UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Online network topology shapes personal narratives and hashtag generation

Permalink

https://escholarship.org/uc/item/6pv4z0j5

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Priniski, Hunter
Linford, Bryce
Krishna, Sai
et al.

Publication Date

2024

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Online network topology shapes personal narratives and hashtag generation

J. Hunter Priniski
Department of Psychology
UCLA

Bryce Linford
Department of Psychology
UCLA

Sai Krishna
Information Sciences Insitute
USC

Fred Morstatter
Information Sciences Institute
USC

Jeff Brantingham
Department of Anthropology
UCLA

Hongjing Lu
Department of Psychology
UCLA

Abstract

While narratives have shaped cognition and cultures for centuries, digital media and online social networks have introduced new narrative phenomena. With increased narrative agency, networked groups of individuals can directly contribute and steer narratives that center our collective discussions of politics, science, and morality. We report the results of an online network experiment on narrative and hashtag generation, in which networked groups of participants interpreted a text-based narrative of a disaster event, and were incentivized to produce matching hashtags with their network neighbors. We found that network structure not only influences the emergence of dominant beliefs through coordination with network neighbors, but also impacts participants' use of causal language in their personal narratives.

Keywords: online networks, hashtags, personal narratives, NLP, causal structure

Introduction

The internet has remapped how people interpret and discuss events. Digital technology enables organizing efforts that transcend geographic limitations, leading to some of the largest demonstrations in history (Dawson, 2020; Yang, 2016). However, these same platforms can create information environments that foster extremism, hate, and anti-democratic ideals. Echochambers, where online communities consume media that confirms their beliefs and identities, are breeding grounds for unreliable information and conspiracy theories (Sasahara et al., 2021). Empirical research is necessary to understand how networked environments shape belief formation at both individual and group levels, so as to better control the dynamics of information spread, and to possibly mitigate the harm of misinformation and social segregation.

Empirically investigating networked behavior is a challenging task because the narratives arising in online contexts are sprawling and unwieldy (Tangherlini, Shahsavari, Shahbazi, Ebrahimzadeh, & Roychowdhury, 2020), largely due to the dynamic and complex nature of modern social media environments. For instance, the interactivity of digital media and online social networks allows for a new sense of *narrative agency* (Yang, 2016). Through the production and sharing of social media data, users can embed personal narratives and express their point of view on an event or social issue (Boyd, Golder, & Lotan, 2010). From these low-level interactions, a collective narrative emerges within a group of individuals, who bring their own prior beliefs and goals when characterizing an event or issue, which directly shapes the narrative shared by a network (Dawson, 2020).

Hashtags are a potent force for narrative interaction on social media; they allow users to tag personal narratives and contribute to online discourse by indexing their produced content with proxy topic labels. Previous research has shown that hashtags are concise representations of the narratives (Giaxoglou, 2018; Dawson, 2020), and connect spatially disorganized groups according to the content of their narratives and goals (Papacharissi, 2015, 2016; Howard & Hussain, 2013). Across an online network, hashtags categorize social media discourse for effective indexing and search, and can allow interpersonal signaling and sense making between instances (Papacharissi, 2015, 2016).

Previous research on hashtags has primarily focused on how they are used in real-world (i.e., "scale-free") networks, by focusing on the linguistic and semantic structure of popular hashtags (Booten, 2016), or modeling the dynamics underlying their spread online (Cunha et al., 2011; Lin, Margolin, Keegan, Baronchelli, & Lazer, 2013). For instance, hashtags fall into one of two categories. *Focal hashtags* tag posts with broad semantic topics to relate posts to broader discussions and movements across an onine network. A second set of *individualistic hashtags* make the distribution of hashtags heavy-tailed, as they co-occur with focal hashtags while allowing users use to signal personal narratives (Booten, 2016). Furthermore, hashtags generally fall into a "winner' or "loser" category (Lin et al., 2013), in that while many hashtags initially compete for popularity, only a small set of hashtags will persist to allow for broader narrative collaboration across the network. It is unclear how endorsed narratives shape the production of hashtags and how hashtag dynamics are sensitive to network structures. To address these important questions, empirical experiments must mirror the interaction structure and media of online environments to effectively account for how groups engage with narratives in real-world networks, and how group behaviors are affected by an individual's representation of the underlying event.

One approach to studying narrative interaction is to run social network experiments, where a group of participants are placed in a social network and interact with one another. This paradigm allows network structure to be manipulated under experimental control. Social network experiments have historically used relatively simple materials to measure how varying social network structure (e.g., node connectivity) influences the adoption of normative behaviors. For example, Centola and Baronchelli (2015) asked participants to coordi-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

nate on naming an image of a face with network neighbors. They found that interactions within *homogeneously-mixed networks*—fully-connected networks where each participant is linked to every other participant—support the emergence of normative behaviors (i.e., the full network aligning on a single name), while *spatially-embedded networks*, where each participant is linked only to a handful of neighbors, did not. Here, we extend this well-established network experiment design by using naturalistic narrative materials and interaction behaviors akin to those in real-world online networks. We hypothesize that repeated interactions in localized neighborhoods will allow groups to coordinate more effectively, but neighborhoods spanning a fully-connected network will be more likely to produce dominant behaviors. To test these hypotheses, we developed fine-grained measures of behavioral coherence at both the level of local coordination between pairs of participants, and at the level of global convergence across the full network. In addition, we applied an NLP model to compare people's causal language use in personal narratives before and after networked interaction.

Online Network Experiment on Personal Narrative and Hashtag Generation

Participants

We sampled a total of $N = 420$ participants from the Prolific and SONA subject pools, and placed them into one of ten network conditions. Conditions vary the size of a network ($N = 20, 50, 100$) as well as its connectivity (homogeneously-mixed/fully-connected; spatially-embedded/ring-like). We collected six network runs for $N = 20$ (three runs per network structure), and single runs for each structure of $N = 50$ and $N = 100$. Participants $N = 20$ and $N = 50$ conditions were sampled using Prolific. For the $N = 100$ condition, we recruited undergraduates in the Department of Psychology at UCLA through SONA . We posted initial recruitment surveys a week prior to each run in SONA and a few hours prior to each run in Prolific. Participants who received the most points at the end of the experiment received an additional $10 bonus.

Materials

Across all network conditions, participants first read a four-paragraph narrative description of the Fukushima nuclear disaster. The narrative explains how a large earthquake triggered a tsunami that caused damage to a nuclear reactor and resulted in radiation leaks, population displacement, and an energy-saving movement "Setsuden". We selected this narrative based on a pilot study demonstrating that it resulted in the most diverse set of hashtags within a set of tested narratives related to natural and financial disasters. This is likely because the narrative describes a rich set of causal relations (a generative causal chain producing a branching common cause sequence) and included both negative (e.g., displacement, poisoning) and positive effects (e.g., energy saving movement). Fig. 1 illustrates the causal structure of the Fukushima disaster narrative.

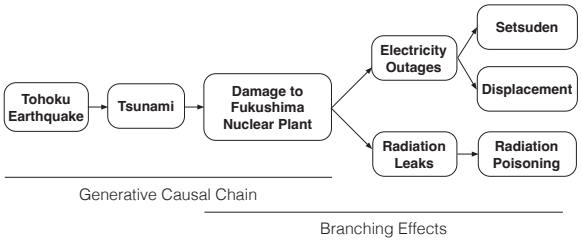


Figure 1: **Causal model communicated in the nuclear disaster narrative**. This diagram is just for illustration purposes, participants did not see this diagram. They read a four-paragraph narrative describing how the Tōhoku earthquake triggered a massive tidal wave that damaged the Fukushima Nuclear Power Plant, resulting in electricity outages, radiation leaks and poisoning, human displacement, and *Setsuden*, a national energy-saving holiday.

Experimental Design and Procedure

Network Experiment Method We used the open-source framework OTree written in Python (Chen, Schonger, & Wickens, 2016), and hosted experiments on a Linux server. Participants joined the experiment through a Qualtrics survey that directed participants to the network experiment.

Our social network experiment proceeded in three steps. First, we randomly assigned each participant as a player in a network that defined who may interact with whom on a given trial. Second, we assigned interactions between individual participants on each trial. Third, we rewarded participants based on the outcome of their interactions. We can specify this process using graph theory notation. The first step is to initialize a fixed graph $G(N,E)$, defined by a set N nodes representing individual participants connected through an edge set E. We discuss below the specific graph structures used. The second step iterates over T trials. On a given trial $t \in T$, connection (edge) configurations follow mixing participants randomly within a participant's neighborhood. The third step is to identify and reward coordinated behavior. If the response from participant n_i on trial t is r_i^t, then participants n_i and n_j coordinate if $r_i^t = r_j^t$.

Procedure The experiment consisted of three blocks: a pre-interaction block, a networked interaction block (described above), and a post-interaction block. This three-block design allowed us to assess behavioral dynamics during the networked interaction block, in addition to examining whether networked interaction can shift beliefs from pre- to post-interaction blocks.

In the **pre-interaction block**, participants read a four paragraph narrative describing the Fukushima nuclear disaster, and then were asked to write a "tweet" (within a 140 word limit) and ten hashtags characterizing the events described in the narrative.

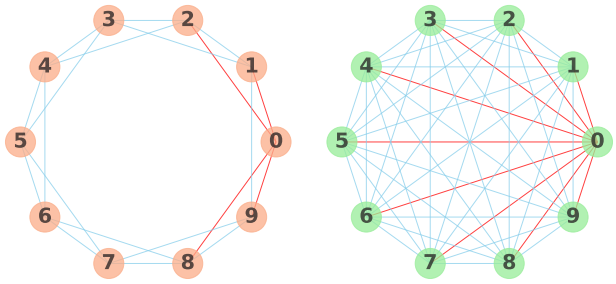


Figure 2: **Two neighborhood structures tested in this experiment:** Spatially-embedded (**left**) and homogeneously-mixed (**right**) networks with $N = 10$ nodes. Red edges represent the neighborhoods for a hypothetical node 0 in both networks. As a network's size grows, the diameter of spatial networks grow whereas homogeneous networks maintain a diameter of 1.

In the **networked interaction block**, participants joined a network experiment through real-time interaction on the online platform. Participants were assigned to one of six experimental conditions based on the size of the network ($N = 20$; 50; 100) and network structure (spatially-embedded and homogeneously-mixed; see Fig. 2). Regardless of network size, nodes in spatial networks have a consistent neighborhood size $k = 4$, meaning each participant would interact with four other participants in the entire experiment. Neighborhood size in homogeneous networks is $N - 1$, as each participant can interact with any of the remaining participants. A consequence is that the network diameter (i.e., the largest geodesic distance in the connected network) was consistently 1 in all tested homogeneous networks, but grows as a function of size in spatial networks. Both of these features of network topology (i.e., size and diameter) uniquely influence the emergence of shared behavior in online networks (Anagnostopoulos, Becchetti, Castillo, Gionis, & Leonardi, 2012).

The networked interaction block consisted of 40 trials, where participants interacted with their partners based on the edge structure in the assigned network. On each trial, participants were instructed to write a single hashtag describing the narrative they read in the pre-interaction block. After participants submitted their hashtag response, they were then presented with a new page showing their own hashtag response, their partner's hashtag response, whether they received a point for matching responses with their partner, and their cumulative reward point.

Following networked interactions, participants entered a **post-interaction block** where they wrote one more "tweet" for the same narrative and another ten hashtags describing the event. Before completing the experiment, they provided demographic information.

Results and Discussion

Global connections support emergence of dominant responses We fit a Bayesian generalized linear model (GLM) to predict how the two network structures (spatially-embedded vs homogeneously-mixed structure) support the emergence of a dominant hashtag response. We assume that the proportion of participants who produced the dominant hashtag on trial t follows a Beta distribution, a commonly used distribution to predict proportion values (McElreath, 2016); we used uninformative priors (i.e., $\mathcal{N}(0, 10)$) over regression coefficients. Specifically, the GLM model predicted the proportion value as a function of trial number (i.e., *Trial*) interacted with network structure (*Spatial* vs *Homogeneous*), while controlling for network *size*. Here, we simply predict the proportion of players in a network producing the dominant response on a given trial, which could be different hashtags across different trials for a single run.

As shown in Fig. 3, shared responses emerge from networked interactions in both network structures ($\beta_{Trial} = 0.04$, 95% CI $[0.04, 0.05]$), however dominant responses emerge more easily in homogeneously-mixed networks than spatially-embedded networks ($\beta_{Trial \times Spatial} = -0.01$, 95% CI $[-0.02, -0.00]$). We found that within the confines of a given experimental run (i.e., 40 interaction trials), the shared responses emerge more quickly in smaller network sizes than larger ones ($\beta_{Size} = -0.01$, 95% CI $[-0.01, -0.00]$) (note that this effect represents the additive effect of increasing network size by one). These results suggest that network structure and size are important characteristics to determine the adoption of shared beliefs, and replicates previous findings from behavioral economics and the computational social sciences (Golub & Jackson, 2010; Centola & Baronchelli, 2015).

Network topology affects entropy dynamics of response distribution Responses from all players in a network produce a distribution of generated hashtags on each trial. The variability of hashtags over time captures the degree of coherence of individual responses across all participants in a network. Hence, entropy is a concise measure of response variability across a network (Avolio et al., 2019; Hallett et al., 2016); the lower the entropy, the more similar the set of responses from all players. We computed the change in entropy of the response distributions across each network run, and fit a Gaussian generalized linear model to predict a network's entropy as a function of time and structure (and their interaction), while controlling for a network's size.

As shown in Fig. 4, the GLM model shows that the starting entropy is similar across network structures ($\beta_0 = 2.59$, 95% CI $[2.47, 2.71]$). Furthermore, the entropy of hashtag responses decreases over time ($\beta_{Trial} = 0.04$, 95% CI $[-0.04, -0.03]$), however entropy decreases at a faster rate in homogeneously-mixed networks than spatially-embedded networks ($\beta_{Spatial} = 0.01$, 95% CI $[0.01, 0.02]$). Because larger networks have more participants contributing responses, we also found a positive effect of network size on entropy ($\beta_{Size} = 0.01$, 95% CI $[0.01, 0.01]$). While a sin-

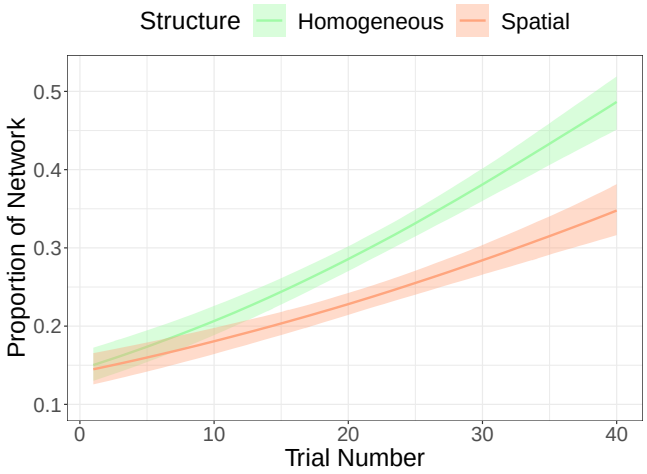


Figure 3: **Proportion of participants generating a dominant hashtag as a function of trials across network structures.** Curves represent marginal effects of the linear model, and error bars represent 95% Bayesian credible intervals. Homogeneously-mixed networks show faster emergence of dominant responses due to information aggregation across network connections.

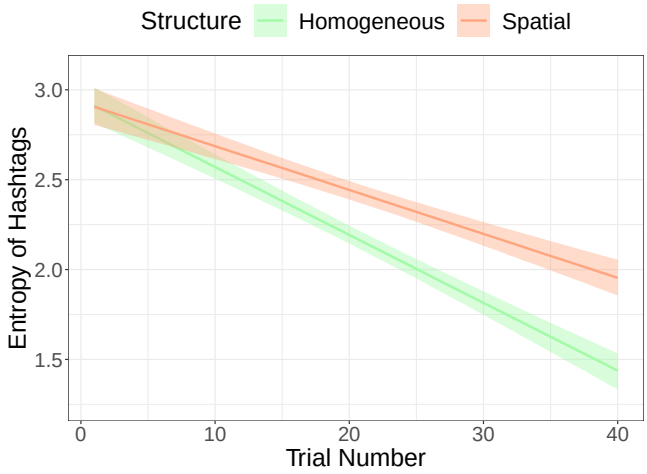


Figure 4: **Entropy of hashtag responses decrease over time across network structures.** Lower values imply greater coherence of responses across participants because generated hashtags are more similar. However, entropy of responses decrease more rapidly in homogeneously-mixed networks due wider communication of information.

gle dominant response may struggle to emerge in spatially-embedded networks, responses are still cohering within local neighborhoods, which in turn decreases the entropy (less variability) of hashtag responses, but at a slower rate than homogeneously-mixed networks. Because separate local-neighborhood groups can align on different hashtags in spatially-embedded networks, this finding is consistent with echochambers found in online networks. We discuss this idea in more detail at the end of this section (see Fig. 6).

Local connections support coordination within neighborhoods Connections across a network support the emergence of a dominant belief. However, increasing the connectivity of a node decreases the amount of times that any given pair of players can coordinate. In the spatially-embedded, partner players coordinate ten times across forty trials, regardless of total network size; whereas in homogeneously-mixed networks repeated interactions significantly decrease as a function of network size (i.e., number of players in the network). To assess how coordination dynamics varied across network structures, we fit a Bernoulli generalized linear model to predict if a pair of participants coordinated on trial t interacted with network structure over time (controlling for the size of the network).

As shown in Fig. 5, participants in both neighborhood structures learned to coordinate with their networked neighbors as trials progressed ($\beta_{Trial} = .04$, 95% CI [0.00, 0.04]), however those in spatially-embedded networks coordinated more effectively than they did in homogeneously-mixed networks ($\beta_{Spatial} = .02$, 95% CI [0.01, 0.03]).

Local coordination can result in separable neighborhoods coordinating on different hashtags. Fig. 6 depicts the full array of responses for a single run of both $N = 20$ network structures. Different local groups align on different hashtags in the spatially-embedded network. For example, 16-18 adopted #NuclearDisaster, while 2-6 and 8-14 aligned on #Nuclear; and nodes 19, 20, and 1 aligned on #Setsuden. The emergence of separate, localized groups coordinating on different hashtags likely hinders a dominant response from emerging in these networks, and limits the decrease of entropy as shown in the earlier analysis of global network coherence. The other $N = 20$ runs and network sizes $N = 50, 100$ display similar results. Participants in localized clusters received high rewards for coordinating within their partners in the local clusters (as indicated by the node size in the networks to the left of the color maps). Although participants were not coordinating as effectively in homogeneously-mixed networks, network topology still supported the emergence of a dominant response globally.

This result suggests that a latent form of information aggregation leads to the emergence of dominant hashtag, rather than being directly due to participants coordinating effectively in the local neighborhood (Golub & Jackson, 2010). Furthermore, note how in Fig. 6, participants 7 (top) and 17 (bottom) both received zero reward for coordinating, and continually generated new hashtags over course of the experiment. Participants are rewarded by adopting shared behaviors that encourage a shift towards consensus, and those that don't learn to exploit environmental regularities are not rewarded.

Probing causal representation change following networked interactions Before and after networked interaction, participants wrote "tweets" describing the narrative.

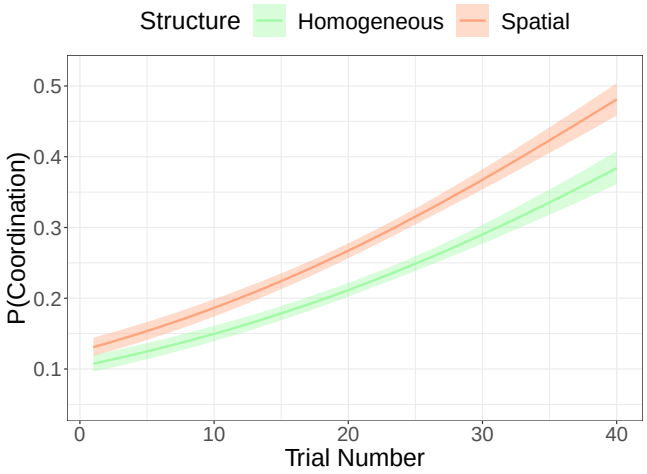


Figure 5: **Emergence of coordinated behavior across network structures.**

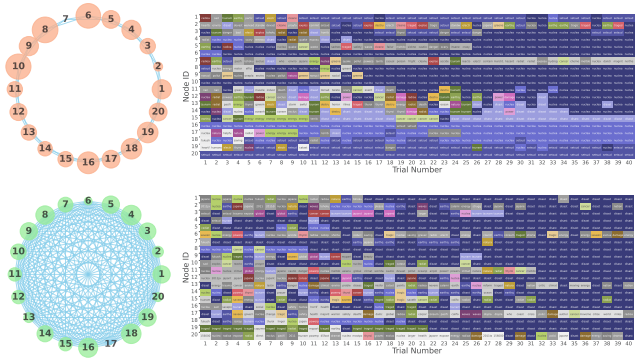


Figure 6: **Rewards and colormaps of hashtag responses across a single $N = 20$ run.** Top panel shows results for a spatially-embedded network, the bottom panel is from a homogeneously-mixed network. **Left:** Network structure with player nodes sized by participants' final rewards for coordinating (range (top) 0-25; (bottom) 0-19). **Right:** Colormap of individual responses, rows represent individual participants' set of responses, columns represent trials. Cells encode the first five letters of the generated hashtag.

Natural language processing (NLP) methods applied to these documents can illuminate which parts of the narrative people focused on when describing the events in a brief format such as tweets. Because causal relations are central to narrative representation (Zwaan, Magliano, & Graesser, 1995; Morrow, Bower, & Greenspan, 1989), we analyzed the causal claims that participants made in the tweet documents.

Priniski et al. (2023) developed a Large Language Model (LLM) system that identifies and extracts causal claims expressed in text documents. The model identifies spans of words serving as inputs to explicitly stated causal relations. Both the cause and effect events, and the underlying causal relation (i.e., a causal trigger) must be explicitly stated for the algorithm to identify the causal claim. The extracted claims are then co-referenced based on BERT embeddings of the identified entities (Devlin, Chang, Lee, & Toutanova, 2019), to produce clusters of semantically similar topics, termed "causal topics". The model additionally encodes the direction of the stated causal relationships linking any two clusters.

Networked interaction shifts causal language expressed in personal narratives To assess if participants expressed more causal language following networked interaction, we fit a hurdle Poisson model to predict the number of causal claims a participant produced at a given phase of interaction. The hurdle Poisson model consists of a logistic classification step for identifying tweets with no identified causal claims, and then a Poisson distribution estimates the counts for remaining documents. Because some participants may be more prone to generate causal reports than others, we fit the model with a random intercept for subject. Effects are to be interpreted as cumulative log odds which describes the expected number of claims in each interaction phase and network condition.

The hurdle parameter predicts that around 54% of documents contained zero causal claims ($hu = .54$, 95% CI [0.51, 0.57]). The intercept equals the expected log count of causal claims before networked interaction in a homogeneously-connected network ($\beta_0 = .19$, 95% CI [-0.04, 0.39]), which equates to a mean of 1.19 claims. Interestingly, after networked interaction, the expected count increases to around 1.61, which is more than pre-interaction counts ($\beta_0 = .30$, 95% CI [0.08, 0.52]). Participants in spatially-embedded networks produced slightly fewer claims than those in homogeneously-mixed networks, however there was not a credible difference ($\beta_{Spatial} = -.12$, 95% CI [-0.34, 0.09]).

Finer-grained analyses of the content of participants' causal claims can reveal what causal topics are most salient in the narrative, and how network structure may shift an individual's representation of the narrative's causal content. We compared differences in claims made after and before interaction to highlight causal relations that may have been enhanced by interaction. As shown in Fig. 7, the model identified fourteen causal clusters, plus a non-topic category of claims that couldn't be clustered based on their embeddings. All causal events expressed in the narrative appear as clusters in the corpus (for reference see Fig. 1), reflecting a tendency to use causal claims when summarizing events, even in short-formatted messages such as tweets. Because causal relations are directional (causes produce effects), Fig. 7 shows the direction of the relationships linking any two topics. Rows denote clusters used as a cause and columns denote clusters used as an effect. Cell values are the averaged differences across networked conditions in the number of claims expressing that causal relation after interaction relative to before (i.e, post - pre-interaction). Due to the global aggregation of information in homogeneously-mixed networks, participants in these networks generate tweet messages including a smaller set of causal relations centering around the

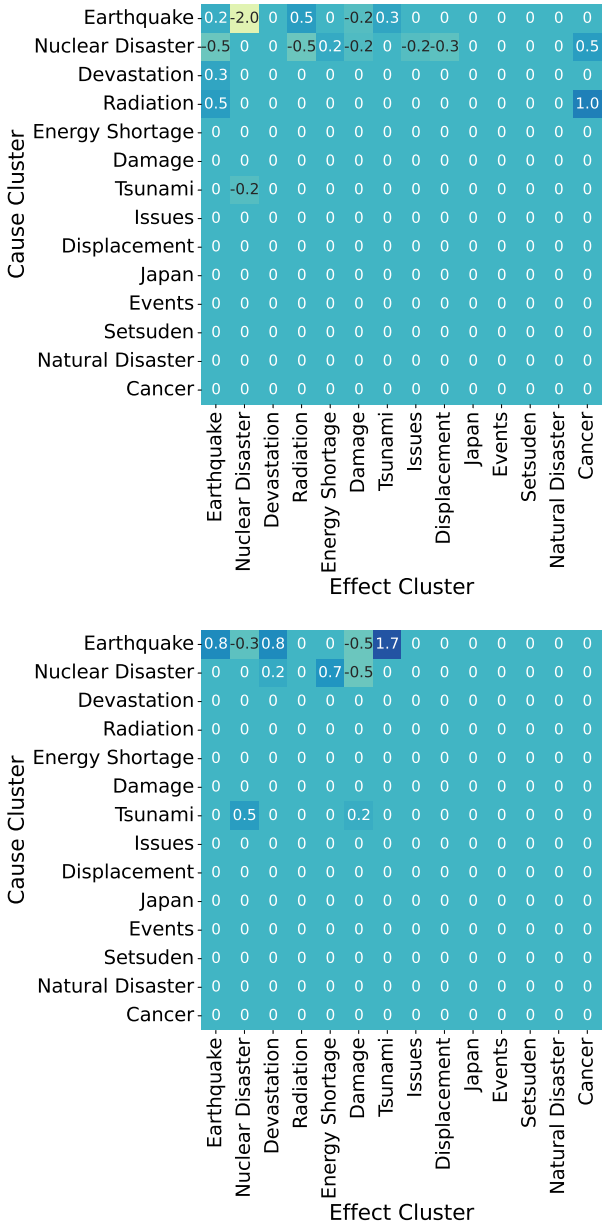


Figure 7: **Shift in causal language following networked interaction. Top:** spatially-embedded networks; **bottom:** homogeneously-mixed networks. Cells represent average difference scores of claims instantiating each causal topic across network structures. Cell i, j represents documents claiming that topic i caused j. Positive values indicate more documents expressed that causal relationship after interaction, negative values indicate less causal claims in the tweet documents after network interaction.

initial causal chain in the narrative. For example, for the groups with homogeneously-mixed interactions, the causal link *earthquake* $\rightarrow$ *tsunami* had increased by 1.7 documents after interaction, and the subsequent causal link *tsunami* $\rightarrow$ *nuclear disaster* had increased by .5. However, the enhancement of this causal chain is much weaker for the groups with spatially-embedded interactions; the difference scores are .3 and $-.2$, respectively.

Discussion

We examined how the similarity between a full set of hashtags across a network, and the similarity between pairs of coordinating nodes is sensitive to a network's neighborhood topology. We replicated the finding that global connections facilitate the emergence of dominant behaviors, while repeated interactions within localized neighborhoods promotes higher rewards based on coordination, with separable communities producing different hashtag responses.

These findings have direct implications for understanding the onset of echochambers and belief polarization in online social networks. High reward values in the spatial networks proxy social reinforcement for stating one's beliefs online. People will increase their confidence in their stated beliefs when those in their neighborhood reaffirm their responses, despite different beliefs reported from others in different communities. After separate clusters of people begin coordinating on different responses, it will become more difficult to find common ground once beliefs are solidified with one's groups (Kahan et al., 2012). Our results suggest that one way to increase global agreement within a online network is to structure neighborhoods as to encourage interactions across a wider-array of nodes in the network. Even if participants don't directly coordinate their beliefs with their direct contacts, access to this information could have an aggregate impact on the global consensus of the network.

In addition to analyzing hashtag reporting during networked interaction, we measured shifts in causal language use before and after networked interaction. We found that more causal topics emerged after network interactions, especially more in homogeneous structure than in spatial structure. These results suggest that network interactions have the potential to deepen people's understanding of causality in complex events. Future NLP analyses of "tweets" should parse a wider-array of semantic relations to model shifts in *situation models*, the memory representations people build when processing text-based narratives (Morrow et al., 1989; Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). This effort could elucidate how networked interactions impact people's narrative representations, and illuminate mechanisms for encouraging healthier discourse online.

Acknowledgements

This work was funded in part by the AFOSR MURI grant No. FA9550-22-1-0380 and the DARPA Army Research Office (ARO), under Contract No. W911NF-21-C-0002. We

thank Yiling Yun, Yiting Wang, and Chole Ji for their help in running the experiments.

References

Anagnostopoulos, A., Becchetti, L., Castillo, C., Gionis, A., & Leonardi, S. (2012). Online team formation in social networks. In *Proceedings of the 21st international conference on world wide web* (pp. 839–848).

Avolio, M. L., Carroll, I. T., Collins, S. L., Houseman, G. R., Hallett, L. M., Isbell, F., ... Wilcox, K. R. (2019). A comprehensive approach to analyzing community dynamics using rank abundance curves. *Ecosphere*, *10*(10), e02881.

Booten, K. (2016). Hashtag drift: Tracing the evolving uses of political hashtags over time. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2401–2405). ACM.

Boyd, D., Golder, S., & Lotan, G. (2010, January). Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In *2010 43rd Hawaii International Conference on System Sciences* (pp. 1–10).

Centola, D., & Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, *112*(7), 1989–1994.

Chen, D. L., Schonger, M., & Wickens, C. (2016, March). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Cunha, E., Magno, G., Comarela, G., Almeida, V., Gonçalves, M. A., & Benevenuto, F. (2011, June). Analyzing the Dynamic Evolution of Hashtags on Twitter: a Language-Based Approach. In M. Nagarajan & M. Gamon (Eds.), *Proceedings of the Workshop on Language in Social Media (LSM 2011)* (pp. 58–65). Portland, Oregon: Association for Computational Linguistics.

Dawson, P. (2020). Hashtag narrative: Emergent storytelling and affective publics in the digital age. *International Journal of Cultural Studies*, *23*(6), 968–983.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019, May). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.* arXiv. (arXiv:1810.04805 [cs])

Giaxoglou, K. (2018). #jesuischarlie? hashtags as narrative resources in contexts of ecstatic sharing. *Discourse, context & media*, *22*, 13–20.

Golub, B., & Jackson, M. O. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, *2*(1), 112–149.

Hallett, L. M., Jones, S. K., MacDonald, A. A. M., Jones, M. B., Flynn, D. F. B., Ripplinger, J., ... Collins, S. L. (2016). codyn: An r package of community dynamics metrics. *Methods in Ecology and Evolution*, *7*(10), 1146–1151.

Howard, P. N., & Hussain, M. M. (2013). *Democracy's Fourth Wave?: Digital Media and the Arab Spring*. Oxford University Press.

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012, October). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732–735.

Lin, Y.-R., Margolin, D., Keegan, B., Baronchelli, A., & Lazer, D. (2013). # bigbirds never die: Understanding social dynamics of emergent hashtags. In *Proceedings of the international aaai conference on web and social media* (Vol. 7, pp. 370–379).

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. New York: Chapman and Hall/CRC. doi: 10.1201/9781315372495

Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of memory and language*, *28*(3), 292–312.

Papacharissi, Z. (2015). *Affective publics: sentiment, technology, and politics*. New York, NY: Oxford University Press.

Papacharissi, Z. (2016). Affective publics and structures of storytelling: sentiment, events and mediality. *Information, Communication & Society*, *19*(3), 307–324.

Priniski, J., Verma, I., & Morstatter, F. (2023). Pipeline for modeling causal beliefs from natural language. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)* (pp. 436–443). Toronto, Canada: Association for Computational Linguistics.

Sasahara, K., Chen, W., Peng, H., Ciampaglia, G. L., Flammini, A., & Menczer, F. (2021). Social influence and unfollowing accelerate the emergence of echo chambers. *Journal of Computational Social Science*, *4*(1), 381–402.

Tangherlini, T. R., Shahsavari, S., Shahbazi, B., Ebrahimzadeh, E., & Roychowdhury, V. (2020, June). An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, Pizzagate and storytelling on the web. *PLOS ONE*, *15*(6), e0233879.

Yang, G. (2016). Narrative Agency in Hashtag Activism: The Case of #BlackLivesMatter. *Media and Communication*, *4*(4), 13–17.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological science*, *6*(5), 292–297.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of experimental psychology: Learning, memory, and cognition*, *21*(2), 386.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, *123*(2), 162.