

A cross-hazard analysis of terse message retransmission on Twitter

Jeannette Sutton^a, C. Ben Gibson^b, Nolan Edward Phillips^b, Emma S. Spiro^c, Cedar League^d, Britta Johnson^d, Sean M. Fitzhugh^b, and Carter T. Butts^{b,e,f,g,1}

^aDepartment of Communication, University of Kentucky, Lexington, KY 40506; ^bDepartment of Sociology, University of California, Irvine, CA 92697; ^cInformation School, University of Washington, Seattle, WA 98195; ^dTrauma, Health and Hazards Center, University of Colorado, Colorado Springs, CO 80918; ^eDepartment of Statistics, University of California, Irvine, CA 92697; ^fDepartment of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697; and ^gInstitute for Mathematical and Behavioral Sciences, University of California, Irvine, CA 92697

Edited by Roger E. Kasperson, Clark University, Worcester, MA, and approved September 30, 2015 (received for review May 12, 2015)

For decades, public warning messages have been relayed via broadcast information channels, including radio and television; more recently, risk communication channels have expanded to include social media sites, where messages can be easily amplified by user retransmission. This research examines the factors that predict the extent of retransmission for official hazard communications disseminated via Twitter. Using data from events involving five different hazards, we identify three types of attributes—local network properties, message content, and message style—that jointly amplify and/or attenuate the retransmission of official communications under imminent threat. We find that the use of an agreed-upon hashtag and the number of users following an official account positively influence message retransmission, as does message content describing hazard impacts or emphasizing cohesion among users. By contrast, messages directed at individuals, expressing gratitude, or including a URL were less widely disseminated than similar messages without these features. Our findings suggest that some measures commonly taken to convey additional information to the public (e.g., URL inclusion) may come at a cost in terms of message amplification; on the other hand, some types of content not traditionally emphasized in guidance on hazard communication may enhance retransmission rates.

disaster | warning | social media | retransmission | communication

Under conditions of imminent threat, rapid communication of warning information to the public is a primary strategy for decreasing loss of life and increasing public safety by eliciting protective actions from those at risk (1). For decades, public warnings have been relayed via mass media channels, including radio, broadcast television, and sirens (2). With the advent of social computing, warnings have begun to be disseminated via online social networks (OSNs), where messages can be more easily propagated and amplified by the user population (3–6). Risk amplification via message retransmission in this setting is important because it enables a message to reach individuals beyond the sender's direct contacts, increasing exposure and potentially leading to lifesaving actions (7). Although such transmission occurs offline as well (8–11), OSNs offer the potential for the rapid retransmission of short messages with higher fidelity—and to more persons—than would typically be feasible via other means.

In addition to enabling message diffusion, the clustered structure of most OSNs (12, 13) allows retransmission to expose individuals to the same message multiple times. Multiple exposures to messages have been linked to greater confidence in message veracity (14, 15), which can lead to further sharing (16, 17). Repeated exposures from multiple network ties are often a prerequisite for the spread of information through networks, and are of particular importance for inducing behavioral change (4, 18–20). Under conditions of imminent threat, exposure to a warning message from a trusted source (such as a neighbor, friend, or family member) strongly affects one's willingness to take protective actions (21, 22). This highlights the need to understand the factors that enhance or suppress the amplification of emergent risk messages

within OSNs, with particular attention on the features of the messages themselves. Such an understanding can inform evidence-based strategies to increase message proliferation, thus allowing risk communicators to achieve a higher level of message penetration and/or to increase the number of exposures per person in the impacted population. This research examines message retransmission—commonly referred to as “retweeting”—on Twitter, identifying network, content, and style features that promote the amplification (23) of terse (i.e., content-constrained) messages within five types of hazard events.

Network Features

Considerable attention has been given to the role of traditional media in amplifying risk; until recently, few studies have focused on the online environment, and even fewer have studied the amplification process in real time (5). Past OSN propagation models have focused on local network structures that influence message amplification (24–28) in disaster and nondisaster settings alike. This body of work demonstrates the importance of local network characteristics for propagation. For example, Tyshchuk et al. (27) found that official emergency management accounts quickly became leaders of information diffusion and maintained that position throughout the period of imminent threat during a hurricane. Importantly, the network structures can create “information cascades” where individuals amplify information received by retweeting, and thus further sharing, the information (29). Moreover, these cascades tend to be wide (i.e., composed of a large number of short-length passing chains)

Significance

Online social networks (OSNs) enable time-resolved measurement of communication behavior during disasters, making it possible to probe the mechanisms by which messages are amplified or suppressed with precision unattainable by traditional data sources. To our knowledge, this research provides the first systematic study of the factors predicting the social amplification of risk communication in OSNs by examining the retransmission of official messages across five hazards. Our findings demonstrate the respective impacts of sender characteristics, message content, and message style in determining whether an official message will be passed on during an emergency, as well whether these vary across hazards. These results contribute to the evidence base for policies guiding the delivery by emergency management organizations of lifesaving information to the public.

Author contributions: J.S. and C.T.B. designed research; J.S., C.L., B.J., and S.M.F. performed research; S.M.F. maintained infrastructure for data collection; J.S., C.B.G., N.E.P., and E.S.S. analyzed data; and J.S., C.B.G., N.E.P., and C.T.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: butts@uci.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508916112/-DCSupplemental.

rather than long (i.e., composed of a small number of long passing chains), which highlights the importance of initial message exposure for subsequent amplification.

Message Content

An emerging line of research mines the textual content that users generate for online media (27, 28). This development has enabled researchers to examine the message content features associated with higher rates of message retransmission (30, 31). Researchers note the emergence of user conventions when retweeting, such as the inclusion of a RT (indicating a retransmitted message or retweet), # (hashtag), and web link or URL (32). These style conventions are related to retransmission: for instance, Starbird and Palen (33) and Bruns and Burgess (34) tracked retweeted messages across events to identify the content and source of individual messages and found that hashtags are used to form ad hoc discussion channels around a particular event. Additionally, researchers have suggested that including a web link to additional information for terse messaging channels, such as Twitter, can increase message clarity because length-constrained messages may not be able to convey all pertinent information (35).

Under quotidian conditions, the inclusion of URLs and hashtags in a message strongly influences the likelihood that the message is retweeted (4, 31); however, researchers have not found this to be true during disaster settings (6, 36, 37). Collectively, they find that URL inclusion significantly reduces retweet counts, whereas the inclusion of a hashtag significantly increases retweet counts during a wildfire (6); Burnap et al. (36) observed similar results for message retransmission after a terrorist attack in the United Kingdom; and Genes et al. (37) noted that the messages most retweeted during a blizzard or hurricane do not include a URL. As discussed below, our results shed light on these discrepancies.

Message Style

Although previous studies of public message retransmission have examined message content by focusing on trending topics (31), performing sentiment analysis (30), or analyzing word choice through its frequency of use (25, 26), these approaches do not include message style features—that is, how the message is written to improve message clarity. Rhetorical strategies—such as the use of punctuation, the choice of sentence style, or the capitalization of words or phrases—are additional message features that may affect perceptions of message salience among recipients and subsequently influence their likelihood to retransmit a message. For instance, warning messages with clear and specific information (2) or that include instructional information (37) are found to increase message understanding, leading to increased likelihood that message receivers will take protective actions. Similarly, imperative sentences are directive and can instruct message receivers (38). Thus far, analyses on the use of “ALL CAPS” have centered on its function to convey sentiment (39) or as an “intensifier” of information (40); capitalization may also serve to bring attention to key warning elements (41).

In this research, we link these three factors—network properties, message content, and message style—in a combined model of message retransmission. By analyzing these factors’ impacts across multiple hazard events, we are able to determine the influence of each on message retransmission net of other factors, and to determine the degree to which these effects are similar, or different, across hazard types. This is a critical first step in generating a normative template for message amplification during hazards, estimating the distinct impact of different message and contextual features on the flow of hazard information through OSNs, and modeling the amplification of risk through message retransmission during imminent threat events.

Materials and Methods

Design, Sample, and Data Collection. We manually coded and analyzed Twitter messages sent from official accounts during five different hazard events, with the goal of identifying features that consistently predict increased or decreased message retransmission under conditions of imminent threat. All data were

collected via the Twitter Application Programming Interface (API), and all messages were publicly available; as no human subject research was performed, IRB approval was not required. We selected five events that represent five different hazard types: terrorist attack (TA), wildfire (WF), blizzard (BL), hurricane (H), and flash flood (FL) with differing population sizes and regions impacted. For each event, the time period of imminent threat and warning varied based upon environmental conditions or containment of the hazard agent. These ranged from 48 hours (H) to 5 days in duration (TA). Table 1 below includes information on each event; additional information for each event can be found in Table S1.

Following previous research (6, 23, 42), we identified and enumerated Twitter accounts for each of our five events, obtaining a census of the population of public officials or agencies at the local, state, and federal level posting in response to the respective event ($n = 188$). Using the Twitter REST API, we collected each account’s posting history over the time period of the hazard event. This includes all public-facing messages posted during the period of imminent threat for each event; for each message, we observed the time of posting and message content. We also obtained the number of “Followers” for each account and the number of times that each message was retweeted: we updated this information daily over the period of observation.

For four cases, we identified the accounts that were posting tweets on the primary event hashtags. (In one case, TA, we were already collecting data from the apropos accounts because we had targeted them for the BL, which occurred in the same geographical area.) Subsequently, we scanned each organization’s connections and local websites to identify additional relevant accounts and added them to our list of targeted accounts. Although an account may exist that was not discovered by our collection procedure, it is highly unlikely to have been central in the response because such an account did not post any messages with an event hashtag and was not referenced by an active organization. We therefore regarded the accounts used for each case as an effective census of the population of official accounts responding to the event in question. We focused on official emergency management accounts, both because these accounts are central information diffusers during disasters (27) and because the retweeting of messages from official accounts provides a direct probe on the amplification of official messages by members of the public (6, 23, 38).

Content Coding. We used a deductive coding approach developed in prior research (23, 41) to code each tweet for message content and style. All data are available in Dataset S1. We coded for the presence of 11 primary content themes: *advisory*, i.e., actions to take or refrain from; *closures/openings* (*close/open*) of facilities, events, or roads; *corrections* to previously posted information; *evacuation/shelter* (*evac./shelter*) information; information on hazard impact (*haz. Impact*) to the environment or population at risk; *general information*, including phone numbers, websites, or help facility locations; information on *reentry* to the affected area; *thanks and prayers* (*thanks*); information on *volunteering* (*volunteer*); *direct responses to requests for help* (*help/dir.*); and *emotive/evaluative* (*emot./eval.*) content, e.g., cohesion-building statements such as “we will recover.” Two additional codes were used for messages with content that was considered undeterminable (*unsure*) or irrelevant to the hazard event (*not on topic*), and these messages were removed from the dataset. Theme code definitions with examples are available in Table S2. This approach enabled us to focus on what the general public views as salient based on what messages content they choose to retransmit and hence amplify.

To assess message style (clarity and specificity), we coded each message for sentence style (declarative, imperative, interrogative, or exclamatory) as well as the inclusion of a word or phrase in ALL CAPS to specify a category signifier (e.g., UPDATE) or to emphasize a word or phrase within the message (e.g., NOW or DO NOT). Additionally, we coded each message’s style features including whether the tweet was directed at or responding to another Twitter user, or referencing a third party (i.e., @name), whether it was a retweet (RT), and the inclusion of a hashtag or URL. Content categories and style features are not mutually exclusive (i.e., multiple themes and sentence styles might be included in a single message—see Table S3 for these counts). Fig. 1 below shows the proportion of themes present in the messages (see Fig. S1 for similar depictions of message style). Finally, we measured message exposure as the number of Followers for each targeted account.

Analytic Strategy. For each disaster event, we used a negative binomial regression model for retransmission activity (total retweets per initial message), whose expectation is a function of network features, message style, and message content. This form accounts for unobserved heterogeneity in the underlying retransmission process (manifesting empirically as overdispersion relative to a Poisson process); Poisson and geometric models were also considered but were found to produce inferior models as selected by the corrected Akaike information criterion (AICc). We used the generalized linear

competing for attention in a noisy environment. Thus, organizations should use empirically informed strategies to amplify their warning messages. Although design of risk messages is currently governed primarily by “best practices” (44, 45), evidence-based message design has the potential to substantially improve communication to the public. Future preventive warning policy for risk communicators of all types who use terse messaging channels should include the design and adoption of message templates that include highly salient content areas and the use of effective stylistic features, while avoiding features, such as URLs, that reduce retransmission. In the absence of a URL, however, additional clarifying information should be made available through alternative sources. The URL finding also highlights the fact that emergency management organizations may have different goals in providing relevant information to the public. If emergency managers want to disseminate detailed information to the public on Twitter, then the inclusion of a URL to that information is advisable; however, emergency managers should be aware that this could attenuate the amplification of that message. Therefore, emergency

managers should use preestablished hashtags, capitalize critical information, and use imperative sentences to increase message amplification.

Our study also shows differences in retransmission patterns across hazards, suggesting the value of a quantitative theory relating hazard characteristics to the social amplification of risk. The development of a database covering a wider range of hazard cases would greatly contribute to this effort and further validate existing frameworks. A systematic theory of cross-hazard message amplification would provide a foundation for the development of “normative principles” that exploit common retransmission patterns across hazard types to enhance the dissemination of official messages. By focusing on message content and style along with network attributes, we are closer to generating templates that can maximize message amplification during imminent threats.

ACKNOWLEDGMENTS. This material is based on research supported by Army Research Office Award ARO-W911NF-14-1-0552 as well as National Science Foundation Awards CMMI-1031853, CMMI-1031779, IIS-1251267, and OIA-1028394.

- Atman CJ, Bostrom A, Fischhoff B, Morgan MG (1994) Designing risk communications: Completing and correcting mental models of hazardous processes, Part I. *Risk Anal* 14(5):779–788.
- Mileti DS, Sorensen JH (1990) Communication of emergency public warnings. *Landslides* 1(6):52–70.
- Kasperson RE, et al. (1988) The social amplification of risk: A conceptual framework. *Risk Anal* 8(2):177–187.
- Suh B, Hong L, Piroli PL, Chi EH (2010) Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. *Second IEEE International Conference on Social Computing (SocialCom)* (IEEE Computer Society, Los Alamitos, CA), pp 177–184.
- Chung JI (2011) Social amplification of risk in the Internet environment. *Risk Anal* 31(12):1883–1896.
- Sutton J, et al. (2014) Warning tweets: Serial transmission of messages during the warning phase of a disaster event. *Inf Commun Soc* 17(6):765–787.
- Vieweg S, Hughes AL, Starbird K, Palen L (2010) Microblogging during two natural hazards events: What Twitter may contribute to situational awareness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, New York), pp 1079–1088.
- Scanlon TJ (1977) Post-disaster rumor chains: A case study. *Mass Emergencies* 2(126):22–27.
- Scanlon J (2007) Sampling an unknown universe: Problems of researching mass casualty incidents (a history of ECRU's field research). *Stat Med* 26(8):1812–1823.
- Erickson BH, Nosanchuk TA, Mostacci L, Dalrymple CF (1978) The flow of crisis information as a probe of work relations. *Can J Social* 3(1):71–87.
- Heath C, Heath D (2007) *Made to Stick: Why Some Ideas Survive and Others Die* (Random House, New York).
- Kumar R, Novak J, Tomkins A (2010) Structure and evolution of online social networks. *Link mining: Models, Algorithms, and Applications*, eds Yu PS, Han J, Faloutsos C (Springer, New York), pp 337–357.
- Kwak H, Lee C, Park H, Moon S (2010) What is Twitter, a social network or a news media? *Proceedings of the 19th International Conference on World Wide Web* (ACM, New York), pp 591–600.
- Arkes HR, Hackett C, Boehm L (1989) The generality of the relation between familiarity and judged validity. *J Behav Decis Making* 2(2):81–94.
- Hawkins SA, Hoch SJ, Meyers-Levy J (2001) Low-involvement learning: Repetition and coherence in familiarity and belief. *J Consum Psychol* 11(1):1–11.
- Rosnow RL, Yost JH, Esposito JL (1986) Belief in rumor and likelihood of rumor transmission. *Lang Commun* 6(3):189–194.
- Fragale AR, Heath C (2004) Evolving informational credentials: The (mis)attribution of believable facts to credible sources. *Pers Soc Psychol Bull* 30(2):225–236.
- Centola D, Macy M (2007) Complex contagions and the weakness of long ties. *AJS* 113(3):702–734.
- Centola D (2010) The spread of behavior in an online social network experiment. *Science* 329(5996):1194–1197.
- Starbird K, Palen L (2012) (How) will the revolution be retweeted?: Information diffusion and the 2011 Egyptian uprising. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (ACM, New York), pp 7–16.
- Peters RG, Covello VT, McCallum DB (1997) The determinants of trust and credibility in environmental risk communication: An empirical study. *Risk Anal* 17(1):43–54.
- Liu BF, Fraustino JD, Jin Y (January 13, 2015) Social media use during disasters: How information form and source influence intended behavioral responses. *Communic Res*, 10.1177/0093650214565917.
- Sutton J, et al. (2014) Terse message amplification in the Boston bombing response. *ISCRAM 2014 Conference Proceedings: 11th International Conference on Information Systems for Crisis Response and Management* (International Community on Information Systems for Crisis Response and Management, University Park, PA), pp 612–621.
- Lee K, Mahmud J, Chen J, Zhou M, Nichols J (2014) Who will retweet this?: Automatically identifying and engaging strangers on Twitter to spread information. *Proceedings of the 19th International Conference on Intelligent User Interfaces* (ACM, New York), pp 247–256.
- Petrovic S, Osborne M, Lavrenko V (2011) RT to win! Predicting message propagation in Twitter. *ICWSM 2011 Conference Proceedings* (Association for the Advancement of Artificial Intelligence, Palo Alto, CA), pp 586–589.
- Yang J, Counts S (2010) Predicting the speed, scale, and range of information diffusion in Twitter. *ICWSM 2010 Conference Proceedings* (Association for the Advancement of Artificial Intelligence, Palo Alto, CA), pp 355–358.
- Tyshchuk Y, Li H, Ji H, Wallace WA (2013) Evolution of communities on Twitter and the role of their leaders during emergencies. *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (ACM, New York), pp 727–733.
- Peng HK, Zhu J, Piao D, Yan R, Zhang Y (2011) Retweet modeling using conditional random fields. *2011 IEEE 11th International Conference on Data Mining Workshops (ICDMW)* (IEEE, Los Alamitos, CA), pp 336–343.
- Hui C, Tyshchuk Y, Wallace WA, Magdon-Smail M, Goldberg M (2012) Information cascades in social media in response to a crisis: A preliminary model and a case study. *Proceedings of the 21st International Conference Companion on World Wide Web* (ACM, New York), pp 653–656.
- Naveed N, Gottrott T, Kunegis J, Alhadi AC (2011) Bad news travel fast: A content-based analysis of interestingness on Twitter. *Proceedings of the 3rd International Web Science Conference* (ACM, New York), pp 8–14.
- Zhu J, Xiong F, Piao D, Liu Y, Zhang Y (2011) Statistically modeling the effectiveness of disaster information in social media. *2011 IEEE Global Humanitarian Technology Conference* (IEEE, Los Alamitos, CA), pp 431–436.
- Boyd D, Golder S, Lotan G (2010) Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. *2010 43rd Hawaii International Conference on System Sciences* (IEEE, Los Alamitos, CA), pp 1–10.
- Starbird K, Palen L (2010) Pass it on?: Retweeting in mass emergency. *Proceedings of the 7th International ISCRAM Conference* (International Community on Information Systems for Crisis Response and Management, Seattle, WA), pp 1–10.
- Bruns A, Burgess J (2012) Researching news discussion on Twitter: New methodologies. *Jpn Stud* 13(5-6):801–814.
- Stephens KK, Barrett AK (April 1, 2014) Communicating briefly technically. *IJBC*, 10.1177/2329488414525463.
- Burnap P, et al. (2014) Tweeting the terror: Modelling the social media reaction to the Woolwich terrorist attack. *Soc Netw Anal Min* 4(1):1–14.
- Genes N, Chary M, Chason K (2014) Analysis of Twitter users' sharing of official New York storm response messages. *Meld* 2 0 3(1):e1.
- Sutton J, League C, Sellnow TL, Sellnow DD (2015) Terse messaging and public health in the midst of natural disasters: The case of the Boulder floods. *Health Commun* 30(2):135–143.
- Kouloumpis E, Wilson T, Moore J (2011) Twitter sentiment analysis: The good the bad and the omg! *ICWSM 2011 Conference Proceedings* (Association for the Advancement of Artificial Intelligence, Palo Alto, CA), pp 538–541.
- Saif H, He Y, Alani H (2012) Semantic sentiment analysis of Twitter. *The Semantic Web—ISWC 2012* (Springer, Berlin), pp 508–524.
- Sutton J, et al. (2013) Tweeting the spill: Online informal communications, social networks, and conversational microstructures during the Deepwater Horizon oil spill. *Int J Inf Syst Crisis Response Manage* 5(1):58–76.
- Hughes AL, St Denis LA, Palen L, Anderson KM (2014) Online public communications by police and fire services during the 2012 Hurricane Sandy. *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems* (ACM, New York), pp 1505–1514.
- Sellnow TL, Sellnow DD, Lane DR, Littlefield RS (2012) The value of instructional communication in crisis situations: Restoring order to chaos. *Risk Anal* 32(4):633–643.
- US Department of Health and Human Services (2012) *CDC's Guide to Writing for Social Media* (US Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta).
- Veil SR, Buehner T, Palenchar MJ (2011) A work-in-process literature review: Incorporating social media in risk and crisis communication. *J Contingencies Crisis Manage* 19(2):110–122.