwork neighbor diagram, however, is that it also depicts the smaller nodes that culture connects to. These connections include: culture→thrives (II), culture→shapes (III), culture→dresses (III), culture→behaves (III), culture→impacts (II), and culture→changes (I). The word pairs are found in postings that we classified as the IAM Phases II, III, III, III, II, I, respectively.

Words shape the formation of concepts, and in any discussion where knowledge is socially constructed there are thousands of words exchanged. For qualitative researchers who are studying social construction of knowledge, the value of a word-network analysis is that it can isolate and depict graphically (refer to Figure 5): (a) the concepts the participants are co-constructing, such as “culture” (big nodes), (b) the relationships between concepts and the words used to develop them (thick edges between other big nodes), and (c) the less frequent words that reify the concept (small nodes connected to the concept node with thin edges). For the IAM in particular, the analysis suggests that smaller nodes may be associated with postings in which the higher-phases of knowledge co-construction are occurring.

Findings from Social Network Analysis

Figure 7 shows the centrality measures table generated by NodeXL. From left to right the columns are: the anonymized IDs that correspond to the students who participated in the online discussion forum; the number of postings by that student; the in-degree centrality; the out-degree centrality; and finally, the betweenness centrality.

Node	# Postings	In-Degree	Out-Degree	Betweenness Centrality
JG	5	4	4	23.13
JL	4	4	1	3.33
MM	2	2	1	8.73
LW	4	2	2	17.33
AJ	3	1	2	1.07
EP	6	3	4	30.30
BS	3	2	4	14.97
CJ	4	3	6	44.60

TK	2	2	0	1.90
KB	3	3	2	13.40
LM	1	4	2	1.67
AN	1	2	1	5.90
DK	2	0	1	0.00
GO	2	0	2	1.67

Figure 7. NodeXL centrality measures table.

In the context of an online discussion forum, as mentioned earlier, number of postings, in-degree, and out-degree account for student overall degree of centrality, which results from the number of interactions that each student had with other students. To illustrate the meaning of centrality, we will now examine different scores from Figure 7.

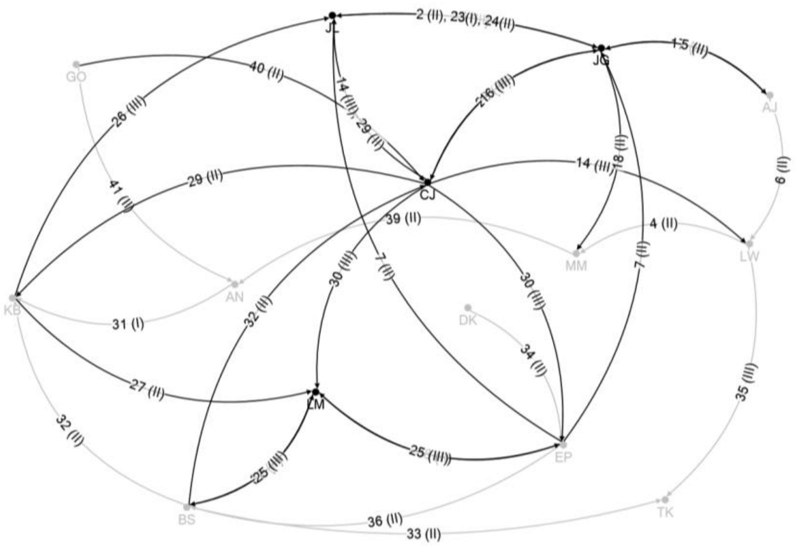
Student CJ posted four times, and had an in-degree of three, which denotes that CJ was mentioned three times by other students. The out-degree of six denotes the number of students mentioned by CJ in his or her postings. Betweenness centrality is more complex, generally, it is a measure of the number of times a node appears on the shortest path between all other nodes in the network. However, the edges in our network are not physical paths but represent mentions, i.e., a kind of conceptual path. Thus for our data, betweenness centrality is an index of a node (person's) aggregation and dissemination of concepts. A high betweenness centrality score for a specific node denotes a person's contribution to the co-construction of knowledge in the social network. Thus CJ's relatively high betweenness centrality score of 44.6, suggests CJ was central to the social construction of knowledge in the discussion forum.

Contrast CJ with AJ. AJ posted three times and had in-degree, out-degree, and betweenness centrality scores of one, two, and 1.07, respectively. Thus AJ was mentioned only once by other users; mentioned only two other users; and was a low factor in the social construction of knowledge.

In short, the number of postings a student posts in a discussion forum may be a misleading indicator of their contribution to knowledge construction, thus the need for researchers to use more sophisticated measures of student centrality that are more indicative of communication dynamics, for instance—as the colloquial expression goes—a student is "in" when he or she is mentioned by others, "out" when not mentioned by others or is the one doing the mentioning, that is, a kind of popularity or influence rating. However, most revealing is a high score of betweenness centrality, which uncovers information brokers.

Two questions emerge: (a) what do students post that makes them information brokers? and (b) how do students become influential? The answer to the first question can be found by looking at the results of the IAM and part of the answer to the second question can be found by analyzing how students position themselves in a discussion forum.

IAM analysts can "x-ray" the position of students in a discussion forum by generating social network diagrams that reveal how students are connected to one another. Figure 8 shows a diagram of interaction patterns that corresponds to Figure 7.



Note: Edges are labeled with the posting number and the maximum IAM phase in parentheses. Nodes with the highest in-degree (JL, JG, LM) and highest betweenness centrality (CJ) are highlighted in black along with associated edges.

Figure 8. Social network diagram of interaction patterns.

Figure 8 has circles known as nodes or actors that represent a student that published a posting. A line with an arrow at the end is known as a directed arc or edge and it represents a reply, i.e., one interaction or connection. This diagram is one possible graphical representation of student-student interaction patterns, which provides supplemental visual information. Let us take the same examples from Figure 7 to further explain the meaning of the social network diagram presented in Figure 8.

There is a circle labeled CJ (center of diagram) which is student CJ's initial posting and there are four directed arcs going out of it representing an out-degree of four because the student mentioned other students in the postings indicated on the diagram four times. There is one directed arc to student CJ representing an in-degree of one because CJ was mentioned by JG.

According to the student’s betweenness centrality score of 44.6, CJ is positioned in the social network in such a way that it is easy for him or her to receive and send information for others, which is a social position that facilitates processing of information—in the sense of negotiation of meaning of others' postings—a necessary element for knowledge construction.

There is a circle labeled AJ (upper-right), which is student AJ's initial posting and there is one directed arc going to this posting for an in-degree of one. There is one directed arc to student AJ due to an in-degree of one. Student AJ is not positioned to be an information broker, according to the student’s betweenness centrality score of 1.07. AJ is positioned in the periphery of the social network so it is difficult for the student to receive and send information for others, in comparison to student CJ.

An important finding that emerged from SNA was that nodes at the IAM Phase III either had high in-degree centrality or high betweeness centrality. This suggests that IAM analysts should focus on people with high centrality when conducting their analysis because their position as information brokers highlights their key role in social construction of knowledge. For example, EP has the highest centrality score and two directed arcs at Phase III pointing to the student's node. In addition, our SNA revealed that students who are most central to the network captured nine out of ten instances of Phase III postings. Note that the majority of these postings were at Phase III A – negotiation or clarification of the meaning of terms. It follows that people who demonstrate the highest levels of knowledge construction will also be the most central to the network because higher levels of knowledge cannot be achieved unless there is interaction with previous postings. However, the two students who provided postings at the highest Phase III operations (D and E) according to our interaction analysis were on the periphery of the network as indicated by their centrality scores. This suggests that knowledge can be constructed by students regardless of their centrality in the network.

The ability of students on the periphery to make such substantive contributions is enabled by those who are most central to the network because they facilitate learner engagement by disseminating critical knowledge. Bonk and Khoo (2014) define learner engagement as effort, involvement, and investment, and note that decades of research has shown that student involvement in the learning process matters when it comes to student achievement. They cite Ingram's (2005) work that suggested that engagement is made up of three variables: (a) deep attention to the learning task or situation, (b) the activation of effective cognitive processes such as strategies of rehearsal, organization, visual imagery, etc., and (c) the social context or community in which learning occurs. Learner engagement is critical to knowledge construction and therefore, those with high centrality measures are very important to the network to get students engaged in the discussion. We conclude that students with high centrality are important for (a) preparing a group for construction of new knowledge and (b) getting students engaged in the discussion.

This study has shown the advantage of utilizing multiple methods to analyze online discussions which provides insight into the social dynamics that accompany the process of knowledge construction. Other researchers (e.g., Aviv, Erlich, Ravid, & Geva, 2003; Li, 2009; and Buraphadeja, 2010) have also utilized more balanced mixed methods approaches to examine the relationship between knowledge construction and social networks. In doing so they have gained insight into the orchestration of online discussion forums, which is still the quintessential communication tool deployed by online faculty to facilitate discussion of content, promote socialization among students, and foster paths to higher levels of knowledge construction.

Limitations

Qualitative research, such as implementing the IAM, is often a time consuming undertaking that requires coders to work with the same dataset in multiple iterations to produce trustworthy findings. The primary limitation we encountered however, was that the discussion forum we analyzed did not contain examples of the higher levels of knowledge construction at Phase IV and V of the IAM. This limited our ability to assess whether the model could effectively identify higher levels of knowledge construction such as testing knowledge against given criteria or applying knowledge to new contexts. Although we demonstrated that pairing analytics with interaction analysis enriches qualitative research, the true depth of the impact cannot be identified without using a discussion forum that reaches Phase V of the IAM. Indeed our findings would have been enriched by using a discussion forum that exemplified all five IAM Phases. For example, since the discussion forum did not provide many examples of dissonance, the negative word list in the social presence lexicon was not sufficiently robust to demonstrate how negative words can both facilitate and hinder knowledge construction. A further limitation is the lexicon that was used in the sentiment analysis was taken from a marketing context. We suggest using a lexicon that comes from an online education context or creating a lexicon that is specific to the research such as the lexicon we created for the social presence analysis. The SNA indicated an association between higher levels of knowledge construction and centrality in a social network, however, further research is necessary to establish this association.

Conclusion

Interaction analysis using the IAM can be enriched by employing learning analytics and SNA to expand our understanding of the socio-emotional dynamic that accompanies the process of knowledge construction. Through combining these different analyses, we were able to perceive that learning is not purely a cognitive process, but is also emotionally loaded and situated within a social context. In addition, learning analytics helped us to identify general trends in the data. For example, our interaction analysis identified posting 25 as one of the few instances of Phase III in the discussion forum. Further clarity was provided regarding the nature of this posting via our sentiment analysis and social presence analysis. The sentiment analysis identified posting 25 as having a very positive sentiment indicated by the highest difference between positive and negative words. The social presence analysis assigned the highest social presence score to posting 25. These findings make it clear that interaction analysis can be enhanced by employing learning analytics. Therefore, we suggest that learning analytics be used by IAM analysts to inform their qualitative results. For example, data scraping, sentiment and social presence analyses, are useful techniques for: (a) highlighting areas that a qualitative researcher should focus on in her data, (b) indicating the socio-emotional context that accompanies knowledge construction, and (c) suggesting hypotheses for future research.

In addition, we illustrated how to use SNA to obtain supplemental measures of student centrality and a diagram of an online discussion forum based on the idea that centrality measures provide sound indicators of a student's ability to transfer information and exert influence over other students in a discussion. Social network diagrams offer a snapshot of the resulting interaction patterns. Therefore, the study of the relationship between social construction of knowledge and student centrality in this context helps researchers gain a better grasp of the characteristics of discussion postings and the degree of student centrality associated with potential paths to higher levels of knowledge construction. In sum, social network diagrams make the social dynamics of online learning tangible which extends the IAM analysis beyond its typical capacity of focusing on cognitive processes.

Figure 9 illustrates the process we suggest for incorporating learning analytics with IAM analysis.

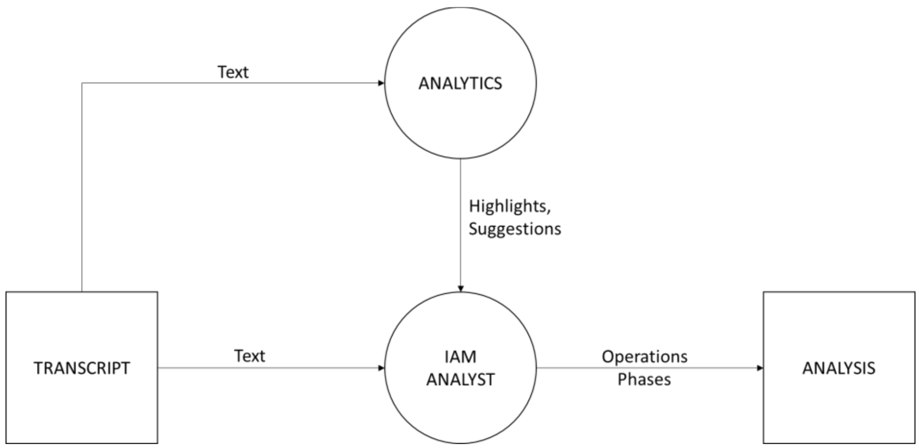


Figure 9. Analytics as a process that assists, not substitutes for, the analyst.

Using the text from a discussion transcript to conduct analytics and equip the IAM analyst with preliminary findings and general themes will enhance qualitative content analysis. Our intent is to provide a quantitative backbone for the IAM analyst because the analyst's lens will always play a key role in analysis and analytics cannot replace this vital perspective. This article demonstrated that the role of analytics is to assist the researcher in analysis by highlighting instances in the data where significant knowledge co-construction may be occurring, and by suggesting hypotheses and topics for further research. Note that the code for the custom parser, sentiment analyzer and social presence analyzer is available upon request for utilization in future research studies.

References

Bonk, C. J., & Khoo, E. (2014). *Adding some TEC-variety: 100+ activities for motivating and retaining learners online.* Bloomington, Indiana: Open World Books. Creative Commons: http://tec-variety.com/TEC-Variety_eBook_5-4.pdf

Buraphadeja, V., & Dawson, K. (2008). Content analysis in computer-mediated communication: Analyzing models for assessing critical thinking through the lens of social constructivism. *The American Journal of Distance Education, 22*(3), 130-145

Chai, C. S., & Tan, S. C. (2009). Professional development of teachers for computer-supported collaborative learning: A knowledge-building approach. *The Teachers College Record, 111*(5), 1296-1327

Dawson, S., Macfadyen, L., Lockyer, L., & Mazzochi-Jones, D. (2011). Using social network metrics to assess the effectiveness of broad-based admission practices. *Australasian Journal of Educational Technology, 27*(1), 16–27

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2007). Applying multilevel modelling to content analysis data: Methodological issues in the study of role assignment in asynchronous discussion groups. *Learning and Instruction, 17*(4), 436-447

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2009). Structuring asynchronous discussion groups: the impact of role assignment and self-assessment on students' levels of knowledge construction through social negotiation. *Journal of Computer Assisted Learning, 25*(2), 177-188

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2010). Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior, 26*(4), 516-523

Formatted: English (United States)

De Wever, B., Van Winckel, M., & Valcke, M. (2008). Discussing Patient Management Online: The Impact of Roles on Knowledge Construction for Students Interning at the Paediatric Ward. *Advances in Health Sciences Education, 13*(1), 25-42

Ding, X., Liu, B., & Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. *In Proceedings of the 2008 International Conference on Web Search and Data Mining (pp. 231-240)*. ACM

Eichmann, D. (1995). Ethical web agents. *Computer Networks and ISDN Systems, 28*, 127-136.

Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language.* Boston: Cambridge University Press

Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities.* New York: Basic Books.

Fauconnier, G., & Turner, M. (2008). *The way we think: Conceptual blending and the mind's hidden complexities.* New York: Basic Books

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*, 82-89

Firdausiah Mansur, A. B., & Yusof, N. (2013). Social learning network analysis model to identify learning patterns using ontology clustering techniques and meaningful learning. *Computers & Education, 63*, 73-86

Flor, N. V., Coulson, S., & Maglio, P. P. (2006). Schema blending and stable structure in online social systems. *International Journal of Web Based Communities, 2*, 143-159

Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications*, *1*, 147-166

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), 397-431

Gunawardena, C. N., & Zittle, F. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *American Journal of Distance Education*, *11*, 8-26

Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufmann

Haythornthwaite, C., & De Laat, M. (2010, May). *Social networks and learning networks: Using social network perspectives to understand social learning.* Paper presented at the 7th International Conference on Networked Learning, Aalborg, Denmark. Retrieved from http://celstec.org.uk/

Heo, H., Lim, K. Y., & Kim, Y. (2010). Exploratory study on the patterns of online interaction and knowledge co-construction in project-based learning. *Computers & Education, 55*(3), 1383–1392

Herlocker, J., Konstan, J., & Riedl, J. (2000). Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on computer supported cooperative work* (pp. 241-250). ACM

Hou, H.-T., Chang, K.-E., & Sung, Y.-T. (2008). Analysis of problem-solving-based online asynchronous discussion pattern. *Educational Technology & Society, 11*(1), 17–28

Hou, H.-T., Chang, K.-E., & Sung, Y.-T. (2009). Using blogs as a professional development tool for teachers: Analysis of interaction behavioral patterns. *Interactive Learning Environments, 17*(4), 325–340

Ingram, A. L. (2005). Engagement in online learning communities. In J. Bourne and J. C. Moore (Eds.), *Elements of quality online education: Engaging communities, Volume 6 in the Sloan-C Series.* Needham, MA: Sloan Consortium. Retrieved from http://kentstate.academia.edu/AlbertIngram/Papers/582901/Engagement_in_online_learning_communities

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we Twitter: understanding microblogging usage and communities. *In Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis (pp. 56-65)*. ACM

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, *4*(1), 39–103

Konstan, J. A., & Riedl, J. (2012). Deconstructing recommender systems: How Amazon and Netflix predict your preferences and prod you to purchase. *IEEE Spectrum*, *49,* 55-58

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage

Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. *In Mining text data (pp. 415-463)*. Chicago: Springer

Liu, B., Hu, M., & Cheng, J. (2005). Opinion observer: Analyzing and comparing opinions on the web. *Proceedings of the 14th International World Wide Web conference (WWW-2005)*, May 10-14, 2005, Chiba, Japan

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

Islas, J. R. L. (2004). Collaborative learning at Monterrey Tech-Virtual University. In Duffy, T. M., & Kirkley, J. R. (2004). *Learner-centered theory and practice in distance education: Cases from higher education* (pp. 297-319). Mahwah, N.J: Lawrence Erlbaum Associates

Lucas, M., & Moreira, A. (2010). Knowledge construction with social web tools. In M. D. Lytras et al. (Eds.), *1st International conference on reforming education and quality of teaching,* CCIS 73 (pp. 278–284). Springer Verlag. doi: 10.1007/978-3-642-13166-0_40 78-284

Lucas, M., Gunawardena, C., & Moreira, A. (2014). Assessing social construction of knowledge online: A critique of the interaction analysis model. *Computers in Human Behavior,* 30, 574-582

Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science, 28*(6), 441-453

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, *2*, 1-135

Paulus, T. (2007). CMC modes for learning tasks at a distance. *Journal of Computer-Mediated Communication, 12*(4)

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

Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? *Computer-Supported Collaborative Learning, 2*, 225–246

Scott, J. (2012). *Social network analysis*. Thousand Oaks, CA: Sage

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications.* London: John Wiley & Sons

Sing, C. C., & Khine, M. S. (2009). An analysis of interaction and participation patterns in online community. *Educational Technology & Society, 9*(1), 250–261

Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the conference on empirical methods in natural language processing (EMNLP)* (Vol. 1631, p. 1642)

Tan, J., Ching, S. C., & Hong, H. Y. (2008). The analysis of small group knowledge building effort among teachers using an interaction analysis model. In *16th conference on computers & education*, Taipei, Taiwan

Toikkanen, T., & Lipponen, L. (2011). The applicability of social network analysis to the study of networked learning. *Interactive Learning Environments, 19*(4), 365-379

Tu, C. H. (2001). How Chinese perceive social presence: An examination of interaction in online learning environment. *Education Media International*, *38*(1), 45-60

Whiteside, A. L., & Dikkers, A. G. (2012). Maximizing multicultural online learning experiences with the social presence model, course examples, and specific strategies. *Computer-mediated communication across cultures: International interactions in online environments,* (pp. 395-413): IGI Global

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

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

Yildiz, S. (2009). Social presence in the Web-based classroom: Implications for intercultural communication. *Journal of Studies in International Education, 13*, 46-67