

November – 2022

# SLOAN: Social Learning Optimization Analysis of Networks

David John Lemay<sup>1</sup>, Tenzin Doleck<sup>2</sup>, and Christopher G. Brinton<sup>3</sup>

<sup>1</sup>Cerence Inc, <sup>2</sup>Simon Fraser University, <sup>3</sup>Purdue University

## Abstract

Online discussion research has mainly been conducted using case methods. This article proposes a method for comparative analysis based on network metrics such as information entropy and global network efficiency as more holistic measures characterizing social learning group dynamics. We applied social learning optimization analysis of networks (SLOAN) to a data set consisting of Coursera courses from a range of disciplines. We examined the relationship of discussion forum uses and measures of network efficiency, characterized by the information flow through the network. Discussion forums vary greatly in size and in use. Courses with a greater prevalence of subject-related versus procedural talk differed significantly in seeking but not disseminating behaviors in massive open online course discussion forums. Subject-related talk was related to higher network efficiency and had higher seeking and disseminating scores overall. We discuss the value of SLOAN for social learning and argue for the experimental study of online discussion optimization using a discussion post recommendation system for maximizing social learning.

*Keywords:* social learning optimization analysis of networks, SLOAN, social cognitive theory, social learning, information theory, network analysis

## SLOAN: Social Learning Optimization Analysis of Networks

The advent of the Internet and information technology is radically reshaping social relations. No field has been unaffected, especially educational technology. New data analysis technologies and new possibilities unlocked by increasing computing power are reshaping research by introducing advanced statistical modeling techniques. Researchers can use powerful open-source information tools to create new fields and new tools of investigation. This same process repeating itself across all fields of human activity, dubbed *technological disruption* by some, shows how information technology can remake human activity. From a social-cultural perspective, information technology as a psychological tool serves knowledge transmission and production. Psychological tools have the characteristic of being endlessly combinatorial and expressive (Vygotsky, 1986; Wertsch, 1985). At the same time, information technology has a material dimension as well: the interconnectedness of information technology, and computational tools and data, can be put to work to serve particular ends that create the conditions for new activities to arise. For Marx (1839–1841/1973), automation leads to the accumulation of productive capital. Somewhat presciently, he divined that these forms of infinite labor surplus of information machines would remake social relations. Thus, information technology in both its material and psychological dimension is *sui generis*, a tool of infinite productive capacity that, in dialectical fashion, expands the sphere of human activity while automating away human labor.

Educational technologists have eagerly incorporated new technological and methodological developments from data science into their research (Wise & Cui, 2018). This is evidenced by the growth in data-focused fields of educational data mining and learning analytics but also in the growth of mobile and multimodal technologies using user data to curate learning experiences. Whereas much educational technology research has focused on developing learning tools and measuring impact, comparatively little research has explored the development of learning interfaces, or

the windows on the world through which a person views information and which cause a certain quality of learning to occur. Interfaces to learning are the cognitive artifacts, the resources for learning, that populate the learning environment and occasion learning (Duchastel, 1996, p. 207).

Although the idea of learning interfaces can be traced back to Duchastel's work with a special committee of the North American Treaty Alliance (NATO) investigating the possibility of advanced educational technology (Liao, 1996), few researchers in the intervening years have focused on educational interface design (Kloos et al., 2020), and learning interfaces have not evolved considerably in the intervening years. We may have more data, but educational technologists are still searching for ways to incorporate learning analytics into learning technologies (Wise et al., 2015). One issue is that learning technologies are still primarily focused on formal learning activities, and content and learning interfaces have thus remained primarily static, orchestrated affairs. Much educational technology is focused on tools that are helpful for teaching and learning knowledge and skills: for instance, an important aspect of the field of self-regulated learning is focused on leveraging multimodal learner data to support the development of meta-cognitive skills (Winne, 2017). However, few have explored how information technology creates new opportunities for knowledge production as well as reproduction or how information technology creates new contexts for learning. One exception has been discussion forums. The seminal work around the Knowledge Forum and

computer-supported collaborative learning (Scardamalia & Bereiter, 2014) provides a glimpse of the potential of information technology for expanding learning possibilities by harnessing network effects of groups of learners. Although much recent research has applied new analytical techniques to the study of online discourse (Rosé, 2017), technology-supported learning environments have not kept pace with new innovations in Internet technologies that exploit advanced statistical techniques and the abundance of data and computing power to create more dynamic and adaptive learning interfaces (Wise et al., 2015), for instance, interfaces that can moderate and curate user discussion threads based on their interactions. Studies of online discourse show that differential patterns of interaction lead to qualitatively different learning outcomes (Fu et al., 2016). For instance, Wise et al., (2014) have demonstrated that *speaking*, or posting, is predicted by *listening*, or attending to others' posts. Yet few studies to date have explored algorithmic methods for fostering those interaction patterns associated with better learning outcomes (Rosé et al., 2008).

Online discussion forums are a ubiquitous part of contemporary connected living. We use them to communicate, share, express, and interact. At the most fundamental level, they are about information sharing. From them, we can get crowd-funded advice, and we can learn from the wisdom of the crowd. Indeed, most turn to the Internet for answers to their questions because they know that their question has already been asked and answered many times before. Thus, we are interested in optimizing discussion forums for social learning. In network analysis terms, we argue that improving the flow of information, or the efficiency of the network, can help create better online learning communities and foster better discourse overall. *Optimization* (Boyd & Vandenberghe, 2004), in mathematical parlance, refers to solutions that minimize error given a specific value function. *Network optimization* refers to finding all the connections that maximize benefit (or, equivalently, minimize error). In the context of social learning networks, the benefit equates to maximizing learning opportunities that can be gained from online connections. In recent years, online discussion forum research has seen an explosion, and analytical methods have been a large focus of this research (Ruipérez-Valiente et al., 2020; Zhu et al., 2020). However, this research has largely been approached in a case-based fashion, with a concomitant variety of approaches. The profusion of analytical methods (Almatrafi & Johri, 2019; O'Riordan et al., 2020) and the lack of controlled experiments make it hard to draw strong conclusions based on the extant literature. The lack of a unifying theoretical framework and method are partly responsible. Information network theory can arguably bridge perspectives and help identify features of discussion forums for comparative analysis and experimental study. Whereas many researchers have applied social network analysis to the study of online discussion forums (Jan et al., 2019; Kim & Ketenci, 2019), most have focused on connections and have not considered the quality of those connections in terms of content and message (Wise et al., 2017)—in other words, the flow of information through the network. In their systematic review of social network analysis in online learning, Jan et al. (2019) found a general lack of consideration to attributional and performance variables in the extant literature. The field of network theory provides tools to analyze the flow of information through networks. Information flow is a fundamental aspect of social learning in online discussion forums. This is central to understanding how discussion features are related to better learning outcomes. We borrow the related notions of information entropy and network efficiency (Brinton et al., 2016; Latora & Marchiori, 2001) to understand the diffusion of information within a group. *Information entropy* refers to the overall structure of the information graph (Dehmer & Mowshowitz, 2011) and *network efficiency* to the degree to which information flows through the graph (Latora & Marchiori, 2001). In the present study, we examine

network efficiency across 60 Coursera courses collected over a one-year period by Rossi and Gnawali (2014). We compare courses to understand how course features might influence discussion forum efficiency. More efficient online discussion forums can promote better social learning and better social outcomes.

## Literature Review

In the following sections, we review social learning research in the context of discussion forums, and we examine the relationship between massive open online course (MOOC) discussions and learning outcomes.

### Social Learning

According to Crittenden (2005), social learning theory “explains human behavior in terms of continuous reciprocal interaction between cognitive, behavioral, and environmental influences” (p. 960). Social learning theories (Deaton, 2015) allow us to examine and understand the social factors that contextualize and influence teaching and learning. According to social learning theories (Hill et al., 2009), collective behavior is considered key in shaping learning in a social context. Indeed, among the dimensions that shape social learning is shared construction of knowledge. In fact, Reed et al. (2010) note that a critical distinguishing characteristic of social learning is that a process—to be considered social learning—ought to occur through social interactions between actors in a network. As such, the collaborative purpose is important to the purpose of the network.

Theories of social learning emphasize the importance of discourse to the performance of groups. In fact, discourse is widely recognized as a ubiquitous and important feature of teaching and learning in and out of classrooms (Gilbert & Dabbagh, 2004). The notion of the importance of discussions has had a strong impact in educational research (Wu & Hiltz, 2004). Discussion is considered an important driver for social learning development (Soter et al., 2008). As such, to successfully coordinate learning, a key objective is to foster and facilitate effective discussions among participants in a group learning setting (Hill et al., 2009; Lee & Recker, 2021), especially in online environments (Raković et al., 2020).

Since the advent of mobile computing, the rapid diffusion of educational technology has offered new opportunities for teaching, learning, and research (Castañeda & Williamson, 2021; Lemay, Doleck, & Bazalais, 2021). The need to better understand social learning is made more acute as more learning and teaching starts to take place in online learning environments (Castro & Tumibay, 2019). Researchers have emphasized ever more the need to examine educational technology’s features and mechanisms that promote effective teaching and learning (Kimmons et al., 2021).

A common theme underlying the literature on online learning environments is that discussion forums are a crucial feature of online learning environments (Lee & Recker, 2021; Tirado et al., 2012), not only because they enable and support interactions and communication (Almatrafi & Johri, 2019) but because they facilitate the sharing of ideas and knowledge (Andresen, 2009; Rovai, 2007). The literature on online learning environments informs us that the utility of online learning environments such as MOOCs is often

tied to their ability to provide conditions that enable effective discourse (Goshtasbpour et al.; Hill et al., 2009).

## **MOOCs and Discussion Forums**

Discussion forums—widely seen as important in fostering and facilitating communication and interaction among participants in learning communities (Hammond, 2005; Rovai, 2007)—are considered crucial for social learning (Almatrafi & Johri, 2019; Goshtasbpour et al., 2021; Thomas, 2002). The topic of computer-mediated discussion forums has drawn significant attention from educational technology researchers (Gay & Betts, 2020). Prior research has extensively studied discussion forums for teaching and learning (Chiu & Hew, 2018), particularly as key learning spaces in online courses and learning management systems (Marra, 2006). For the current study, we focus on a specific context of MOOCs. The argument for prioritizing MOOCs in the current study starts with an observation that “the MOOC environment has great potential for leveraging social learning on a global scale” (Loizzo & Ertmer, 2016, p. 1028). Indeed, MOOCs have attracted scholarly attention that continues to grow (Ruipérez-Valiente et al., 2020). In fact, a recent review notes that social learning is a key topic of research on MOOCs (Zhu et al., 2020). Common to most studies on MOOC discussion forums is an acknowledgment that MOOCs often do not offer individual instructor support to students (Moore et al., 2020)—that in most instances of MOOC discussion forums, they “are the only channel for support and for information exchange between peers” (Boroujeni et al., 2017, p. 128).

Early work on MOOC discussion forums relied heavily on frequency counts and other quantitative measures (Marra, 2006; O’Riordan et al., 2020). In fact, O’Riordan et al. (2020) note that research on MOOC discussion forums has been “dominated by assessments of the quantity rather than the quality of interaction” (p. 691). Yet Moore et al. (2020) note that the volume of text in MOOC discussion forums is both an opportunity and challenge for researchers.

With the availability of fine-grained data, the analyses of discussion forums have expanded due to the creative use of data and advanced analytical methods. Researchers have used different methods for understanding various aspects of discussion forums and for extracting actionable insights from forum data across a wide variety of MOOCs (Ruipérez-Valiente et al., 2020). Almatrafi and Johri (2019) conducted a review of MOOC discussion forums and extracted the following common methods for analyzing them: observation, qualitative data, statistics, data mining, visualization, and social network analysis. Furthermore, several researchers underscore the need to examine the links between discussion forums and learning outcomes (Almatrafi & Johri, 2019; Galikyan et al., 2021; Joksimović et al., 2017). Such efforts are also salient for understanding if participating (or not participating) in MOOC discussion forums helps or hinders learning and learning outcomes.

## **MOOC Discussion Forums and Learning Outcomes**

Discussion forums serve as an important means by which participants overcome communication constraints. As noted at the outset, discussions can impact teaching and learning and, in turn, learning outcomes. Importantly, discussion forums are fertile grounds for research providing data sources for

(a) measures of engagement, by tracking users' forum viewing patterns; (b) measures of mastery, understanding, or affect, generated by applying natural language processing to the raw text of forum posts; and (c) social network data by assembling graphs where various connections in the fora constitute edges. (Gardner & Brooks, 2018, p. 138)

Given the rich data available from discussion forums, a recent and growing body of work has specifically focused on the association between MOOC discussion forums and learning outcomes.

The role of discourse behavior and forum activities has been explored in a number of studies (e.g., O'Riordan et al., 2020), with some documenting an association between discourse behavior and learning (e.g., Wang et al., 2015) and others finding correlations between forum activity and course success (e.g., Santos et al., 2014). To contextualize the link between discourse on MOOCs and performance, Dowell et al. (2015) note that discourse features accounted for 5% of the variance in the performance in their analysis. Yet other research finds active participants contributing posts unrelated to the course at a higher rate (Feng et al., 2015). Almatrafi and Johri (2019), in reviewing the literature on discussion forums in MOOCs, summarize the findings apropos participation and performance by noting that "there is a correlation between participation in the forums and completion and performance" (p. 420).

Studies in this area have also focused on the cognitive dimension of learning. With a focus on the content generated by learners in a MOOC discussion forum, Galikyan et al. (2021) examined the link between learner cognitive engagement and performance, finding a negative relationship. Regarding social interactions in MOOCs, one strand of research focuses on social presence. For instance, Zou et al. (2021) find that certain social presence is linked with higher network prestige in MOOC discussion forums.

Comparative analysis has also been an active area of analysis in MOOC forum research. For example, some studies adopting a comparative perspective have documented differences across MOOCs (e.g., Jiang et al., 2014; Joksimović et al., 2016). Other research has examined differences between contributors and non-contributors to discussion forums (Wise & Cui, 2018). The authors document a higher rate of passing the course for contributors. Similarly, other studies have documented better scores for learners participating in MOOC discussions (Tseng et al., 2016). Research has also suggested that MOOC discussion forum activities can influence learning and achievement differently (e.g., Chiu & Hew, 2018). These differences highlight the importance of comparing not only between MOOCs but also between different types of learners.

Focusing on the methodological issues and challenges, prior research has raised concerns regarding operationalizing discussion forum use (Bergner et al., 2015; Tang et al., 2018; Zhu et al., 2016). In addition to the operational issues of discussion forum use, concerns have also been raised about the variability in the definition and operationalization of performance. All this speaks to the need for precise estimations of both forum use and success so that we can better understand the conditions that are necessary or adequate for optimizing the use of discussion forums for improved learning experiences and outcomes.

However, viewed more generally, the empirical research is equivocal about the association between discussion forum participation and learners' outcomes, especially with respect to which aspects of MOOC discussions are consistently effective. Rather than merely describing MOOC discussion forum use, we

suggest that it is crucial to better understand how learners' participation in MOOC discussion forums is associated with learning outcomes as a way to elaborate a theory of social network learning. We build upon prior research in MOOC discussion forums, employing optimization theory to measure and compare the efficiencies of discussion forums. In network theory terms, *efficiency* refers to how well information flows through a network. Thus, we aim to compare network features of discussion and how they are related to MOOC learning outcomes.

## Study Purpose

Our overarching purpose is to develop learning interfaces (Duchastel, 1996) that augment human ability as Engelbart (1962) and the early trailblazers of personal computing envisioned, wherein the distributed intelligence can empower learners and maximize learning outcomes through online interactions. As discussion forum activity is associated with learning, we seek to develop tools (e.g., algorithms) that empower learning in discussion forums. In the present study, we outline an approach to the study of social learning networks employing a convex optimization algorithm to calculate the learning benefit (network efficiency) of social learning in MOOC discussion forums, which we call social learning optimization analysis of networks (SLOAN). Using a comparative approach, we sought to answer the following question: How do MOOC discussion forums compare in terms of the efficiency of their social learning networks?

## Methods

### Theoretical Framework

Our research is informed by socio-constructivist and social cognitive theory. A fundamental aspect of learning resides in its social dimension—that is to say, learning is a social activity, and focusing strictly on the individual cognitive aspect ignores how learning functions in social groups. Specifically, we appeal to the dual processes of externalizing and internalizing learning (Vygotsky, 1986) that posit that internal intra-cognitive processes first arise as external inter-cognitive processes before being internalized by the individual. We also invoke the social learning behaviors of modeling—demonstrating a target skill or behavior—and vicarious learning—learning from the experiences of others (Bandura, 1986). A MOOC discussion forum helps users exchange information; however, it is also a social artifact that preserves the interactions and also supports modeling and vicarious learning from others' experiences.

### Study Design

We employed SLOAN, developed by Brinton et al. (2016, 2018). In a previous study, we presented an implementation and performed a confirmatory analysis (Doleck et al., 2021); we now report on a replication study. As we used pre-existing data from older courses, we employed a retrospective comparative study design.

## Data Source

We employed an open-source data set of 60 Coursera discussion forums collected and made available by Rossi and Gnawali (2014). The anonymized data consist of the complete data set of these courses, many having been taught many times to many sections. The sample sizes of courses ( $N$ ) range from as little as 46 to as many as 5,172. The courses cover a wide range of topics, as can be discerned by their course titles (Table 1), but skew toward finance and science, in particular, computer science and programming.

**Table 1**

*Coursera MOOC Data Set*

Course	$N$	Total threads	Total posts
Numerical Analysis (French)	46	125	843
Asset Pricing	170	681	3,158
Automata	290	472	3,269
Big Data and Education	423	604	5,126
Bioinformatics	580	1,191	10,245
Blended Learning	1,913	4,734	27,762
Synapse, Neurons, and Brains	1,261	1,181	17,081
Climate Literacy	636	1,105	19,222
Compilers	221	457	2,668
Computational Methods for Data Analysis	127	193	1,135
Cryptography	310	433	4,170
Physical Sciences (Spanish)	105	121	1,464
Data Analysis	1,247	1,979	23,165
Data Science	3,928	4,802	52,927
Design	503	618	10,206
Designing Cities	788	523	10,033
Digital Media	1,732	2,562	23,245
Data Structures and Algorithms (Chinese)	148	284	1,227
Digital Signal Processing	261	668	3,529
E-learning and Digital Cultures	480	438	7,523
Understanding Einstein	981	1,587	19,590
Finance	873	941	12,407
Networks: Friends, Money, and Bytes	68	113	627
Game Theory	355	368	5,012
Gamification	2,431	1,766	47,570

**SLOAN: Social Learning Optimization Analysis of Networks**  
**Lemay, Doleck, and Brinton**

Genomic Science	176	238	2,654
Global Warming	355	1,047	7,799
Human–Computer Interaction	557	1,041	9,416
History of Rock	524	392	9588
Humankind	2,936	3,353	69,313
Intro to Programming (French)	427	674	5,676
Intro to Java (French)	434	687	7,310
Intro to EU Law	831	808	11,510
Intro to Psychology	4,589	9,970	103,449
Intro to Statistics	1,512	1,197	17,012
Inspiring Leadership Through Emotional Intelligence	10447	8,482	65,852
Linear Programming	409	776	6,201
Mathematical Methods for Quantitative Finance	172	196	2,950
Mental Health	969	2,989	21,049
Machine Learning	4,262	5,653	64,362
Nanotechnology	504	860	9,394
Neural Networks	817	1,368	11,677
Natural Language Processing	759	1,261	12,888
Online Games	882	899	18,009
Organizational Analysis	2,623	2,579	64,831
Probabilistic Graphical Models	464	941	5,270
Bioinformatics: Introduction and Methods (Chinese)	238	342	3,651
Introduction to Computing (Chinese)	418	873	5,811
Precalculus (Spanish)	179	317	2,324
Functional Programming	1,042	1,622	12,255
Programming 1	3,042	3,590	46,666
Programming 2	842	1,196	8,690
Relationships	1,851	2,109	55,926
Scientific Writing	1,691	1,528	79,042
Social Network Analysis	545	938	9,202
Startups	5,172	6,730	76,890
Statistics 1	2,953	2,938	35,892
Useful Genetics	443	881	7,697
Video Games in Learning	1,531	9,261	35,804

Note. MOOC = Massive open online course; EU = European Union.

## Instrument

SLOAN is a convex optimization algorithm (Brinton et al., 2016) that maximizes connections (equivalently formulated as minimizing the distance) between knowledge seekers and knowledge disseminators in online learning discussion forums subject to constraints. Convex optimization is a family of algorithms for solving multivariate problems that are linear in their constraints. Convex graph optimization problems refer to points that are linearly connected within a space and do not have analytic—that is, exact—solutions. Minimums and maximums are discovered using numerical methods including gradient descent. SLOAN uses a variant of gradient descent where the points are projected on a solution region (defined by the constraints). In Figure 1, we observe that users' (u) knowledge-seeking ( $s_{u,k}$ ) and knowledge-disseminating ( $d_{u,k}$ ) behaviors on a given topic (k) can be defined as functions of users' question-asking tendency ( $q_{u,r,k}$ ). The log term represents the diminishing returns of multiple posts (p) on a given topic (k), effectively a penalty for multiple postings. An adjacency matrix representing user–user connections, that is, their posts and replies, across topics is optimized using a variation of the alternating method of multipliers employing projected gradient descent (Boyd & Vandenberghe, 2004; Brinton et al., 2016) to produce a network of optimal connections between knowledge seekers and knowledge disseminators on specific topics.

### Figure 1

*Knowledge-seeking and knowledge demonstration as functions of question-asking tendency*

$$q_{u,r,k} = \frac{\sum_{p \in P_{r,u}} \theta_{p,k} \cdot Q(p)}{\sum_{p \in P_{r,u}} \theta_{p,k}}$$

$$d_{u,k} = \sum_r (1 - q_{u,r,k}) \cdot \log \left( 1 + \sum_{p \in P_{r,u}} \theta_{p,k} \cdot |x'_p| \right)$$

$$s_{u,k} = \sum_r q_{u,r,k} \cdot \log \left( 1 + \sum_{p \in P_{r,u}} \theta_{p,k} \cdot |x'_p| \right)$$

Note. Users' (u) knowledge-seeking ( $s_{u,k}$ ) and knowledge-disseminating ( $d_{u,k}$ ) behavior on a given topic (k) can be defined as functions of users' question-asking tendency ( $q_{u,r,k}$ ).

## Procedure

Each course's discussion forum was analyzed using SLOAN. Data were prepared by sanitizing inputs using the Beautiful Soup Python package to remove markup language and segmented and lemmatized using the Stanford Natural Language Toolkit (nltk) as standard practice in natural language processing to facilitate

the application of algorithms that rely on frequentist distributions such as topic induction. Topics were inferred using Latent Dirichlet allocation (Blei, Ng, & Jordan, 2003). Seeking and disseminating behaviors were calculated based on weighted averages of posting frequencies by topic. We report on three measures of network efficiency for reliability and validity: the benefit calculated through SLOAN, the ratio of observed and optimized eigenvalues as a measure of overall network connectedness, and the global efficiency (Latora & Marchiori, 2001), which is a related measure of information flow through the network. Both SLOAN and global efficiency algorithms can be considered variants of shortest path algorithms, which are used extensively in network analysis (Dehmer & Mowshowitz, 2011).

Subsequently we compared MOOCs based on the dominating talk within the discussion forum, whether it was subject-related or procedural talk, using simple *t*-tests. Forums dominated by subject-related talk and those dominated by procedural talk were grouped based on the inferred topics. A course with a higher prevalence of keywords such as *assignment*, *quiz*, and *homework* was classified as dominated by procedural talk. A course with a higher prevalence of subject-related keywords such as "state, european, member, union" for a course on European Union Law was classified as dominated by subject-related talk.

## Results

Seeking and disseminating scores and observed and optimized social learning network efficiency (Benefit) are presented for the following: courses with highest seeking tendency (Table 2), courses with highest disseminating tendency (Table 3), and courses with highest network efficiency (Table 4).

As can be observed in Figures 2–5, observed and optimized social learning network efficiency (Benefit) appears negatively related to seeking and disseminating tendency. Interestingly, we noticed an increased disseminating tendency in humanities-oriented courses and increased seeking in science-oriented courses. Notice how optimized benefits (on the y-axis) are a greater order of magnitude compared with observed benefits for both seeking and disseminating tendencies across all courses. Smaller courses are grouped among the courses with highest seeking and disseminating scores. However, larger courses, with thousands of students, also display high disseminating tendencies, suggesting a network effect where once a course reaches a certain size, disseminating behavior becomes more generalized and knowledge more accessible in a positive feedback loop.

Seeking and disseminating scores differ by orders of magnitude. Interestingly, seeking behavior is much less prevalent than disseminating behavior in MOOC discussion forums across all groups. The lower seeking scores can be understood by the fact that *seeking* here is defined in terms of question posting and ignores all the other ways individuals may seek information before resorting to asking a question in the discussion forum (e.g., asking a classmate or the instructor directly, searching in the course materials or using a search engine, etc.).

**Table 2**

*Courses with Highest Seeking Tendency*

Course	<i>N</i>	Seeking	Disseminating	Observed benefit	Optimized benefit
Networks: Friends, Money, and Bytes	68	1.16E-04	6.7675	9.5269	246.2457
Numerical Analysis	46	1.08E-04	2.9393	23.9728	17.3465
Computational Methods for Data Analysis	127	1.86E-05	0.8897	32.0250	26.2753
Physical Sciences	105	1.37E-05	1.7282	26.9862	364.8678
Digital Signal Processing	261	1.35E-05	0.9365	61.8857	908.1452
Precalculus	179	1.25E-05	1.2726	27.1547	197.1336
Asset Pricing	170	1.24E-05	0.7224	60.0251	49.7035
Compilers	221	1.09E-05	0.9093	76.1573	1,241.8557
Experimental Genome Science	176	9.40E-06	1.3903	24.4761	548.3799

**Table 3**

*Courses with Highest Disseminating Tendency*

Course	<i>N</i>	Seeking	Disseminating	Observed benefit	Optimized benefit
Data Structures and Algorithms	148	2.43E-06	29.2140	1.4256	2.4562
Introduction to Computing	418	3.01E-07	11.4953	4.7747	50.8587
Writing in Sciences	1691	1.10E-07	10.9102	12.9485	302.9709
Introduction to Psychology	4589	2.25E-08	9.3390	134.4230	6,656.4653
Networks: Friends, Money, and Bytes	68	1.16E-04	6.7675	9.5269	246.2457
History of Rock, Part One	524	3.66E-06	3.1643	28.5779	374.6067
Numerical Analysis	46	1.08E-	2.9393	23.9728	17.3465

		04			
Gamification	2431	8.28E-08	2.8050	32.8729	1,534.2317
Bioinformatics: Introduction and Methods	238	4.81E-06	2.3905	17.0457	171.7716
The Law of the European Union: An Introduction	831	6.60E-07	2.0681	64.8254	3,808.6908

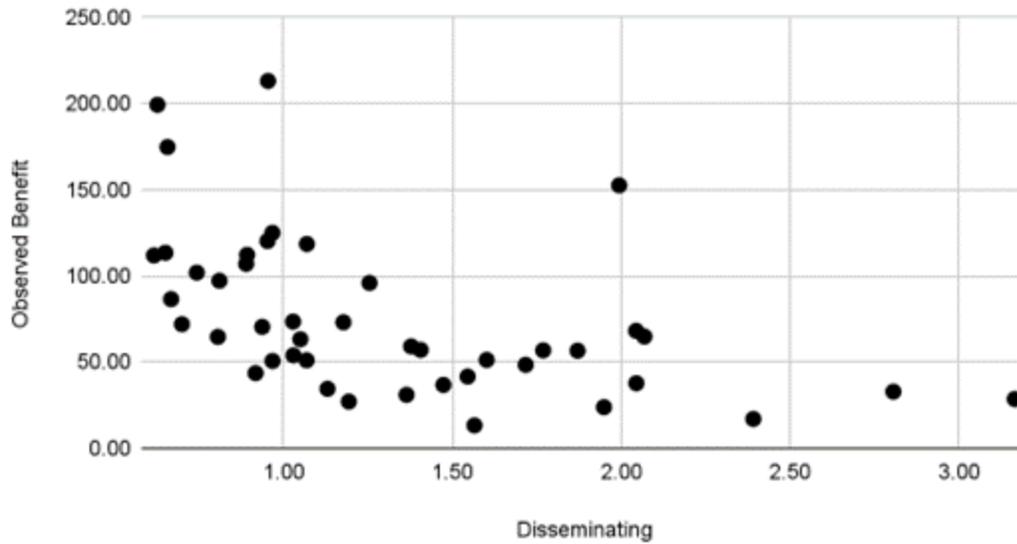
**Table 4**

*Courses with Highest Network Efficiency*

Course	N	Seeking	Disseminating	Observed benefit	Optimized benefit
Startup Engineering	5172	1.64E-08	2.04	68.15	3,250.63
Machine Learning	4262	2.43E-08	1.41	57.09	1,901.45
Introduction to Data Science	3928	2.72E-08	1.38	59.11	2,520.50
Learn to Program: The Fundamentals	3042	3.59E-08	1.18	73.08	1,920.63
Statistics One	2953	4.76E-08	0.94	70.44	2,251.49
A Brief History of Humankind	2936	4.30E-08	0.63	199.30	3,329.09
Organizational Analysis	2623	8.76E-08	1.77	56.82	1,090.97
Gamification	2431	8.28E-08	2.81	32.87	1,534.23
Blended Learning: Personalizing Education for Students	1913	1.78E-07	1.99	152.56	3,435.24
The Fiction of Relationship	1851	8.55E-08	0.81	97.20	613.89

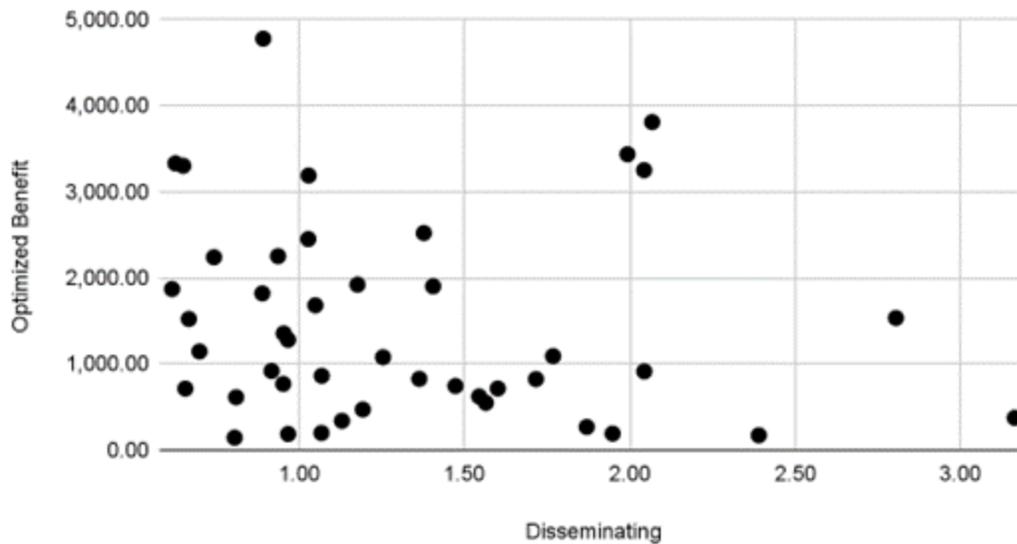
**Figure 2**

*Network Efficiency by Observed Disseminating Tendency*



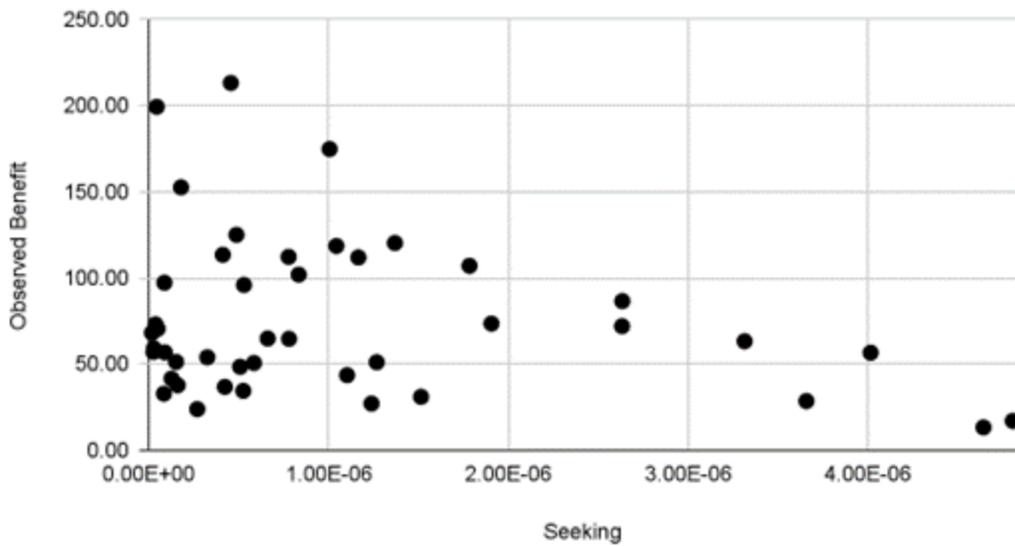
**Figure 3**

*Network Efficiency by Optimized Disseminating Tendency*



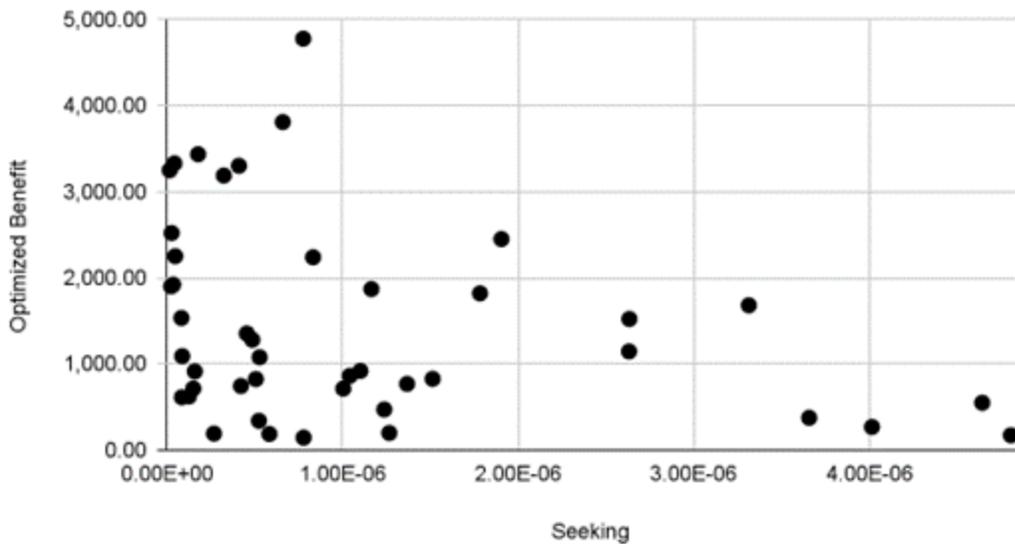
**Figure 4**

*Network Efficiency by Observed Seeking Tendency*



**Figure 5**

*Network Efficiency by Optimized Seeking Tendency*



Although the correlation coefficients revealed moderate effects, the statistics were not significant, as reported in Table 5.

**Table 5**

*Pearson Correlation Coefficient for Seeking and Disseminating by Network Efficiency*

Ratios	<i>r</i>	<i>t</i>	<i>p</i>
Seeking/observed	-0.2567	-2.0927	.1715
Seeking/optimized	-0.2785	-2.2995	.1482
Disseminating/observed	-0.2898	-2.4091	.1376
Disseminating/optimized	-0.2873	-2.3845	.1399

When comparing topics across courses, we found that courses made two broad uses of discussion forums: course-related and topic-related discussions. Some, like Blended Learning (Table 6) and Intro to European Union Law (Table 7), present topics that are more related to the subject matter. Others present topics that are more related to coursework. In Computational Methods (Table 8), we note an increased prevalence of course-directed words such as assignment, lecture, Coursera, and date.

**Table 6**

*Subject-Related Talk: Blended Learning*

Topic no.	Top word associations
0	Words: 0.029*“assign” + 0.020*“week” + 0.020*“video” + 0.016*“post” + 0.015*“link” + 0.015*“cours” + 0.014*“work” + 0.013*“final” + 0.013*“submit” + 0.013*“grade”
1	Words: 0.036*“thank” + 0.033*“cours” + 0.017*“think” + 0.015*“learn” + 0.012*“great” + 0.011*“share” + 0.010*“teach” + 0.009*“student” + 0.009*“good” + 0.009*“interest”
2	Words: 0.019*“student” + 0.018*“agre” + 0.014*“good” + 0.012*“learn” + 0.011*“definit” + 0.010*“teacher” + 0.009*“like” + 0.008*“peopl” + 0.008*“chang” + 0.008*“work”
3	Words: 0.032*“student” + 0.029*“teacher” + 0.022*“learn” + 0.018*“time” + 0.018*“think” + 0.017*“school” + 0.012*“work” + 0.010*“teach” + 0.007*“want” + 0.006*“like”
4	Words: 0.047*“learn” + 0.030*“student” + 0.019*“teacher” + 0.017*“blend” + 0.015*“think” + 0.015*“teach” + 0.013*“onlin” + 0.013*“time” + 0.011*“cours” + 0.009*“school”
5	Words: 0.041*“color” + 0.041*“size” + 0.040*“font” + 0.025*“think” + 0.019*“definit” + 0.016*“learn” + 0.014*“student” + 0.011*“khan” + 0.010*“like” + 0.009*“academi”
6	Words: 0.102*“learn” + 0.053*“student” + 0.038*“blend” + 0.016*“teacher” + 0.013*“high” + 0.013*“technolog” + 0.011*“qualiti” + 0.010*“person” + 0.010*“pace” + 0.009*“educ”
7	Words: 0.028*“learn” + 0.024*“student” + 0.020*“class” + 0.015*“technolog” + 0.013*“blend” + 0.012*“work” + 0.011*“like” + 0.011*“teach” + 0.010*“teacher” + 0.007*“school”
8	Words: 0.066*“student” + 0.018*“work” + 0.014*“learn” + 0.013*“teacher” + 0.012*“need” + 0.011*“group” + 0.010*“class” + 0.008*“classroom” + 0.008*“think” + 0.008*“like”

9 Words: 0.035\*“learn” + 0.032\*“school” + 0.027\*“student” + 0.024\*“teacher” +  
0.018\*“blend” + 0.015\*“educ” + 0.011\*“year” + 0.010\*“work” + 0.009\*“classroom” +  
0.008\*“teach”

**Table 7**

*Subject-Related Talk: Intro to European Union Law—001*

Topic no.	Top word associations
0	Words: 0.010*“cours” + 0.007*“state” + 0.006*“time” + 0.006*“peer” + 0.006*“think” + 0.005*“question” + 0.005*“like” + 0.005*“review” + 0.005*“thank” + 0.004*“great”
1	Words: 0.023*“state” + 0.018*“european” + 0.016*“member” + 0.011*“union” + 0.010*“countri” + 0.008*“europ” + 0.007*“market” + 0.007*“council” + 0.007*“econom” + 0.006*“treati”
2	Words: 0.018*“languag” + 0.011*“question” + 0.009*“think” + 0.008*“answer” + 0.008*“know” + 0.008*“countri” + 0.007*“european” + 0.007*“citizen” + 0.006*“peopl” + 0.006*“right”
3	Words: 0.018*“cours” + 0.011*“question” + 0.011*“direct” + 0.010*“case” + 0.010*“answer” + 0.009*“thank” + 0.009*“right” + 0.007*“point” + 0.007*“like” + 0.006*“good”
4	Words: 0.019*“state” + 0.015*“countri” + 0.013*“member” + 0.013*“right” + 0.011*“nation” + 0.011*“citizen” + 0.010*“case” + 0.009*“direct” + 0.008*“articl” + 0.008*“european”
5	Words: 0.013*“cours” + 0.010*“think” + 0.010*“learn” + 0.008*“peopl” + 0.008*“interest” + 0.007*“countri” + 0.007*“good” + 0.006*“time” + 0.006*“student” + 0.005*“like”
6	Words: 0.008*“peopl” + 0.008*“cours” + 0.008*“direct” + 0.007*“mean” + 0.007*“european” + 0.007*“think” + 0.007*“state” + 0.007*“ukrain” + 0.006*“work” + 0.006*“learn”
7	Words: 0.028*“cours” + 0.016*“week” + 0.013*“student” + 0.013*“read” + 0.011*“thank” + 0.011*“time” + 0.011*“video” + 0.009*“test” + 0.009*“think” + 0.009*“quiz”
8	Words: 0.014*“cours” + 0.012*“work” + 0.011*“countri” + 0.010*“live” + 0.009*“peopl” +

0.009\*“think” + 0.008\*“differ” + 0.008\*“good” + 0.008\*“hello” + 0.007\*“time”

Words: 0.027\*“thank” + 0.009\*“time” + 0.007\*“book” + 0.007\*“cours” + 0.006\*“chegg” +

9 0.006\*“problem” + 0.006\*“answer” + 0.005\*“final” + 0.005\*“download” + 0.005\*“like”

**Table 8**

*Procedural Talk: Computational Methods*

Topic no.	Top word associations
0	Words: 0.022*“cours” + 0.017*“frequenc” + 0.014*“time” + 0.009*“transform” + 0.009*“signal” + 0.009*“data” + 0.009*“function” + 0.009*“lectur” + 0.009*“think” + 0.008*“right”
1	Words: 0.027*“thank” + 0.021*“time” + 0.019*“frequenc” + 0.018*“plot” + 0.016*“domain” + 0.015*“point” + 0.011*“work” + 0.011*“grid” + 0.010*“wave” + 0.008*“number”
2	Words: 0.017*“matlab” + 0.014*“code” + 0.013*“imag” + 0.010*“filter” + 0.010*“slice” + 0.010*“problem” + 0.010*“fftshift” + 0.009*“subplot” + 0.008*“work” + 0.008*“meshgrid”
3	Words: 0.023*“time” + 0.014*“frequenc” + 0.014*“window” + 0.012*“cours” + 0.012*“domain” + 0.011*“answer” + 0.010*“signal” + 0.009*“function” + 0.008*“filter” + 0.006*“number”
4	Words: 0.018*“class” + 0.017*“plot” + 0.015*“forum” + 0.015*“thread” + 0.011*“coursera” + 0.011*“question” + 0.010*“compmethod” + 0.009*“thread_id” + 0.008*“like” + 0.008*“https”
5	Words: 0.019*“matlab” + 0.017*“thank” + 0.009*“valu” + 0.009*“know” + 0.009*“filter” + 0.008*“cours” + 0.008*“look” + 0.007*“right” + 0.007*“differ” + 0.007*“lectur”
6	Words: 0.024*“array” + 0.014*“plot” + 0.013*“octav” + 0.010*“like” + 0.009*“frac” + 0.008*“problem” + 0.007*“want” + 0.007*“line” + 0.007*“http” + 0.007*“function”
7	Words: 0.024*“frequenc” + 0.014*“nois” + 0.014*“imag” + 0.013*“signal” + 0.009*“time” +

0.009\*“https” + 0.009\*“filter” + 0.008\*“matlab” + 0.008\*“lectur” + 0.008\*“like”

Words: 0.018\*“matlab” + 0.014\*“valu” + 0.013\*“frequenc” + 0.013\*“time” + 0.011\*“signal”

8 + 0.010\*“like” + 0.010\*“function” + 0.009\*“work” + 0.009\*“domain” + 0.008\*“fftshift”

Words: 0.020\*“plot” + 0.018\*“complex” + 0.016\*“valu” + 0.014\*“right” + 0.013\*“leav” +

9 0.011\*“real” + 0.011\*“time” + 0.010\*“nois” + 0.010\*“domain” + 0.009\*“signal”

Overall, SLOAN managed to find important efficiencies in the networks, over 90% for all but the two courses with the smallest population (< 200 students), in which global efficiency of the optimized network did not surpass 60%, and Introduction to Computing, with a score of 83%, which was an outlier with the maximum disseminating score of 11.50.

While SLOAN is very effective, often converging after a single iteration, it is subject to error when optimizing learning networks where the topics are not varied enough. We note five networks where the optimized solution was simply a fully connected network. In Table 9, we observe one such network solution for the Compilers course, where the topics are barely differentiated. Note the repeated occurrence of the same words across topics.

**Table 9**

*Failed Optimization: Undifferentiated Topics, Compilers Course*

Topic no.	Word associations
0	Words: 0.014*“state” + 0.013*“cours” + 0.013*“charact” + 0.012*“compil” + 0.011*“thank” + 0.009*“string” + 0.009*“work” + 0.009*“assign” + 0.008*“error” + 0.008*“understand”
1	Words: 0.023*“class” + 0.012*“return” + 0.012*“string” + 0.012*“cool” + 0.010*“case” + 0.010*“rule” + 0.010*“type” + 0.009*“express” + 0.009*“state” + 0.009*“symbol”
2	Words: 0.014*“work” + 0.012*“string” + 0.011*“program” + 0.010*“cach” + 0.009*“compil” + 0.009*“languag” + 0.008*“need” + 0.007*“good” + 0.007*“thank” + 0.006*“know”
3	Words: 0.020*“cours” + 0.016*“error” + 0.010*“compil” + 0.010*“pars” + 0.009*“class” + 0.009*“lectur” + 0.009*“assign” + 0.008*“test” + 0.008*“program” + 0.008*“wiki”
4	Words: 0.014*“error” + 0.012*“cours” + 0.011*“java” + 0.011*“class” + 0.010*“compil” + 0.010*“program” + 0.009*“like” + 0.009*“think” + 0.008*“rule” + 0.008*“start”
5	Words: 0.026*“compil” + 0.019*“cool” + 0.018*“class” + 0.018*“assign” + 0.012*“file” + 0.011*“program” + 0.011*“code” + 0.009*“lexer” + 0.008*“output” + 0.008*“line”
6	Words: 0.017*“state” + 0.014*“step” + 0.011*“class” + 0.011*“input” + 0.010*“thank” + 0.009*“think” + 0.009*“epsilon” + 0.008*“work” + 0.007*“script” + 0.007*“assign”
7	Words: 0.019*“class” + 0.015*“work” + 0.013*“expr” + 0.011*“code” + 0.010*“express”

+ 0.009\*“look” + 0.009\*“languag” + 0.008\*“rule” + 0.008\*“like” + 0.008\*“method”

8 Words: 0.027\*“thank” + 0.024\*“string” + 0.014\*“rule” + 0.011\*“error” +  
0.009\*“charact” + 0.009\*“problem” + 0.008\*“like” + 0.008\*“return” + 0.008\*“match”  
+ 0.007\*“token”

9 Words: 0.016\*“express” + 0.010\*“cool” + 0.010\*“time” + 0.009\*“string” +  
0.009\*“assign” + 0.008\*“exampl” + 0.008\*“line” + 0.008\*“thank” + 0.007\*“token” +  
0.007\*“like”

Finally, we compared courses where talk was dominated by procedural matters—assignment questions, lecture notes, quizzes—with courses where the talk was dominated by subject-related matters—that is, centered on course topics (Table 10). After removing outliers and non-English courses, we calculated the *t*-statistic for both groups and found they were significantly different on seeking tendency though not on disseminating tendency, as well as significantly different on observed and optimized network efficiency.

**Table 10**

*Differences Between Procedural and Subject-Related Talk*

	<i>t</i> -statistic	<i>P</i> -value
Observed network efficiency		
Procedural talk	0.00E+00	0.022094755160
Subject-related talk	0.00E+00	
Optimized network efficiency		
Procedural talk	0.00E+00	0.044738732610
Subject-related talk	0.00E+00	
Seeking		
Procedural talk	1.00E+00	0.001951708988
Subject-related talk	1.00E+00	
Disseminating		
Procedural talk	2.92E-01	0.552466099200
Subject-related talk	3.39E-01	

## Discussion

We conducted a social learning optimization analysis of networks (SLOAN) in MOOC discussion forums in 60 Coursera MOOCs. We found that MOOC forums differ in terms of the optimization potential. Courses with smaller enrollment numbers appear to struggle more with generating enough discussion and connecting knowledge seekers with knowledge disseminators. However, we find that networks of all sizes can be optimized to improve their overall efficiency. Most interesting of all, we find that subject-related-talk-dominated discussion forums and procedural-talk-dominated discussion forums are significantly different in overall network efficiency and seeking tendency, but not in disseminating tendency.

It is striking how seeking tendency is orders of magnitude lower than disseminating tendency. This is possibly due to the nature of the discourse in the discussion forums, where learners are sharing information but not necessarily engaging extensively and collectively in discussion but rather are engaging in smaller sporadic exchanges. It may also be because knowledge seeking and knowledge disseminating are functions of question-asking tendency; this ignores the many other ways that knowledge-seeking behavior can manifest itself in a forum, such as composing a search query or simply reading through posts. Indeed, researchers have demonstrated that a range of online discussion behaviors are associated with differential learning outcomes (Almatrafi & Johri, 2019; O’Riordan et al., 2020; Wise et al., 2014). As one of the few studies of a larger sample of MOOC forums (Rossi & Gnawali, 2014), we find that these differences can be explained in part by the different uses made of discussion forums for MOOC learning. Wise and colleagues (Wise, 2018; Wise et al., 2017) have compared course-related and non-course-related discussion in MOOC discussion forums. They did not find significant differences in learning outcomes based on their analysis; however, they admitted this could be due to their analytical scheme. In contrast, we found significant differences in disseminating behavior between courses with a greater focus on subject versus procedural talk in the MOOC discussion forums.

Social network analysis has been applied extensively to the study of online learning (Jan et al., 2019). However, much of this research has not considered the quality of connections in social graph, such as, assessing the informativeness of the connections in a social learning network; although interaction patterns are quite variable and are not all equally conducive to learning (O’Riordan et al., 2020; Wise et al., 2017). Our study shows that content and connections can be reconciled, and these complex, multivariate interactions can be modeled and optimized using advanced numerical techniques. Integrating social learning analysis with content analysis in such a fashion provides quantitative tools for studying and augmenting social learning in technology-supported environments in the wild.

## Implications

In the present study, we have argued for conducting educational technology research in the learning interface (Duchastel, 1996), that is, developing algorithms and technologies that augment human intellectual activity (Engelbart, 1962) through responsive and adaptive tools that support and extend our abilities. We have presented a novel approach to conducting educational technology research using optimization theory to maximize learning through online discussion forums. We believe that optimization research holds much promise for the field of educational technology and distance/online learning research. It aligns with the design experimentation ethos (Brown, 1992; Cobb et al., 2003) as it is also concerned with improving teaching and learning through iterative innovation. It is suited to constructivist-based research

(Lemay & Doleck, 2020) with its sociocultural focus on tools as social artifacts and its central place in human social activity (Gee & Green, 1998). Optimization theory is being applied to improve tools and methods in data-intensive disciplines such as neuroscience, where it is helping to improve neural interfaces (Fathima & Kore, 2021). As we have shown, optimization theory can also help improve the learning interface. Online and distance learning have well-established research traditions and well-developed instructional theories such as communities of inquiry (Garrison & Akyol, 2012) and computer-supported collaborative learning (Koshmann, 2011). Interface tools like SLOAN can help improve knowledge sharing and discovery in online learning platforms.

The current study also has implications for discussions in other settings and contexts beyond MOOCs. Such an optimization algorithm can be applied to any information tool that incorporates knowledge-sharing mechanisms, such as Slack or Microsoft Teams in the workplace, to help support knowledge discovery and synergy.

### **Limitations**

The present study is limited by its use of a convenience sample; however, we believe that this is mitigated by its range and its relative size, considering most studies consider few forums in their investigations. This study is also limited by its quasi-experimental nature. We did not conduct an experiment and cannot conclude any causal relationships; however, we believe this exploratory research shows SLOAN's potential for interventionist study and particularly the potential benefits of SLOAN for social learning in online discussion forums in general and in MOOCs in particular. Given the unequal distribution of courses in our convenience sample, we did not make disciplinary comparisons. However, it is possible that disciplinary features led to systematic differences in discussion forum uses.

## **Conclusions and Future Directions**

Online social learning is quickly becoming an indispensable part of learning at every stage and in every environment, from schools to workplaces and into our daily life. The Internet is a great repository of knowledge, but it is primarily a tool for leveraging our collective wisdom. This has been facilitated by specialized knowledge-sharing forums for particular user communities, such as StackExchange and StackOverflow, and by more general knowledge-sharing platforms, such as Quora for questions and answers and Reddit for interest groups. However, the creation of filter bubbles (Turkle, 2010) has shown the dangers of ill-moderated online discussion. We believe that the best way to counter misinformation and intolerance is through discussion for knowledge and perspective sharing. Tools that can optimize social learning online can have multiplicative effects: as people learn to harness the power of online social learning networks, their agency is increased and potentials are extended (Bandura, 2001). Given the polarized landscape and the poor level of much online discourse, fostering better online discourse appears to be a moral imperative.

## References

- Almatrafi, O., & Johri, A. (2019). Systematic review of discussion forums in massive open online courses (MOOCs). *IEEE Transactions on Learning Technologies*, *12*(3), 413–428.  
<https://doi.org/10.1109/tlt.2018.2859304>
- Andresen, M. A. (2009). Asynchronous discussion forums: Success factors, outcomes, assessments, and limitations. *Educational Technology & Society*, *12*(1), 249–257.  
<https://www.researchgate.net/publication/220374740> [Asynchronous Discussion Forums Success Factors Outcomes Assessments and Limitations](https://www.researchgate.net/publication/220374740)
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bergner, Y., Kerr, D., & Pritchard, D. E. (2015, June 26–29). Methodological challenges in the analysis of MOOC data for exploring the relationship between discussion forum views and learning outcomes. In *EDM '15: Proceedings of the 8th International Conference on Educational Data Mining* (pp. 234–241). International Educational Data Mining Society.  
<https://files.eric.ed.gov/fulltext/ED560510.pdf>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022. <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Boroujeni, M. S., Hecking, T., Hoppe, H. U., & Dillenbourg, P. (2017, March). Dynamics of MOOC discussion forums. In *LAK '17: Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 128–137). Association for Computing Machinery.  
<https://doi.org/10.1145/3027385.3027391>
- Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511804441>
- Brinton, C. G., Buccapatnam, S., Wong, F. M. F., Chiang, M., & Poor, H. V. (2016, April 10–14). *Social learning networks: Efficiency optimization for MOOC forums*. IEEE INFOCOM 2016—The 35th Annual IEEE International Conference on Computer Communications, San Francisco, CA.  
<https://doi.org/10.1109/INFOCOM.2016.7524579>
- Brinton, C. G., Buccapatnam, S., Zheng, L., Cao, D., Lan, A. S., Wong, F. M. F., Ha, S., Chiang, M., & Poor, H. V. (2018). On the efficiency of online social learning networks. *IEEE/ACM Transactions on Networking*, *26*(5), 2076–2089. <https://doi.org/10.1109/TNET.2018.2859325>
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, *2*(2), 141–178.  
[https://doi.org/10.1207/s15327809jls0202\\_2](https://doi.org/10.1207/s15327809jls0202_2)

- Castañeda, L., & Williamson, B. (2021). Assembling new toolboxes of methods and theories for innovative critical research on educational technology. *Journal of New Approaches in Educational Research*, 9(2), 1–14. <https://doi.org/10.7821/naer.2021.1.703>
- Castro, M., & Tumibay, G. (2019). A literature review: Efficacy of online learning courses for higher education institution using meta-analysis. *Education and Information Technologies*, 26, 1367–1385. <https://doi.org/10.1007/s10639-019-10027-z>
- Chiu, T. K., & Hew, T. K. (2018). Factors influencing peer learning and performance in MOOC asynchronous online discussion forum. *Australasian Journal of Educational Technology*, 34(4), 16–28. <https://doi.org/10.14742/ajet.3240>
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9–13. <https://doi.org/10.3102/0013189X032001009>
- Crittenden, W. (2005). A social learning theory of cross-functional case education. *Journal of Business Research*, 58(7), 960–966. <https://doi.org/10.1016/j.jbusres.2003.12.005>
- Deaton, S. (2015). Social learning theory in the age of social media: Implications for educational practitioners. *I-Manager's Journal of Educational Technology*, 12(1), 1–6. <https://doi.org/10.26634/jet.12.1.3430>
- Dehmer, M. and Mowshowitz, A. (2011). Generalized graph entropies. *Complexity*, 17(2) 45–50.
- Doleck, T., Lemay, D. J., & Brinton, C. G. (2021). Evaluating the efficiency of social learning networks: Perspectives for harnessing learning analytics to improve discussions. *Computers & Education*, 164, Article 104124. <https://doi.org/10.1016/j.compedu.2021.104124>
- Dowell, N. M., Skrypnik, O., Joksimovic, S., Graesser, A. C., Dawson, S., Gasevic, D., Hennis, T. A., de Vries, P., & Kovanovic, V. (2015, June 26–29). Modeling learners' social centrality and performance through language and discourse. In *EDM '15: Proceedings of the 8th International Conference on Educational Data Mining* (pp. 250–257). International Educational Data Mining Society. <https://files.eric.ed.gov/fulltext/ED560532.pdf>
- Duchastel, P. (1996). Learning interfaces. In T. Liao (Ed.), *Advanced educational technology: Research issues and future potential* (pp. 206–217). Springer. [https://doi.org/10.1007/978-3-642-60968-8\\_13](https://doi.org/10.1007/978-3-642-60968-8_13)
- Engelbart, D. C. (1962, October) Augmenting Human Intellect: A Conceptual Framework. SRI Summary Report AFOSR-3223. Prepared for: Director of Information Sciences, Air Force Office of Scientific Research, Washington DC, Contract AF 49(638)-1024. SRI Project No. 3578.
- Fathima, S., & Kore, S. K. (2021). Formulation of the challenges in brain–computer interfaces as optimization problems—a review. *Frontiers in Neuroscience*, 14, Article 546656. <https://doi.org/10.3389/fnins.2020.546656>

- Feng, Y., Chen, D., Zhao, Z., Chen, H., & Xi, P. (2015, August). The impact of students and TAs' participation on students' academic performance in MOOC. In J. Pei, F. Silvestri., & J. Tang (Eds.), *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015* (pp. 1149–1154). Association for Computing Machinery. <https://doi.org/10.1145/2808797.2809428>
- Fu, E. L. F., van Aalst, J., & Chan, C. K. K. (2016). Toward a classification of discourse patterns in asynchronous online discussions. *International Journal of Computer-Supported Collaborative Learning*, 11, 441–478. <https://doi.org/10.1007/s11412-016-9245-3>
- Galikyan, I., Admiraal, W., & Kester, L. (2021). MOOC discussion forums: The interplay of the cognitive and the social. *Computers & Education*, 165, Article 104133. <https://doi.org/10.1016/j.compedu.2021.104133>
- Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs. *User Modeling and User-Adapted Interaction*, 28(2), 127–203. <https://doi.org/10.1007/s11257-018-9203-z>
- Garrison, D., & Akyol, Z. (2012). The Community of Inquiry theoretical framework. In M. G. Moore (Ed.), *Handbook of distance education* (3rd ed., pp. 104–120). Taylor & Francis. <https://doi.org/10.4324/9780203803738.ch7>
- Gay, G. H., & Betts, K. (2020). From discussion forums to eMeetings: Integrating high touch strategies to increase student engagement, academic performance, and retention in large online courses. *Online Learning*, 24(1), 92–117. <https://files.eric.ed.gov/fulltext/EJ1249245.pdf>
- Gee, J. P., & Green, J. L. (1998). Chapter 4: Discourse analysis, learning, and social practice: A methodological study. *Review of Research in Education*, 23(1), 119–169. <https://doi.org/10.3102/0091732X023001119>
- Gilbert, P., & Dabbagh, N. (2004). How to structure online discussions for meaningful discourse: A case study. *British Journal of Educational Technology*, 36(1), 5–18. <https://doi.org/10.1111/j.1467-8535.2005.00434.x>
- Goshtasbpour, F., Swinnerton, B., & Pickering, J. (2021). Twelve tips for engaging learners in online discussions. *Medical Teacher*, 44(3), 244–248. <https://doi.org/10.1080/0142159x.2021.1898571>
- Hammond, M. (2005). A review of recent papers on online discussion in teaching and learning in higher education. *Journal of Asynchronous Learning Networks*, 9(3), 9–23. <https://doi.org/10.24059/olj.v9i3.1782>
- Hill, J., Song, L. and West, R. (2009). Social learning theory and web-based learning environments: A review of research and discussion of implications. *American Journal of Distance Education*, 23(2), 88–103. <https://doi.org/10.1080/08923640902857713>

- Jan, S. K., Vlachopoulos, P., & Parsell, M. (2019). Social network analysis and learning communities in higher education online learning: A systematic literature review. *Online Learning Journal*, 23, 249–264. <https://doi.org/10.24059/olj.v23i1.1398>
- Jiang, S., Fitzhugh, S.M., & Warschauer, M. (2014). Social positioning and performance in MOOCs. EDM.
- Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., & De Kereki, I. F. (2016). Translating network position into performance: importance of centrality in different network configurations. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp.314-323). ACM.
- Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A. C., & Brooks, C. (2017). How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*, 88(1), 43–86. <https://doi.org/10.3102/0034654317740335>
- Kim, M. K., & Ketenci, T. (2019). Learner participation profiles in an asynchronous online collaboration context. *The Internet and Higher Education*, 41, 62–76. <https://doi.org/10.1016/j.iheduc.2019.02.002>
- Kimmons, R., Rosenberg, J., & Allman, B. (2021). Trends in educational technology: What Facebook, Twitter, and Scopus can tell us about current research and practice. *TechTrends*, 65(2), 125–136. <https://doi.org/10.1007/s11528-021-00589-6>
- Kloos, C. D., Alario-Hoyos, C., Muñoz-Merino, P. J., Ibáñez, M. B., Estévez-Ayres, I., & Fernández-Panadero, C. (2020). Educational technology in the age of natural interfaces and deep learning. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 15(1), 26–33. <https://doi.org/10.1109/RITA.2020.2979165>
- Koschmann, T. D. (2011). *Theories of learning and studies of instructional practice*. Springer.
- Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters*, 87(19), Article 198701. <https://doi.org/10.1103/PhysRevLett.87.198701>
- Lee, J. and Recker, M., 2021. The effects of instructors' use of online discussions strategies on student participation and performance in university online introductory mathematics courses. *Computers & Education*, 162, 104084.
- Lemay, D. J., & Doleck, T. (2020). Constructivist educational technology: Re-examining the foundations and state of the literature. *British Journal of Educational Technology*, 51(6), 1905-1906. <https://doi.org/10.1111/bjet.13042>
- Lemay, D. J., Doleck, T., & Bazalais, P. (2021). Transition to online teaching during the COVID-19 pandemic. *Interactive Learning Environments*, 1-12. <https://doi.org/10.1080/10494820.2021.1871633>

- Liao, T. (1996). *Advanced educational technology: Research issues and future potential*. Springer.
- Loizzo, J., & Ertmer, P. (2016). MOOCocracy: The learning culture of massive open online courses. *Educational Technology Research and Development*, 64, 1013–1032.  
<https://doi.org/10.1007/s11423-016-9444-7>
- Marra, R. (2006). A review of research methods for assessing content of computer-mediated discussion forums. *Journal of Interactive Learning Research*, 17(3), 243–267.  
<https://www.learntechlib.org/primary/p/6290/>
- Marx, K. (1973). Fragment on the machines. In *Grundrisse: Foundations of the critique of political economy* (M. Nicolaus, Trans.; ch. 13). Marxists Internet Archive.  
<https://www.marxists.org/archive/marx/works/1857/grundrisse/ch13.htm> (Original work published 1939–1941)
- Moore, R., Yen, C., & Powers, F. (2020). Exploring the relationship between clout and cognitive processing in MOOC discussion forums. *British Journal of Educational Technology*, 52(1), 482–497. <https://doi.org/10.1111/bjet.13033>
- O’Riordan, T., Millard, D., & Schulz, J. (2020). Is critical thinking happening? Testing content analysis schemes applied to MOOC discussion forums. *Computer Applications in Engineering Education*, 29(4), 690–709. <https://doi.org/10.1002/cae.22314>
- Raković, M., Marzouk, Z., Liaqat, A., Winne, P. H., & Nesbit, J. C. (2020). Fine grained analysis of students’ online discussion posts. *Computers & Education*, 157, Article 103982.  
<https://doi.org/10.1016/j.compedu.2020.103982>
- Reed, M. S., Evely, A. C., Cundill, G., Fazey, I., Glass, J., Laing, A., Newig, J., Parrish, B., Prell, C., Raymond C., & Stringer, L. C. (2010). What is social learning? *Ecology and Society*, 15(4), Article r1. <http://www.ecologyandsociety.org/vol15/iss4/resp1/>
- Rosé, C. (2017). Discourse analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gaevic (Eds.), *Handbook of learning analytics* (pp. 105–114). Society for Learning Analytics Research (SoLAR).  
<https://doi.org/10.18608/hla17.009>
- Rosé, C., Wang, Y., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A. and Fischer, F.(2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, 3(3), 237-271. <https://doi.org/10.1007/s11412-007-9034-0>
- Rossi, L. A., & Gnawali, O. (2014, August 13–15). Language independent analysis and classification of discussion threads in Coursera MOOC forums. In J. Joshi, E. Bertino, B. Thuraisingham, & L. Liu (Eds.), *Proceedings of the IEEE International Conference on Information Reuse and Integration (IEEE IRI 2014)* (pp. 654–661). IEEE Systems, Man, and Cybernetics Society (SMC).  
<https://doi.org/10.1109/IRI.2014.7051952>

- Rovai, A. (2007). Facilitating online discussions effectively. *The Internet and Higher Education*, 10(1), 77–88. <https://doi.org/10.1016/j.iheduc.2006.10.001>
- Ruipérez-Valiente, J., Halawa, S., Slama, R., & Reich, J. (2020). Using multi-platform learning analytics to compare regional and global MOOC learning in the Arab world. *Computers & Education*, 146, Article 103776. <https://doi.org/10.1016/j.compedu.2019.103776>
- Santos, J. L., Klerkx, J., Duval, E., Gago, D., & Rodríguez, L. (2014, March). Success, activity and drop-outs in MOOCs an exploratory study on the UNED COMA courses. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge* (pp. 98–102). Association for Computing Machinery. <https://doi.org/10.1145/2567574.2567627>
- Scardamalia, M., & Bereiter, C. (2014). Knowledge building and knowledge creation: Theory, pedagogy, and technology. In R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 397–417). Cambridge University Press. <https://doi.org/10.1017/CBO9781139519526.025>
- Soter, A. O., Wilkinson, I. A., Murphy, P. K., Rudge, L., Reninger, K., & Edwards, M. (2008). What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research*, 47(6), 372–391. <https://doi.org/10.1016/j.ijer.2009.01.001>
- Tang, H., Xing, W., & Pei, B. (2018). Exploring the temporal dimension of forum participation in MOOCs. *Distance Education*, 39(3), 353–372. <https://doi.org/10.1080/01587919.2018.1476841>
- Thomas, M. (2002). Learning within incoherent structures: The space of online discussion forums. *Journal of Computer Assisted Learning*, 18(3), 351–366. <https://doi.org/10.1046/j.0266-4909.2002.03800.x>
- Tirado, R., Hernando, A., & Aguaded, J. I. (2012). The effect of centralization and cohesion on the social construction of knowledge in discussion forums. *Interactive Learning Environments*, 23(3), 293–316. <https://doi.org/10.1080/10494820.2012.745437>
- Tseng, S., Tsao, Y., Yu, L., Chan, C., & Lai, K. (2016). Who will pass? Analyzing learner behaviors in MOOCs. *Research and Practice in Technology Enhanced Learning*, 11, Article 8. <https://doi.org/10.1186/s41039-016-0033-5>
- Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.
- Vygotsky, L. S. (1986). *Thought and language*. MIT Press. <https://mitpress.mit.edu/books/thought-and-language>
- Wang, X., Yang, D., Wen, M., Koedinger, K., & Rosé, C. P. (2015, June 26–29). Investigating how student's cognitive behavior in MOOC discussion forums affect learning gains. In *EDM '15: Proceedings of the 8th International Conference on Educational Data Mining* (pp. 226–233). International Educational Data Mining Society. <https://files.eric.ed.gov/fulltext/ED560568.pdf>

- Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Harvard University Press.  
<https://www.hup.harvard.edu/catalog.php?isbn=9780674943513>
- Winne, P. (2017). Learning analytics for self-regulated learning. In C. Lang, G. Siemens, A. F. Wise, & D. Gaevic (Eds.), *Handbook of learning analytics* (pp. 241–249). Society for Learning Analytics Research (SoLAR). <https://doi.org/10.18608/hla17.021>
- Wise, A. F., Azevedo, R., Stegmann, K., Malmberg, J., Rosé, C. P., Mudrick, N., Taub, M., Martin, S. A., Farnsworth, J., Mu, J., Järvenoja, H., Järvelä, S., Wen, M., Yang, D., & Fischer, F. (2015). CSCL and learning analytics: Opportunities to support social interaction, self-regulation and socially shared regulation. In *CSCL 2015 Proceedings* (pp. 607–614). International Society of the Learning Sciences.
- Wise, A & Cui, Y. (2018). Unpacking the relationship between discussion forum participation and learning in MOOCs: content is key. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp.330-339). Association for Computing Machinery.  
<https://doi.org/10.1145/3170358.3170403>
- Wise, A. F., Cui, Y., Jin, W., & Vytasek, J. (2017). Mining for gold: Identifying content-related MOOC discussion threads across domains through linguistic modeling. *The Internet and Higher Education*, 32, 11–28, <https://doi.org/10.1016/j.iheduc.2016.08.001>
- Wise, A., Hausknecht, S. and Zhao, Y., (2014). Attending to others’ posts in asynchronous discussions: Learners’ online “listening” and its relationship to speaking. *International Journal of Computer-Supported Collaborative Learning*, 9(2), 185-209. <https://doi.org/10.1007/s11412-014-9192-9>
- Wu, D., & Hiltz, S. (2004). Predicting learning from asynchronous online discussions. *Journal of Asynchronous Learning Networks*, 8(2), 139–152. <https://doi.org/10.24059/olj.v8i2.1832>
- Zhu, M., Bergner, Y., Zhang, Y., Baker, R., Wang, Y., & Paquette, L. (2016, April 25–29). Longitudinal engagement, performance, and social connectivity: A MOOC case study using exponential random graph models. In *LAK '16, Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 223–230). Association for Computing Machinery.  
<https://doi.org/10.1145/2883851.2883934>
- Zhu, M., Sari, A., & Lee, M. (2020). A comprehensive systematic review of MOOC research: Research techniques, topics, and trends from 2009 to 2019. *Educational Technology Research and Development*, 68(4), 1685–1710. <https://doi.org/10.1007/s11423-020-09798-x>
- Zou, W., Hu, X., Pan, Z., Li, C., Cai, Y. and Liu, M. (2021). Exploring the relationship between social presence and learners’ prestige in MOOC discussion forums using automated content analysis and social network analysis. *Computers in Human Behavior*, 115, Article 106582.  
<https://doi.org/10.1016/j.chb.2020.106582>

