

# Online Learning Communities in the COVID-19 Pandemic

## Social Learning Network Analysis of Twitter During the Shutdown

<https://doi.org/10.3991/ijai.v2i1.15427>

David John Lemay  
McGill University, Montreal, QC

Tenzin Doleck <sup>(✉)</sup>  
University of Southern California, Los Angeles, CA  
doleck@usc.edu

**Abstract**—This paper presents a social learning network analysis of Twitter during the 2020 global shutdown due to the COVID-19 pandemic. Research concerning online learning environments is focused on the reproduction of conventional teaching arrangements, whereas social media technologies afford new channels for the dissemination of information and sharing of knowledge and expertise. We examine Twitter feed around the hashtags online learning and online teaching during the global shutdown to examine the spontaneous development of online learning communities. We find relatively small and ephemeral communities on the two topics. Most users make spontaneous contributions to the discussion but do not maintain a presence in the Twitter discourse. Optimizing the social learning network, we find many potential efficiencies to be gained through more proactive efforts to connect knowledge seekers and knowledge disseminators. Considerations and prospects for supporting online informal social learning networks are discussed.

**Keywords**—Social learning network analysis, Twitter, online learning, online teaching, COVID-19, online discourse communities, discussion forums

## 1 Teaching and Learning During The 2020 Global Pandemic

The current pandemic has altered social behaviors on a massive scale. With the global COVID-19 pandemic sparked by the emergence of a novel coronavirus in late 2019 [1], large portions of the global population have seen their livelihoods drastically shifted online. At one point, nearly a third of the world population was under lockdown [2] and nearly all students found educational institutions suddenly closed for the indefinite future [3]. Parents scrambled online to find educational activities as they suddenly found schools and the public realm closed indefinitely and had to find ways to occupy their children during the pandemic.

The present context offers an opportunity to explore the formation of online communities over the course of the global shutdown which saw many states enforce public measures such as social distancing to reduce the spread of the virus.

With everyone suddenly housebound, many school-based activities shifted online, facilitated by such open access platforms such as Google Classroom, Zoom, and countless educational tutoring software providers. For budding homeschoolers, search and discovery of new educational activities operates through social media channels. Twitter is one of the predominant means of joining social conversations with an audience of potential global reach. On Twitter, conversations are discovered through a tagging system using the commonplace hashtag (#) to enable keyword filtering of conversational threads. Thus, to search out threads about online teaching and learning, you might simply search for #onlinelearning or #onlineteaching.

In the present study, we mined Twitter posts to explore how the Twitter conversational communities around online teaching and learning evolved over the course of the global shutdown. We compiled Twitter post data for the hashtags #onlinelearning and #onlineteaching from mid-March to end of April 2020 to answer the following question.

### **1.1 Research question**

How did the online conversational communities evolve on Twitter during the global shutdown due to the COVID-2019 pandemic?

Below, we review literature on online learning environments research before describing our analytical approach and presenting and discussing the results of our temporal analysis of the evolution of social learning networks on Twitter during the 2020 global shutdown.

## **2 Literature Review**

Whereas there have not been many studies of informal discourse communities in the educational technology literature, Twitter post data has been mined to study the evolution of Twitter discourse concurrent to large scale events such as political rallies and academic conferences. Twitter, which is now closing in on 330 million users [4], has become an important backchannel for conversations at academic conferences [5]. A conference specific hashtag is usually used to aggregate real-time conversations around a conference [6].

To understand the potential value of interacting and communicating via conference-specific hashtags, it is important to systematically analyze the structure and content of such conversation spaces. Indeed, analysis of conference specific tweets can reveal salient topics during the conference and can provide a posteriori overview of a conference [7]. Moreover, Xie and Luo [8] highlight two key benefits of using Twitter as a backchannel for conferences: (1) broadening immediate participation and (2) diversity in user types and discourse.

In recent years, concomitant with the increasing use of Twitter conference backchannels, there has been a growing interest in studying various aspects of Twitter use in academic conferences in varying disciplines [6] [8-11]. Our current work builds on this stream of research by examining how social learning networks develop in new communication spaces, such as Twitter conference backchannels.

Most online learning environments research has focused on the reproduction of conventional teaching arrangements, whereas social media technologies [12] afford new channels for the dissemination of information and sharing of knowledge and expertise. Research concerning informal online learning environments is benefiting from the development of statistical tools for the study of social networks and the proliferation of online usage data [13], as a result of a long-term trend and recent massive online migration of public life due to social distancing efforts to mitigate the effects of the pandemic.

Little research has examined informal online learning networks as manifested in discussion forums and social media [14]. However, an increasing amount of information is being communicated through these channels [6] [8-11], influencing opinion and shaping conversations. Indeed, increasing fractions of individuals get their news from social media rather than traditional media. In the context of massive open online courses (MOOC), researchers have examined the relationship between discursive interactions on MOOC platforms and academic performance [15]. However, few have studied how social learning networks develop in informal settings such as in conversational threads on Twitter.

Merely because there are no objective measures of performance should not limit the study of informal or social learning as other measures of performance can be devised [14] [16]. For instance, Brinton et al. [16] derived a measure of social learning network efficiency as the average benefit derived from the connections between knowledge seeker and knowledge disseminators on inductively derived topics. Their social learning network optimization algorithm helps connect participants to increase the overall learning benefit. As a variant of the shortest path algorithm it also raises the overall efficiency (or informativeness) of the network graph [17].

### **3 Method**

#### **3.1 Theoretical framework**

This work is grounded in the communities of practice framework, particularly the notion of boundary crossing [18], as it applies to the site of informal learning networks and exchanges where ideas and artifacts are exchanged between a constellation of communities with their own members and practices. These boundary interactions can be short lived and focused as they can be longer lasting in some formalized structure, but they do not proceed from identification and participation.

Concretely, in a public forum like Twitter, we do not assume any long-lasting connections in the data. In fact, in this study, we examine follow-up and repeated

interactions to study the nature of the interactions that occur in Twitter conversation on online teaching and learning.

### **3.2 Research design**

This retrospective study uses an exploratory case study methodology [19] employing multiple measures to describe the phenomena of interest in sufficient detail to enable reproduction of our work in other contexts.

### **3.3 Data**

Twitter posts with the hashtags #onlinelearning and #onlineteaching were collected from mid-March to end of April 2020 using: R programming language [20], RStudio [21], and the R tweet package [22]. The R tweet package retrieves tweets from the last 6–9 days. For this study, we created five datasets using tweets at approximately 9-day intervals.

### **3.4 Analytical procedure**

Social network analysis was conducted using the Python programming ecosystem, including Pandas for data manipulation and NetworkX for social network analysis. We employed Python 3.7.6 and NetworkX 2.4 in our analyses. Undirected network graphs were constructed from the Twitter post and response data using the NetworkX package [23]. Several metrics were calculated based on the adjacency matrix of the undirected network graph, including descriptive statistics such as nodes and edges, as well as average neighbor degree, number of connected components, density, global efficiency, and community distribution. Global efficiency is a small-world metric that assesses the overall informativeness, or how efficiently information is communicated, of a network calculated as the multiplicative inverse of the shortest path between successive nodes [17]. The average degree refers to the average number of connections for each node. The number of connected components tells us the number of connected groups within the data. The density is the ratio of connections to the overall graph of possible connections. All these measures are implemented in the NetworkX social networking analysis package. We inspected the circular graphs of the networks over the six-week period.

The edges represent post-reply connections. We take the liberal assumption that these edges form an undirected graph. Whereas such connections can reasonably be interpreted as uni-directional, we believe it is justified on the basis of our study's objectives. Our goal is to map out the community interactions and these are defined as bi-directional as members are posting on the same topic and are assumed to be engaging on the topic of discussion, that is, whether or not they reply to specific posts, it is reasonable to assume that they are engaging with the discussion threads around online teaching and learning. Hence, post-replies represent instances of interactions among community members. From a pragmatic perspective, the relaxing of this assumption

also allows us to include metrics that have only been implemented for undirected graphs such as the calculation of connected components and local and global efficiency.

We employ Brinton et al.'s [16] social learning network optimizer as implemented in [14] to identify potential gains from connecting users based on the information they are likely to interact on. Briefly, the algorithm casts the problem as a convex optimization problem that seeks to connect knowledge seekers and knowledge disseminators based on the weighted average of the questioning and answering tendencies per topic. Topics are inductively derived using latent Dirichlet allocation [24]. The solution is computed using alternating direction method of multipliers, which uses Lagrange multipliers to find minima subject to some linear constraints. Brinton et al. [16] demonstrate that their approach has convergence guarantees and is performant for networks with millions of parameters.

For analyzing the tweets, the following packages were used: tidytext; dplyr; syuzhet; stringr; tm; and ggplot2. Before conducting sentiment analysis, to the extracted tweets, we applied the following preprocessing steps: removed stopwords, removed urls, converted text to lowercase, and removed punctuation.

## 4 Results

The results are presented in the following order: social network analysis followed by sentiment analysis.

### 4.1 Social network analysis

We observed an early peak in interest in online teaching and learning coinciding with the beginning of the global social distancing measures. As can be observed in Table 1, we see a maximum number of nodes and edges and the highest average degree at T1, declining at T2 and further at T3. This is reflected in the number of conversational groups (see Table 2) which drop from 623 to 409 (for online learning) and 183 to 59 (for online teaching).

Whereas the average neighbor degree and global efficiency remain stable (see Table 2). This is interpreted as a result of the limited interactions on the Twitter threads relative to the untapped potential interactions. In fact, the great majority of communities are distributed between communities of degree two and communities of degree four for both #onlinelearning and #onlineteaching (see Table 3).

In Figures 1-10, we find the circular graphs representing the user interactions over at five time points (approximately nine-day intervals) over the duration of the study. In the case of the discussion threads on online teaching and learning, most communities contain between two and four members only (See Table 3). Indeed, we do not see large clusters or connections in the circular graphs but a uniform distribution of connections since each node only has few connections. Moreover, we notice an initial burst of interaction steadily declining across subsequent time points.

Finally, we combined the post data for both #onlinelearning and #onlineteaching to generate the full adjacency graph for the six-week period and passed it to the social

learning network optimizer. We set the threshold, or the gain from each iteration, to 0.1, the seeking constraint to 1.25 and the disseminating constraint to 0.75 following Brinton et al. [16]. The constraints capture the diminishing benefits received from specific interactions. The optimization converged after three iterations. The observed learning benefit (before optimization) was 0.65, or essentially zero. The optimized network found a potential learning benefit of 24.78, which could be derived from a more efficient distribution of connections. The observed network is manifestly inefficient, with  $0.65/24.78 \times 100 = 2.63\%$  efficiency.

Figures 1 to 10 present the circular graphs for the two conversational communities (#onlinelearning and #onlineteaching) over the six weeks of the study. Visual inspection of the graphs reveals a steadily increasing sparsity from the initial peak of interest.

**Table 1.** Descriptive Statistics

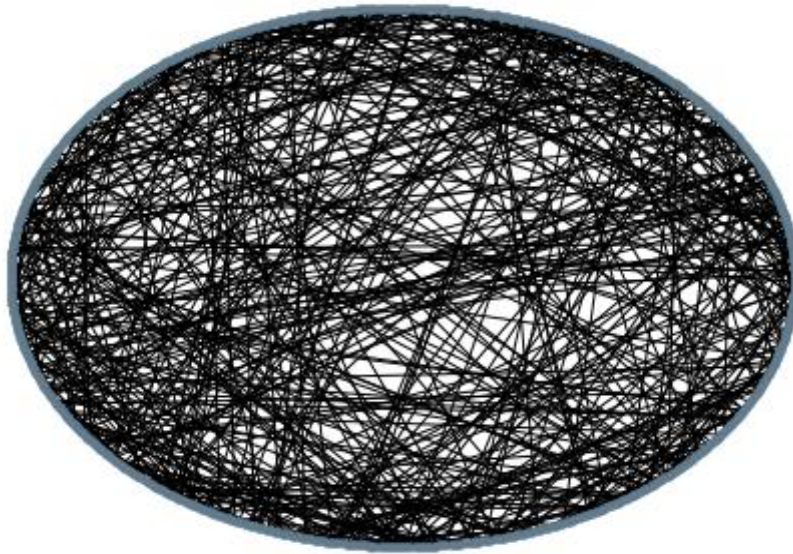
Online Learning	Nodes	Edges	Average Degree
T1	1156	758	1.3114
T2	961	593	1.2341
T3	678	430	1.2684
T4	726	474	1.3058
T5	742	491	1.3235
Online Teaching	Nodes	Edges	Average Degree
T1	329	208	1.2644
T2	134	82	1.2239
T3	113	70	1.239
T4	120	78	1.3000
T5	103	64	1.2427

**Table 2.** Network Connectivity

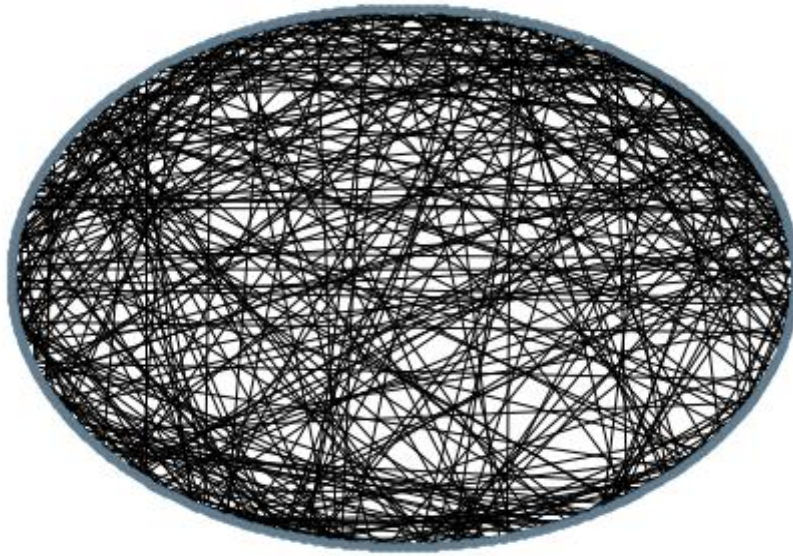
Online Learning	Connected Components	Density	Average Neighbor Degree	Global Efficiency
T1	623	0.0011	1.3654	0.0010
T2	519	0.0013	1.1794	0.0011
T3	367	0.0019	1.3883	0.0099
T4	406	0.0018	1.3993	0.0015
T5	409	0.0018	1.4691	0.0015
Online Teaching	Connected Components	Density	Average Neighbor Degree	Global Efficiency
T1	183	0.0039	1.2957	0.0031
T2	71	0.0092	1.2711	0.0080
T3	63	0.0111	1.1475	0.0084
T4	70	0.0109	1.2000	0.0078
T5	59	0.0122	1.0971	0.0122

**Table 3.** Degree Distribution

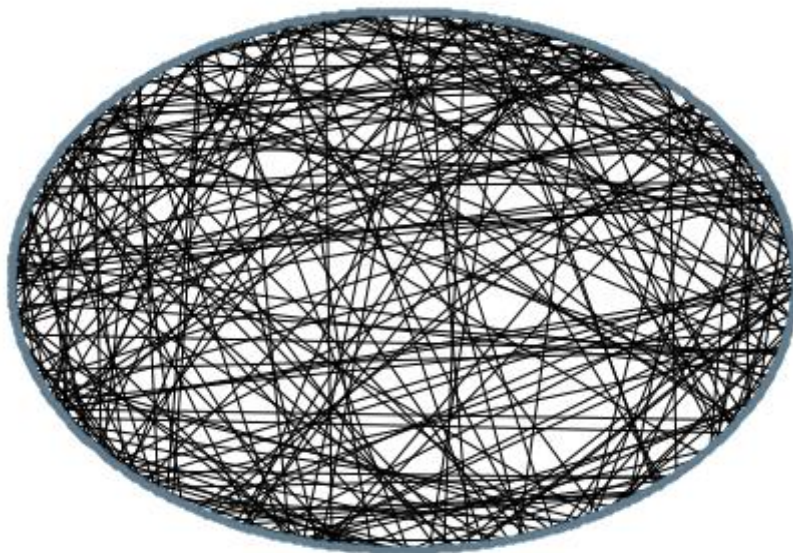
<b>Online Learning</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
T1	0	864	257	20	7	3	2	2	0	1
T2	0	758	185	14	4	0	0	0	0	0
T3	0	531	130	11	2	1	1	0	1	1
T4	0	547	171	13	7	2	0	1	0	1
T5	0	542	165	12	5	1	0	0	0	0
<b>Online Teaching</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
T1	0	257	64	6	1	0	0	0	0	0
T2	0	109	22	2	0	1	0	0	0	0
T3	0	89	21	3	1	0	0	0	0	0
T4	0	87	31	1	1	0	0	0	0	0
T5	0	78	25	0	0	0	0	0	0	0



**Fig. 1.** Online Learning T1

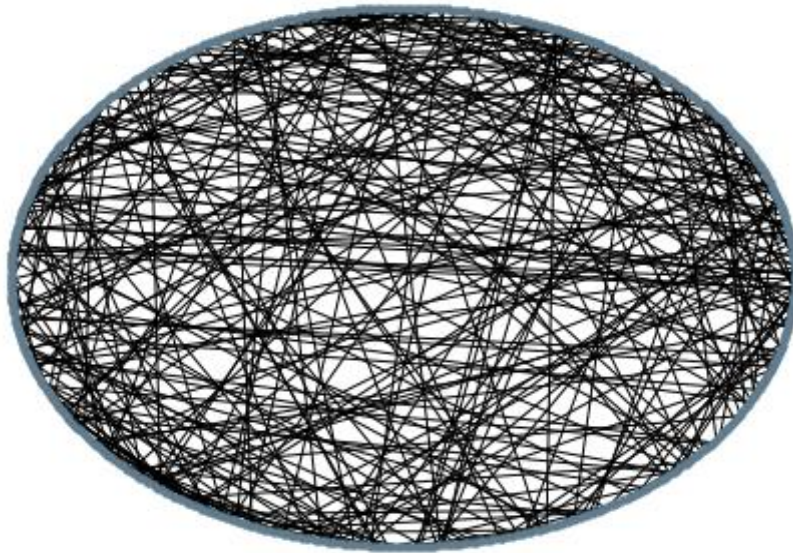


**Fig. 2.** Online Learning T2

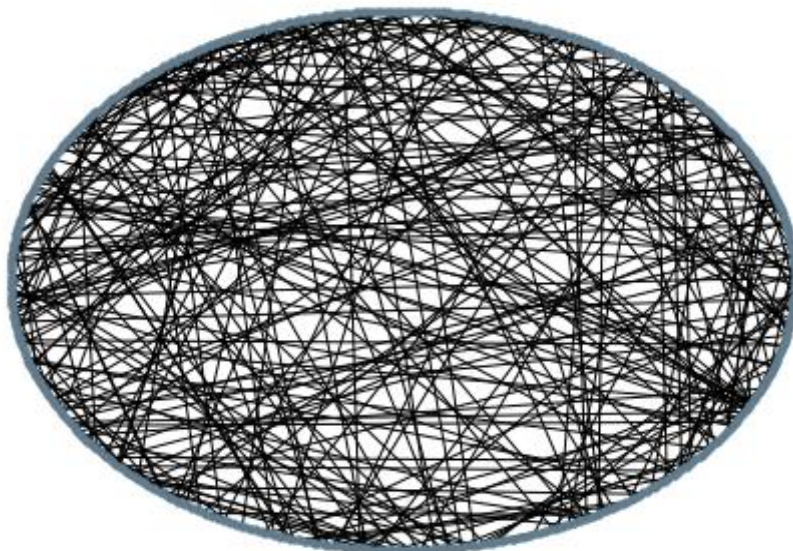


**Fig. 3.** Online Learning T3

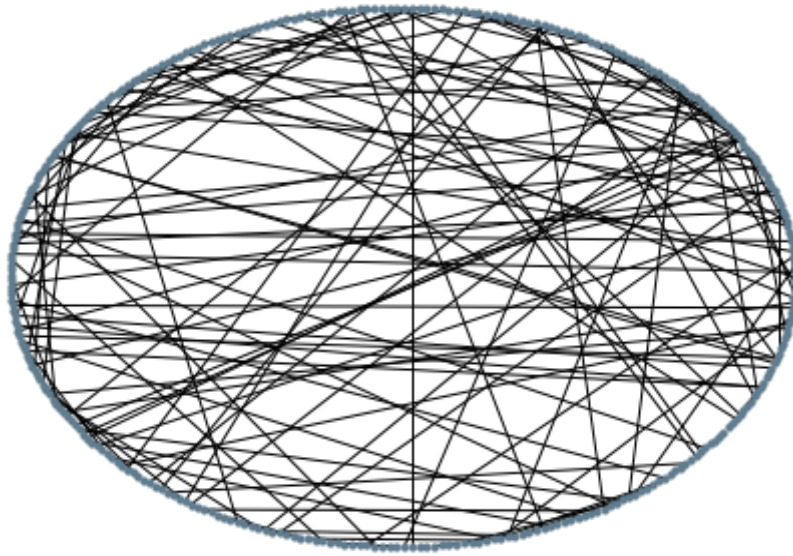




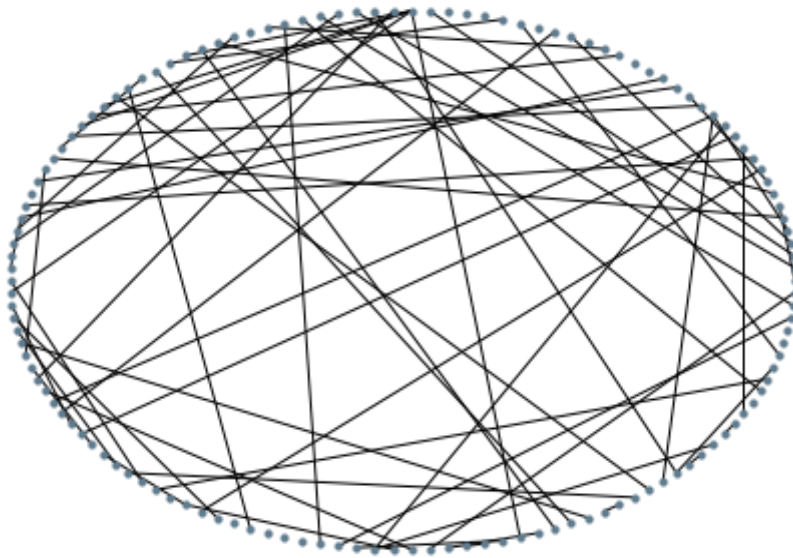
**Fig. 4.** Online Learning T4



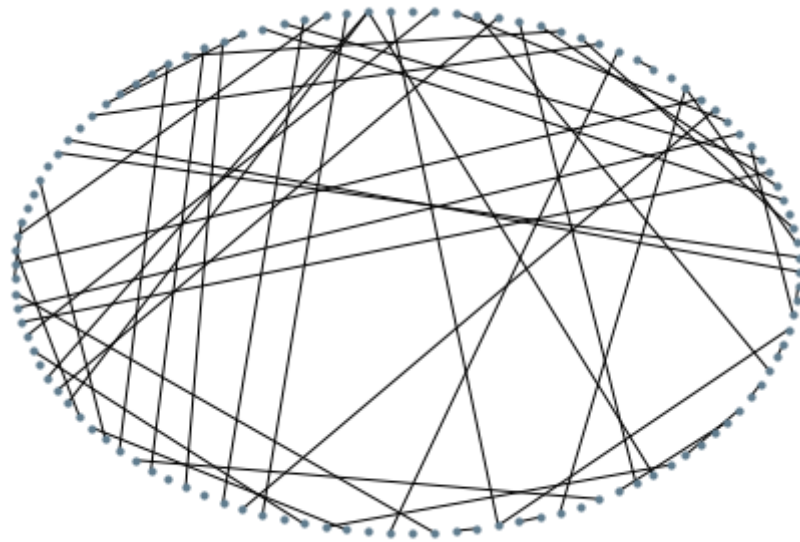
**Fig. 5.** Online Learning T5



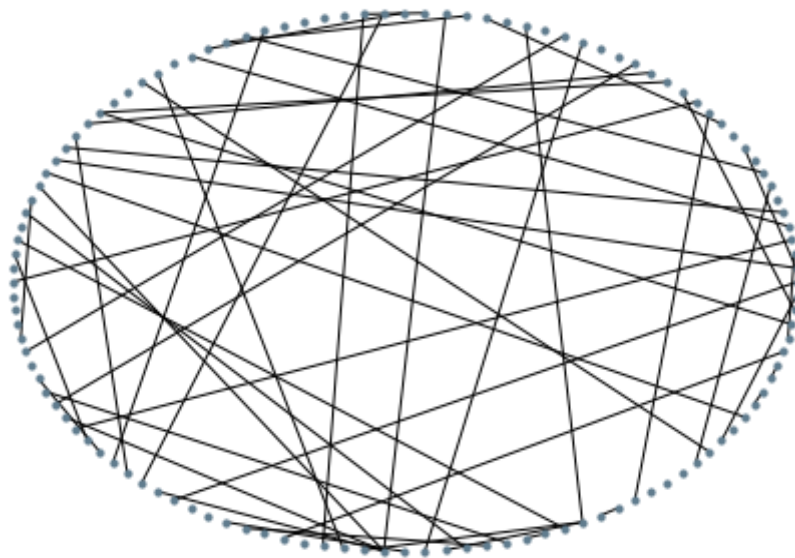
**Fig. 6.** Online Teaching T1



**Fig. 7.** Online Teaching T2



**Fig. 8.** Online Teaching T3



**Fig. 9.** Online Teaching T4

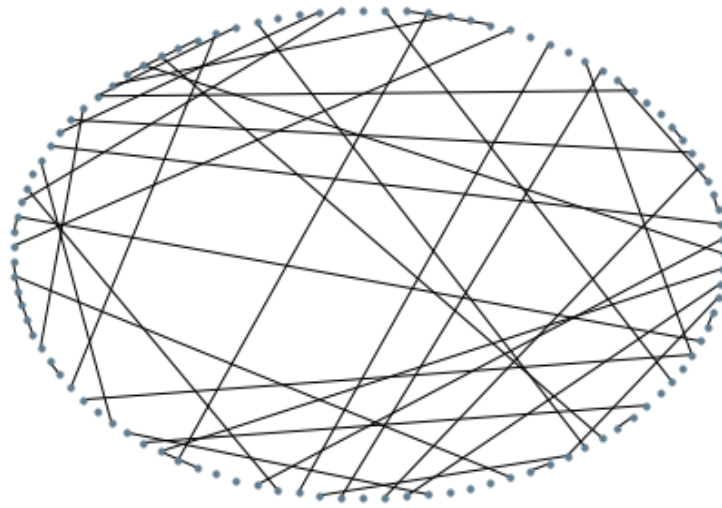


Fig. 10. Online Teaching T5

## 4.2 Sentiment analysis

The Syuzhet R [25] package was used to analyze the sentiment about the content of the tweets. Syuzhet provides the following sentiment measures using the NRC emotion lexicon [26]: eight emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (positive and negative). Overall, the positive sentiment was the most prevalent sentiment for both #onlinelearning (see Figure 11) and #onlineteaching (see Figure 12).

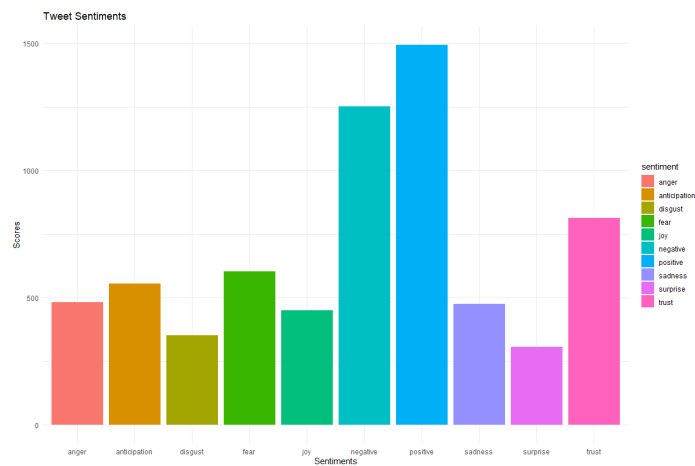


Fig. 11. Online Learning-Overall Sentiments

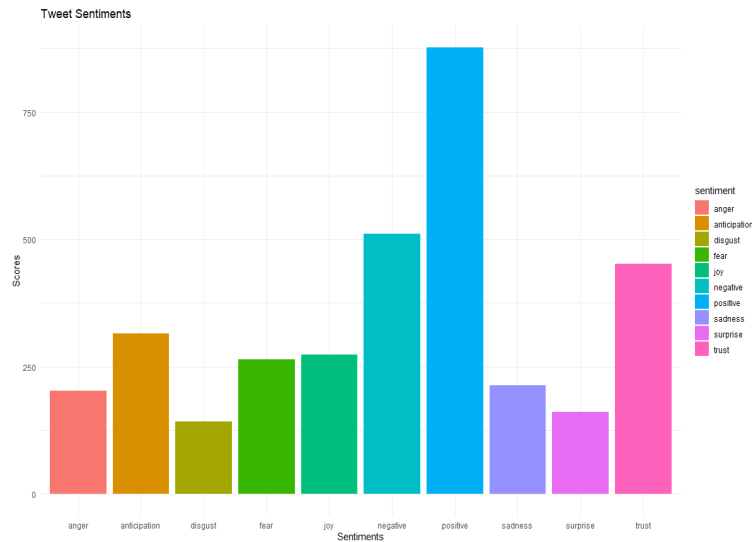


Fig. 12. Online Teaching-Overall Sentiments

## 5 Discussion

Our study finds, for both both #onlinelearning and #onlineteaching, sparse networks and limited, generally positive, exchanges where the great majority of conversational groups contain between two and four members. That is, in the Twitter conversations, many posts are restricted to atomic exchanges [27-28] of one post and one reply. Only very few posts lead to follow-up exchanges. The size is relatively stable, though we identify a spike of interest in online learning and teaching that gradually tapers off over the course of the period considered in the present study.

These results highlight the limited efficiency of the Twitter conversations as a medium for informative communicative exchanges [17]. Such findings support the view that informative or instructive conversational exchanges need to be supported as they do not spontaneously form [16] [29]. In more conventional online learning environments, many tools are at the disposal of educational platform providers and distance learning instructors to support interaction and learning and facilitate discursive exchanges from optimizing social learning network algorithmically [16] to embedding discursive exchanges in online activities and artefacts [29].

It appears evident that supporting informative (and instructive) conversations could benefit from outside support or some scaffolding to improve the quality of online conversational interactions. Indeed, when applying Brinton et al.'s [16] social learning network optimizer, we found large potential learning benefits from connecting knowledge seekers and knowledge disseminators. This suggests potential gains from algorithmically connecting users to sustain conversational interactions and increase knowledge discovery. Indeed, many social media platforms employ such tools to

enhance the quality of social interactions. However, it appears clear that there is room for improving communication on Twitter, especially for creating instructive exchanges.

### 5.1 Limitations

Given the retrospective nature of the study, it is not possible to infer any causal relationships. We are limited to describing overall trends in the Twitter data. Our analysis is limited to two conversational threads (#onlinelearning and #onlineteaching) and six weeks is a short period to study trends. Although, we believe that our study is warranted given the exceptional circumstances. Our social learning network analysis could be extended by extending the number of threads, platforms, and the duration of the sampling period.

### 5.2 Future directions

As an exploratory study of social learning networks on the Twitter platform, we believe the present study serves instrumentally to demonstrate the value of social network analysis methodologies to study the social graph of knowledge creation and dissemination in online social media platforms. Given the democratizing effects of online access to information [30], we call on researchers to extend the study of online learning environments to informal learning networks emergent on social media, and other non-conventional learning environments afforded by new information and communication technologies [31-32]. Study of social learning networks can describe the social processes underpinning shared knowledge construction.

## 6 References

- [1] E. Dong, H. Du and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time", *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533-534, 2020. Available: [https://doi.org/10.1016/s1473-3099\(20\)30120-1](https://doi.org/10.1016/s1473-3099(20)30120-1)
- [2] K. Buchholz, "Infographic: What Share of the World Population Is Already on COVID-19 Lockdown?" *Statista Infographics*, 2020. [Online]. Available: <https://www.statista.com/chart/21240/enforced-covid-19-lockdowns-by-people-affected-per-country/>
- [3] "COVID-19 Educational Disruption and Response", *UNESCO*, 2020. [Online]. Available: <https://en.unesco.org/covid19/educationresponse>.
- [4] Y. Lin, "10 Twitter Statistics Every Marketer Should Know in 2020 [Infographic]", *Oberlo*, 2020. [Online]. Available: <https://www.oberlo.ca/blog/Twitter-statistics>.
- [5] H. Risser and G. Waddell, "Beyond the backchannel: tweeting patterns after two educational conferences", *Educational Media International*, vol. 55, no. 3, pp. 199-212, 2018. Available: <https://doi.org/10.1080/09523987.2018.1512449>
- [6] C. Greenhow, J. Li and M. Mai, "From tweeting to meeting: Expansive professional learning and the academic conference backchannel", *British Journal of Educational Technology*, vol. 50, no. 4, pp. 1656-1672, 2019. Available: <https://doi.org/10.1111/bjet.12817>

- [7] J. Letierce, A. Passant, J. Breslin and S. Decker, "Using Twitter during an academic conference: The#iswc2009 use-case", in *Fourth International AAAI Conference on Weblogs and Social Media*, Washington, DC, 2010.
- [8] Q. Xie and T. Luo, "Examining User Participation and Network Structure via an Analysis of a Twitter-Supported Conference Backchannel", *Journal of Educational Computing Research*, vol. 57, no. 5, pp. 1160-1185, 2018. Available: <https://doi.org/10.1177/0735633118791262>.
- [9] R. Kimmons and G. Veletsianos, "Education scholars' evolving uses of Twitter as a conference backchannel and social commentary platform", *British Journal of Educational Technology*, vol. 47, no. 3, pp. 445-464, 2016. Available: <https://doi.org/10.1111/bjet.12428>.
- [10] D. Lemay, R. Basnet, T. Doleck and P. Bazalais, "Social network analysis of Twitter use during the AERA 2017 annual conference", *Education and Information Technologies*, vol. 24, no. 1, pp. 459-470, 2018. Available: <https://doi.org/10.1007/s10639-018-9787-2>
- [11] A. Wong, I. Capel and M. Malbrain, "Social media in critical care: Fad or a new standard in medical education? An analysis of international critical care conferences between 2014 and 2017", *Journal of the Intensive Care Society*, vol. 20, no. 4, pp. 341-346, 2019. Available: <https://doi.org/10.1177/1751143719846432>
- [12] P. Tess, "The role of social media in higher education classes (real and virtual) – A literature review", *Computers in Human Behavior*, vol. 29, no. 5, pp. A60-A68, 2013. Available: <https://doi.org/10.1016/j.chb.2012.12.032>
- [13] D. Lemay and T. Doleck, "Predicting completion of massive open online course (MOOC) assignments from video viewing behavior", *Interactive Learning Environments*, pp. 1-12, 2020. Available: <https://doi.org/10.1080/10494820.2020.1746673>.
- [14] T. Doleck, D. Lemay and C. Brinton, "Evaluating the Effectiveness of Social Learning Networks: Perspectives for Harnessing Learning Analytics to Improve Discussions." [under review].
- [15] T. Chiu and T. Hew, "Factors influencing peer learning and performance in MOOC asynchronous online discussion forum", *Australasian Journal of Educational Technology*, vol. 34, no. 4, 2018. Available: <https://doi.org/10.14742/ajet.3240>
- [16] C. Brinton et al., "On the Efficiency of Online Social Learning Networks", *IEEE/ACM Transactions on Networking*, vol. 26, no. 5, pp. 2076-2089, 2018. Available: [10.1109/tnet.2018.2859325](https://doi.org/10.1109/tnet.2018.2859325).
- [17] V. Latora and M. Marchiori, "Efficient Behavior of Small-World Networks", *Physical Review Letters*, vol. 87, no. 19, 2001. Available: <https://doi.org/10.1103/physrevlett.87.198701>
- [18] E. Wenger, *Communities of practice: Learning, meaning, and identity*. Cambridge university press, 1999.
- [19] R. Stake, *The art of case study research*. Thousand Oaks, California: Sage, 2010.
- [20] *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020.
- [21] *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc., 2015.
- [22] M. Kearney, "rtweet: Collecting and analyzing Twitter data", *Journal of Open Source Software*, vol. 4, no. 42, p. 1829, 2019. Available: <https://doi.org/10.21105/joss.01829>.
- [23] A. Hagberg, D. Schult and P. Swart, "Exploring network structure, dynamics, and function using NetworkX", in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Pasadena, CA, 2008, pp. 11–15.

- [24] D. Blei, A. Ng and M. Jordan, "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, vol. 3, pp. 993-1022, 2003.
- [25] M. Jockers, *Package 'syuzhet'*. 2017.
- [26] "NRC Emotion Lexicon", *Saifmohammad.com*, 2020. [Online]. Available: <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>.
- [27] H. Nassaji, "What's the use of 'triadic dialogue'? an investigation of teacher-student interaction", *Applied Linguistics*, vol. 21, no. 3, pp. 376-406, 2000. Available: <https://doi.org/10.1093/applin/21.3.376>
- [28] J. Sinclair, *Towards an Analysis of Discourse*. London: OUP, 1978.
- [29] V. Dennen and K. Wieland, "From Interaction to Intersubjectivity: Facilitating online group discourse processes", *Distance Education*, vol. 28, no. 3, pp. 281-297, 2007. Available: <https://doi.org/10.1080/01587910701611328>
- [30] D. Lemay, T. Doleck and P. Bazelais, "Context and technology use: Opportunities and challenges of the situated perspective in technology acceptance research", *British Journal of Educational Technology*, vol. 50, no. 5, pp. 2450-2465, 2019. Available: <https://doi.org/10.1111/bjet.12859>
- [31] T. Doleck, S. Lajoie and P. Bazelais, "Social networking and academic performance: A net benefits perspective", *Education and Information Technologies*, vol. 24, no. 5, pp. 3053-3073, 2019. Available: <https://doi.org/10.1007/s10639-019-09913-3>
- [32] P. Bazelais, T. Doleck and D. Lemay, "Investigating the predictive power of TAM: A case study of CEGEP students' intentions to use online learning technologies", *Education and Information Technologies*, vol. 23, no. 1, pp. 93-111, 2017. Available: <https://doi.org/10.1007/s10639-017-9587-0>

## 7 Authors

**David John Lemay** is a postdoctoral fellow at McGill University. Educational psychologist, experienced in quantitative and qualitative research methods, employed in program evaluation, policy analysis and research. Focused on the study of communicative interaction and technology use.

**Tenzin Doleck** is a postdoctoral fellow at the University of Southern California.

Email: [doleck@usc.edu](mailto:doleck@usc.edu)

Article submitted 2020-05-05. Resubmitted 2020-05-28. Final acceptance 2020-05-28. Final version published as submitted by the authors.