2018

Sentiment, richness, authority, and relevance model of information sharing during social Crises—the case of #MH370 tweets

Weiai Xu
University of Massachusetts Amherst

Congcong Zhang
University of Massachusetts Amherst

Recommended Citation

Xu, Weiai and Zhang, Congcong, "Sentiment, richness, authority, and relevance model of information sharing during social Crises—the case of #MH370 tweets" (2018). *Computers in Human Behavior*. 64.
https://doi.org/10.1016/j.chb.2018.07.041

Sentiment, Richness, Authority, and Relevance Model of Information Sharing during Social Crises—the Case of #MH370 Tweets

Abstract

The study introduces a model of crisis information sharing based on Twitter discussions of the missing Malaysian Airlines Flight 370. Grounded in the Elaboration Likelihood Model, the study tests four salient factors: *Sentiment*, *Richness*, *Authority*, and *Relevance*, which can be measured by peripheral cues in tweets and in user profiles. Findings suggest that information sharing is positively associated with the presence of peripheral cues indicative of a confident, self-revealing and positive emotional language style, and is negatively related to an angry and informal style. Additionally, information sharing is related to the presence of multimedia cues and cues indicating source popularity.

Keywords: Social media, crisis communication, information sharing, peripheral cues, diffusion, Twitter, ELM

1. Introduction

On March 8th, 2014, Malaysian Airlines Flight 370 (MH370) disappeared from radar. Subsequently, the most expensive multinational search in aviation history was launched for the missing plane. During the crisis, social media were teeming with public conversations, amidst which were speculations, rumors, conspiracy theories, and nationalist rhetoric. MH370 thus presents a useful case for examining the public use of social media in a crisis. A crisis creates informational and emotional gaps (Oh, Agrawal, & Rao, 2013; Shklovski, Palen, & Sutton, 2008). Social media step in to fill the gaps by providing an alternative outlet for crowd-sourcing and intelligence gathering (Kwon et al., 2016) as well as a safety valve for releasing collective anxiety and grief (Neubaum et al., 2014). Considering the volume and velocity of online conversations being produced, the public uses the sharing features (e.g., retweeting) to elevate certain content and agendas to prominence (Meraz & Papacharissi, 2013). Particularly interesting to researchers is the fact that only a fraction of content goes viral (Adamic & Huberman, 2000). The disproportional content virality begs the question: *what are the factors that drive the public to favor certain kinds of information over others?* Guided by this question, we consider predictive factors in crisis-time information sharing to be rooted in characteristics of messages and sources. To identify such factors, we first ground our model in dual process theories, arguing that information sharing is in part driven by the presence of persuasive cues in language, and such cues are reflected by word choices and user characteristics. We then present four testable prepositions to build a predictive model incorporating four factors: *Sentiment*, *Richness*, *Authority*, and *Relevance*. The literature review below provides explanations for why including the four factors in the context of crisis communication and how they can be measured using innovative computational methods.

2. Literature review

2.1. Twitter in crisis communication

Social media are characterized by crowd-based consumption and production of content. Their interactive and inclusive nature make them ideal for information dissemination and community mobilization during a crisis. A rich body of scholarship has examined the uses, effects, and pitfalls of social media in crisis communication (see Reuter, Hughes, & Kaufhold, 2018 for a review). Many have acknowledged unique affordances of Twitter, a popular micro-blogging platform. Tweets are spontaneously produced and shared instantly. Unlike public opinion collected through institutionalized and structured opinion polling, tweets are affect-laden, impromptu public discussions, revealing people's most intimate and instant thoughts concerning a crisis (Kwon et al., 2016). And unlike social networking sites (e.g., Facebook) built on strong-tie networks (i.e., family and friends), Twitter contacts are weak ties by shared interests (Takhteyev, Gruzd, & Wellman, 2012). This characteristic enables diverse, free-spirited, and transnational discussions of various public issues.

The literature has noted benefits and burdens of Twitter (and social media in general) in a crisis situation. The tool facilitates collective intelligence-sharing to build situational awareness and resilience (Neubaum et al., 2014; Stieglitz et al., 2018). The collective sharing serves as a bellwether for the concerns and needs of varying publics (Getchell & Sellnow, 2016). The tool also has emotional and therapeutic benefits. By sharing emotions, users feel part of a like-minded and empathetic community (Neubaum et al., 2014). Yet, social media could breed rumors and disinformation, at the times when there is a great amount of uncertainty and hostility (Kwon et al., 2016). The mass production of content on social media also creates an information glut that prevents important messages from gaining public attention (Rice & Spence, 2016).

Scholars are interested in studying kinds of viral messages during a crisis. Li, Vishwanath and Rao (2014) categorized social media posts sent during the Fukushima Nuclear Crisis to show the changing popularity of reassuring content throughout the course of a crisis. Boulianne, Minaker and Haney (2018) noted that expressions of care and concern were prominent in the viral content after wildfires. Rice and Spence (2016) instead focused on the proportion of instructional content in social media content. Among a growing number of studies that seek to build integrated message diffusion models, key factors predicting viral diffusion are identified, including information ambiguity, source credibility, social influence, anxiety, issue involvement, emotionality, and medium-specific features, namely hashtags and URLs (Chua, Tee, Pang, & Lim, 2016; Oh, Agrawal, & Rao, 2013; Liu, Burton-Jones, & Xu, 2014). However, the choice of factors lacks a theoretical guidance (Shi, Hu, Lai, & Chen, 2018). The current study thus attempts to build a theoretically grounded model.

2.2. Linking social media sharing to information processing

Social media sharing is an outcome of persuasion in that public attention has become a scarce resource necessary for attitude and behavioral changes in the digital era (Davenport & Beck, 2001). Sharing indicates that the audience attends to the message and deems it important (Li, Vishwanath, A., & Rao, 2014). Sharing is also a networked gatekeeping process to collectively frame issues and set agendas (Meraz & Papacharissi, 2013). In recognizing the link of information sharing to persuasion, we see sharing as driven by audience's processing and evaluation of social media posts. Evidence suggests that people process information in a dual process fashion (Teng, Wei Khong, Wei Goh, & Yee Loong Chong, 2014), leading some scholars to use dual process theories to study message diffusion during crises (Liu, Burton-Jones, & Xu, 2014).

Dual process theories, such as Petty and Cacioppo's (1986) Elaboration Likelihood Model (ELM), suggest two processing modes. The first mode takes a *central route*, marked by an effortful, logical and rational assessment of a message's merits. The second process takes a *peripheral route*, referring to an automatic, unconscious and fast way of thinking. In peripheral processing, people use cognitive shortcuts in messages to form attitudes. ELM posits that people are motivated to do peripheral processing when they are overloaded with information and distracted by tasks. This aptly describes the situation facing internet users. Online messages (such as tweets) are brief yet overwhelming, considering the speed at which they are being produced and pushed to users. The online audience is unlikely to scrutinize every detail. In that sense, they more likely rely on peripheral cues.

ELM also aptly describes information processing during a crisis, as the public experiencing a crisis would attend to peripheral cues to fulfill their informational and emotional needs. Earlier ELM studies revealed a set of peripheral cues including sentiment, source credibility/attractiveness, visual prominence, personal relevance, and complexity of information (Petty & Cacioppo, 1986). These studies show that an audience is more receptive to a message when it arouses the audience's affective states, comes from a credible or attractive source, has appealing visual elements, contains more and complex arguments, and when it is deemed personally relevant. These types of peripheral cues are abundant in online discussions of social crises. The unpredictability of events, coupled with grave and often fatal impacts, social crises typically arouse collective anxiety, and information dearth as well as information overload (Allport & Postman, 1947). A review of the literature shows that the public is sensitive to message and source-related characteristics indicative of anxiety, ambiguity, media richness, source credibility, attractiveness and personal and social relevance (Oh, Agrawal, & Rao, 2013;

Oh et al., 2018; Liu, Burton-Jones, & Xu, 2014). The current study synthesizes the factors into four categories: *Sentiment*, *Richness*, *Authority*, and *Relevance*.

2.3. Peripheral cues and information sharing

2.3.1 Sentiment

Sentiment entails emotions that have arousing effects which drive attention and attitude change (Petty & Cacioppo, 1986). Cues indicative of negative thoughts, such as fear, anxiety, and sadness, are found highly persuasive (Nabi, 2002), particularly in message diffusion during a crisis (Liu, Burton-Jones, & Xu, 2014; Zhang, Xu, & Zhang, 2017). A social crisis naturally evokes negative sentiments as the public mourns lost lives and as people grow anxious about the uncertain future. Blaming and divisive language is commonplace when people perceive out-group members at fault (Kwon et al., 2016). After the MH370 incident, the Malaysian government was widely criticized, notably by Chinese nationals, for its ill-prepared handling of the crash. Social media, as an unfiltered outlet for emotional releases and an echo chamber for likeminded people, will see a deluge of negative sentiments during crises. Studies show that viral content is characterized by a high degree of emotionality (Stieglitz & Dang-Xuan, 2012; Veltri & Atanasova, 2015; Zhang, Xu, & Zhang, 2017). During a crisis, anxiety and fear are dominant sentiments driving virality (Chen & Sakamoto, 2013; Oh, Agrawal, & Rao, 2013; Li, Vishwanath, & Rao, 2014). The shared negativity can help the public build rapport and rationalize their experience (Liu, Burton-Jones, & Xu, 2014). Social media also offer a place to seek assurance and empathy in the time of uncertainty (Neubaum et al., 2014). Thus, positive content such as those showing praise and appreciation may resonate with the public as well (Li, Vishwanath, & Rao, 2014; Pang & Ng, 2016). In synthesis, we propose the following to test the role of sentiment in social media sharing.

Proposition 1: Message cues indicative of emotionality (both negative and positive) predict information sharing.

2.3.2. Richness

In the current context, *Richness* is an umbrella term to include a set of message cues reflecting whether information is adequate, specific and analytical. Notwithstanding this being a subjective quality, a couple of cues can be used to evaluate message/argument quality. One such peripheral cue is *the amount of information*, with the assumption that more information equates with better quality (Petty & Cacioppo, 1986). Social media messages are typically short, but additional information can be packed into multiple media elements such as video clips, images, or URLs. These cues enhance the "telepresence", "media richness" and "vividness" of a message in that they create more direct sensory experience (Liu, Ji, North, & Yang, 2017). Prior studies show that content with more multimedia cues predicts a higher chance of retweeting (Chung, 2017; Liu, et al., 2017).

Another peripheral cue for evaluating message quality is writing styles. Formal, logically ordered, specific, and consistent writing gives a convincing impression (Kaufman, Stasson, & Hart, 1999; Slater & Rouner, 1996). Thus, audiences can form impressions based on the structure and style of online content (Choi, & Stvilia, 2015; Rowley & Johnson, 2013). Because crises create informational gaps (Oh, Agrawal, & Rao, 2013), the public is compelled to seek clarity and certainty. In doing so, they are wired to avoid noisy and chaotic information (Allport & Postman, 1947), and become more receptive to information that is presented in a formal, consistent, and specific fashion. A past study, specific to the context of rumor retransmission, finds that ambiguous content leads to less message sharing (Liu, Burton-Jones, & Xu, 2014). Consider all aforementioned evidence, we expect a similar pattern in the current study.

Proposition 2: Message cues indicative of the richness of information positively predict information sharing.

2.3.3. Authority

The construct of *Authority* first refers to the existing influence of a message source. Audiences analyze source characteristics to infer whether a message is trustworthy (Sussman & Siegal, 2003). Traditionally, this kind of authority is defined by expert/non-expert status (Petty & Cacioppo, 1986). But the distinction is blurred in online communication where traditional opinion leaders may no longer sway votes in the virtual world (Xu, Sang, Blasiola, & Park, 2014). During crises, established and mainstream sources could be challenged by the skeptical public (Mills et al. 2009; Sutton et al. 2008). What matters then is whether the message comes from an online opinion leader. Online opinion leadership is indicated by having a large social following (Song, Dai, & Wang, 2016; Xu, Sang, Blasiola, & Park, 2014; Zhang, et al., 2014) and certain status symbols (Zhu, Yin, & He, 2014). A large follower count indicates source influence after social vetting and can be used to influence an audience's judgment of source credibility (Xu, 2013).

Authority is also reflected in certain language styles. Using dominant language conveys confidence and authority, leading the audience to form a preferred impression (Sparks & Areni, 2008; Toma & D'Angelo, 2015). Prior studies suggest that in online conversations on health and political topics, confident and assertive language is associated with a better diffusion outcome (Choi, 2014: Kim, Hou, Han, & Himelboim, 2016). Based on the above, the following is proposed:

Proposition 3: Cues indicative of authority positively predict information sharing.

2.3.4. Relevance

Relevance means whether the content contains personally relevant topics or is presented in a style relevant to the audience. This construct is derived from findings that suggest audiences are more likely to pay attention to personally relevant information (Petty & Cacioppo, 1986). Petty and Cacioppo (1986) treated relevance as the extent to which a message has "intrinsic importance" to an audience. Using that definition, we assume that the public universally perceives some topics (such as death, family, and work) as more important than others because these topics either involve the loss frame (Kahneman & Tversky, 1979) or are closely related to someone's economic and social well-being during a crisis. Additionally, relevance is manifested in the use of a personal and informal tone (Payne, 2007).

Proposition 4: Cues indicative of a higher degree of personal relevance positively predict information sharing.

2.4. Peripheral cues in online content

In building the four propositions, we point to a set of cues located in two areas: Social media users' profiles and their writings. The profile cues are relatively straightforward, such as the number of followers/friends, which provides a clear indication of influence and popularity. Cues in writings are latent and are manifested as style differences in the use of language (Payne, 2007), which are referred to as *linguistic cues*. Evidence suggests that an audience actively uses linguistic cues to evaluate content (Toma & D'Angelo, 2015). Sociolinguistic studies have revealed that language styles unveil varied levels of sophistication based on the amount of information unpacked through words (Brysbaert, Warriner, & Kuperman, 2014). Language styles also reveal personality, motivation and cognitive process (Pennebaker, Boyd, Jordan, & Blackburn, 2015; Tausczik & Pennebaker, 2010). Recent studies have also linked linguistic cues to persuasion and diffusion outcomes (Hagen el al., 2016; Hwong, Oliver, Van Kranendonk, Sammut, & Seroussi, 2017; Toma & D'Angelo, 2015). However, despite the growing evidence on how linguistic characteristics, coupled with source-level factors predict diffusion, the crisis communication literature has not yet fully embraced the linguistic approach. We thus seek to fill that gap by mapping linguistic cues to the four factors included in the model.

Regarding sentiments, emotions are expressed using a variety of words related to sadness, anger, anxiety and happiness. Prior studies show that the use of emotional words in citizen conversations and media coverage fluctuates throughout crises (Back, Küfner, & Egloff, 2010 Doré, Ort, Braverman, & Ochsner, 2015; Gortner, & Pennebaker, 2003). Concerning richness, studies show that people use text length and the presence of multimedia elements to inform judgment (Toma & D'Angelo, 2015). Additionally, the use of articles, preposition and conjunctions are associated with concrete, formal and cognitively complex thinking (Pennebaker, et al., 2014). Nouns, adjectives and verbs also vary by the concreteness of the information they convey (Brysbaert, Warriner, & Kuperman, 2014). Regarding authority, a category of words indicative of certainty, dominance and confidence are associated with the notion of *Clout* or authority (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2013). Concerning relevance, previous studies have produced a list of words pertaining to personal concerns, including the topics of death, family, friends, and money, as well as words that are intimate and informal (Pennebaker, Boyd, Jordan, & Blackburn, 2015).

3. Methods

3.1. Data collection

Twitter is the site of observation for this study. As one of the world's largest online social networks, Twitter is an ideal platform for the global public to speculate on global events. Twitter had its largest share of audience in Asia Pacific nations in 2014 (Emarketer, 2014), the areas most affected by MH370. Admittedly, the focus on Twitter and its content in English left out a significant portion of conversations that occurred on Chinese social media outlets such as Weibo and tweets in local languages (e.g., Malay). We wrote a Python script to collect tweets containing the hashtag #MH370. The script was executed daily to query the latest 100 pages of tweets returned from Twitter API. We endeavored to reach a maximal coverage of all tweets available through Twitter API. However, Twitter is notoriously secretive about the indexing method used in its API. So, our dataset is at best described as a convenient sample of the global Twitter stream of #MH370.

The raw database had 83,192 tweets, sent by 55,349 unique users between March 8th, 2014 and March 31st, 2014. The time frame covered all critical developments on the ground. We conducted iterations of automated and manual text cleaning to first remove non-original tweets (i.e., retweeted content distinguishable by RT @ or MT @ in the string). Such tweets simply convey simply other-provided content or include a minuscule amount of original content to influence the tweet's virality. We then excluded directed tweets (distinguished by @ SCREENNAME) in that they were used for dyadic interactions instead of publicity. Lastly, we removed tweets not related to MH370, and exceedingly short tweets (shorter than five words). The final dataset had 13,322 tweets by 9,711 users.

3.2. Measures

The collected Twitter data contains tweets and Twitter users' profile information such as follower count. We applied computational linguistic programs to analyze tweet content to extract linguistic cues. The measures of linguistic cues were primarily based on the latest version of Linguistic Inquiry and Word Count (LIWC2015) (Pennebaker, Boyd, Jordan, & Blackburn, 2015). LIWC is a software for counting the portion of words pertaining to 90 linguistic categories.

Sentiment was measured by the LIWC categories of *positive emotion*, *anger*, *anxiety*, and *sadness*. The positive emotion dictionary in LIWC contains 620 words and includes exemplary words such as *love*, *sweet* and *nice*. The anger dictionary contains 230 words, including exemplary words such as *hate*, *kill* and *annoyed*. The anxiety dictionary contains 116 words, including words such as *worried* and *fearful*. The sadness dictionary contains 136 words, including words such as *crying*, *grief* and *sad*.

Richness was measured firstly by the LIWC category *Analytic*. The category reveals the degree of analytical, logical and consistent thinking, as opposed to more intuitive, narrative writing (Pennebaker, Boyd, Jordan, & Blackburn, 2015). This category is derived from prior studies linking the use of articles, prepositions and conjunctions to logical and analytical thinking (Pennebaker, et al., 2014). Secondly, Richness was measured by *Word Count (WC)* and *words per sentence (WPS)*, with the premise that longer tweets and longer sentences convey richer information. Thirdly, richness was measured by a number of multi-media elements (i.e., image/video and URLs). Lastly, to complement the existing LIWC measures, richness was also assessed based on the average concreteness rating of all words in a tweet, using Brysbaert, Warriner, & Kuperman's (2014) concreteness dictionary.

Authority was measured by the number of followers a Twitter user has. Follower count is a straightforward indicator of source influence. While previous studies used network-level measures (e.g., centrality) as a proxy of source influence (Xu, Sang, Blasiola, & Park,), such measures are related to a source's influence in local, topical Twitter networks. We believe the follower count is a better proxy because the count is readily available to all users and captures source influence in the entire Twitterverse. Authority was also gauged by the presence of powerful and confident language style. Precisely, this language style is captured by the LIWC category *Clout* (Pennebaker, Boyd, Jordan, & Blackburn, 2015). The Clout category is derived from prior research that correlates pronoun uses with social hierarchy (Kacewicz et al., 2013). Higher Clout score is marked by using more we-words and social words and fewer I-words, negations (e.g., no, not), and swear words.

Relevance was first measured by the topical categories in LIWC: death, home, work, religion, etc. The *Death* dictionary of 74 words includes words such as *bury*, *coffin*, and *kill*. The *Money* dictionary contains 226 words, including exemplary words such as *audit*, *cash*, and *owe*. The *Religion* dictionary contains 174 words, including words such as *church* and *altar*. The Home dictionary contains 100 words, including words such as *kitchen* and *landlord*. The Work dictionary contains 444 words, including such words as *job*, *majors*, and *Xerox*. Secondly, relevance was measured by informality, with more informal language indicating a higher degree of relevance. In LIWC, the *Informal* language dictionary contains 380 words, including such categories as *Swear* words (e.g., *f**k*, *damn*, *shit*), *Netspeak* (e.g., *btw*, *lol*, *thx*), *Assent* (e.g., *agree*, *OK*, *yes*), *Nonfluencies* (e.g., *er*, *hm*, *umm*), and *Fillers* (e.g., *I mean*, *you know*). Lastly, relevance was measured by the LIWC category *Authenticity*, which indicates to what extent the language used is personal and self-revealing, rather than detached and guarded (Pennebaker, Boyd, Jordan, & Blackburn, 2015). Language scoring high on authenticity is marked by more I-words, present-tense verbs, and relativity words (e.g., old, far, here) and fewer she-he words and discrepancies.

The outcome variable, social media sharing, was measured by the retweet count obtained through the Twitter API. The retweet count captures the virality of the message at the time of data collection. On average, the studied tweets were captured seven hours after the time of posting. The difference between time of posting and time of data collection is generally constant for all tweets (Mean = 7.38, SD = .64). Figure 1 presents a visualization of the conceptualization and measures of the model.

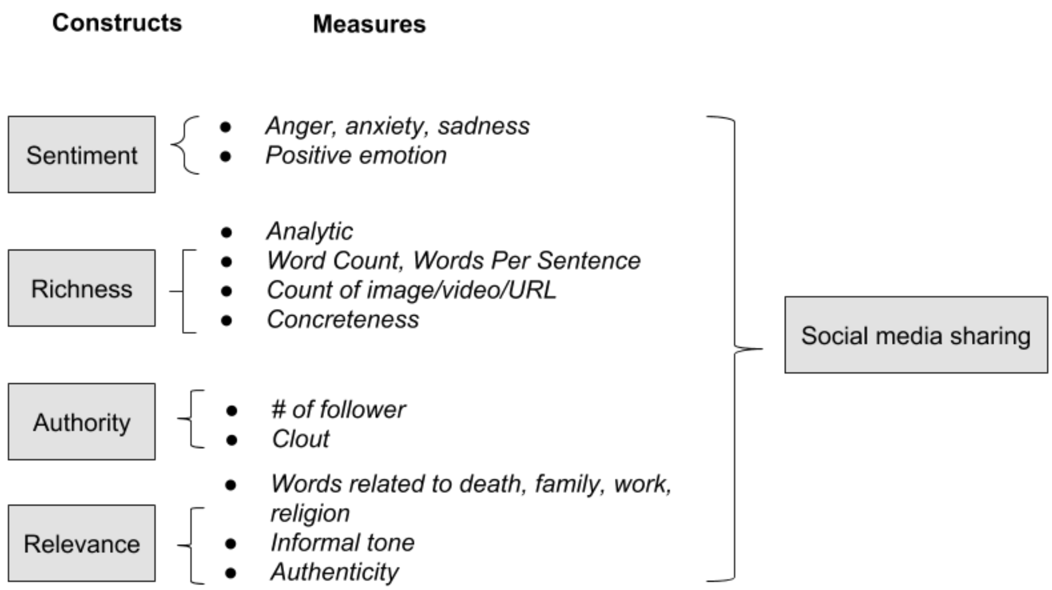

Figure 1. A visualization of the predictive model

Significant events on the ground may have spurious influence on sharing. That is, people are more likely pay attention to and retweet about the issue in light of breaking news or major developments. To account for that, tweets sent on the day of a key event and the following days (n=4267) were labelled to create a dichotomous variable named Important event. The following dates were considered: March 8th and 9th (the plane's reported disappearance), March 15th and 16th (it was reported that MH370 remained in contact with the satellite hours after it had gone off radar), March 24th and 25th (the final announcement of the tragic ending of the plane by the Malaysian authorities).

Thematic differences between tweets might account for variations of linguistic cues and retweeting patterns. In such cases, social media sharing is driven not necessarily by the attention to linguistic cues, but by the audience's interest in certain themes. Therefore, we used semantic network analysis and unsupervised topic modeling (Latent Dirichlet allocation) to explore the distribution of themes in the top 20% and the bottom 20% of tweets ranked by respective LIWC measures (i.e., Analytic, Clout, Authentic, Positive Emotions, Anxiety, Anger, Sad, and etc.). The analysis shows that although there are variations in the salience of central concepts (indicated by central keywords in semantic networks), both the top and bottom ranked tweets consistently contain three topics: breaking news about MH370, speculation about debris and causes, expressed empathy and prayers. The detailed procedures and outputs are presented as supplementary materials.

3.3. Choice of models

In choosing an appropriate statistical model, we had several considerations. The outcome variable, the number of retweets, is a type of count data and is highly skewed. The nature of the data makes the standard Ordinary Least Squares (OLS) regression an unfit. Typically, models such as Poisson or negative binomial regression can be used for such skewed count data. Retweet count had an excessive number of zeros. Therefore, we used a hurdle model (Zeileis, Kleiber, & Jackman, 2008) to address the zero-inflation issue. A hurdle model contains two components: a negative binomial model for the positive counts (the count portion), and a hurdle component model for modeling zero vs. large counts (the zero portion) (Zeileis, Kleiber, & Jackman, 2008). In the present case, the hurdle model contains a part for analyzing the tweets with at least one retweet and another part for analyzing those tweets with zero retweets. Before running the hurdle model, we used the variance inflation factor (VIF) scores to detect the presence of multicollinearity. The independent variables have VIF scores ranging between 1 and 1.5 indicating no multicollinearity issues. To run the hurdle model, we used the hurdle() function in the R library *pscl* (Zeileis, Kleiber, & Jackman, 2008). To compare the fitness of the model, we compared the model to zero-inflated negative binomial model and Poisson model using standard goodness-of-fit measures.

3.4. Validity check

LIWC, the measurement schemes our model primarily relied on, has been rigorously tested for its reliability and external validity using different textual data (Pennebaker, Boyd, Jordan, & Blackburn, 2015). To further ensure the fit of the approach to the present study context, we conducted a crude validity check by comparing longitudinal trends of the key LIWC measures used in the study (i.e., Analytic, Clout, Authentic, Positive Emotions, Anxiety, Anger, Sad, and etc.). If the LIWC measures were valid, they should reflect changing public responses following the development of events on the ground. On March 15th, the news broke that MH370 remained in contact with the satellite hours after it had gone off radar, and on March 24th the final announcement of the tragic ending of the plane was made by the Malaysian authorities. The trends of LIWC measures indicate that positive sentiment gradually decreased. Sadness score, peaked around the time of the two milestone events. The Anxiety score, however, bottomed around the time of the news of MH370's remaining in satellite contact, possibly indicating a sign of rekindled hope for survivors. The Analytic score peaked upon the release of the news of MH370's mysterious satellite contact, while the Clout score bottomed. This could reflect that the public's collective thinking intensified when there was an escalation of events on the ground. And when a puzzling piece of news came out (the plane's mysterious satellite contact), the public conversations showed a sign of hesitation and doubt. Overall, the trends are generally in line with the events on the grounds. Detailed outputs of the validity check are presented as supplementary materials.

4. Results

Table 1 presents the descriptive statistics of all variables, while Table 2 presents the results from the hurdle model. Based on the zero portion of the hurdle model, the model that predicts whether or not a tweet generates a retweet, tweets from a highly followed user are more likely to receive at least one retweet. Such tweets also contain more words, include images, demonstrate a sense of power and confidence (a higher Clout score), sound more authentic, elicit anxiety, and mention home and work-related topics. Retweeting less likely occurs in tweets that contain angry language and URLs.

Table 1. Descriptive statistics

Factors	***Variables***	**Mean**	**S.D**	**Twitter average**	**Min.**	**Max**
Sentiment	*Positive Sentiment*	2.5	4.92	5.48	0	50
	Anxiety	0.33	1.49	0.24	0	22.22
	Anger	0.37	1.64	0.75	0	28.57
	Sadness	1.25	2.85	0.43	0	31.58
Richness	*WC*	17.22	5.97	NA	5	34
	WPS	12.01	6.19	12.1	1.4	33
	Concreteness	1.61	0.48		0	3.1
	URL count	0.31	0.48	NA	0	3
	image/video Count	0.09	0.29	NA	0	1
	Analytic	71.95	31.37	61.94	1	99
	Follower count	14237	291201	NA	0	16,066,591
	Clout	56.26	26.98	63.02	1	99

Authority						
	Authentic	32.41	35.1	50.39	1	99
	Informal	3.61	4.5	4.68	0	40
	Death-related topic	0.17	1.12	0.19	0	25
	Money-related topic	0.24	1.34	0.74	0	22.22
Relevance	*Religion-related topic*	1.22	3.84	0.35	0	70
	Home-related topic	0.45	1.79	0.43	0	25
	Work-related topic	1.23	2.92	2.16	0	31.58
Outcome variable	*Retweet count*	1.39	11.45	NA	0	555

Note: Twitter Average is obtained from Pennebaker, Boyd, Jordan and Blackburn (2015)

The count portion of the hurdle model, the part that predicts the number of retweets, is tested on the tweets with at least one retweet, a tweet accumulates more retweets when it is from a highly followed user, contains more words, has a sense of power and confidence, contains positive emotion, and includes images/videos and mentions work-related topics. A tweet is less likely to accumulate retweets when it includes URLs and uses an informal tone.

For goodness-of-fit measures, we also ran a zero-inflated negative binomial model and a Poisson model, using the R function zeroinfl() and glm() respectively. The Akaike information criterion (AIC) for the hurdle model is 27493.14, lower than the AIC for the negative binomial model (27953.58) and the Poisson model (83037.29), signifying that the hurdle model is the best fit among the three. The Vuong statistics in the comparisons across the three models also show that the hurdle model consistently fits the data better.

Table 2. The hurdle models

		Zero Portion				Count Portion			
		Estimate (SE)	OR	*z* value	*p*	Estimate (SE)	IRR	*z* value	*p*
	(Intercept)	-0.64 (0.05)	0.53	-13.62	***	-14.26 (55.78)	0.00	-0.26	
	Positive Sentiment	0.03(0.02)	1.03	1.51		0.09 (0.04)	1.09	2.42	*
	Anxiety	0.04 (0.02)	1.04	2.17	*	-0.07 (0.04)	0.93	-2.01	*
Sentiment	*Anger*	-0.05 (0.02)	0.95	-1.98	*	-0.06 (0.04)	0.94	-1.45	
	Sadness	0.01 (0.02)	1.01	0.61		0.05 (0.03)	1.05	1.60	

	WC	0.20 (0.03)	1.22	7.67	***	0.31 (0.05)	1.36	6.59	***
	WPS	0.03 (0.02)	1.03	1.26		-0.06 (0.04)	0.94	-1.42	
	Concreteness	-0.00 (0.03)	1.00	-0.13		-0.04 (0.04)	0.96	-0.85	
Richness	*URL count*	-0.17 (0.03)	0.84	-5.88	***	-0.10 (0.05)	0.90	-2.13	*
	IMAGE/VIDEO COUNT	0.23(0.02)	1.26	10.57	***	0.48 (0.04)	1.62	13.18	***
	Analytic	-0.04 (0.02)	0.96	-1.62		-0.001 (0.04)	1.00	-0.04	
	Follower count	13.69 (0.95)	882046	14.44	***	10.32 (1.04)	30333	9.97	***
Authority	*Clout*	0.07 (0.02)	1.07	3.35	***	0.15 (0.04)	1.16	3.92	***

	Authentic	0.05(0.02)	1.05	2.20	*	0.06 (0.04)	1.06	1.31	
	Informal	-0.02 (0.03)	0.98	-0.71		-0.26 (0.05)	0.77	-5.35	***
	Death-related topic	-0.03 (0.02)	0.97	-1.49		-0.08 (0.04)	0.92	-1.79	
	Money-related topic	-0.01 (0.02)	0.99	-0.68		-0.01 (0.03)	.99	-0.23	
Relevance	*Religion-related topic*	0.02 (0.02)	1.04	0.77		0.04 (0.03)	1.04	1.26	
	Home-related topic	0.04(0.02)	1.05	2.02	*	-0.03 (0.03)	0.97	-0.86	
	Work-related topic	0.04 (0.02)	1.05	2.02	*	0.16 (0.04)	1.17	3.97	***
Control variable	*Key events (yes/no)*	.32 (.04)	1.38	7.39	***	0.20 (0.08)	1.22	2.69	***

Log(theta)	-16.04	0.00	-0.29
	(55.78)		

Note: All variables were converted into Z scores. * $p < .05$ (two-tailed), ** $p < .01$ (two-tailed). Different from lineal regression models, a hurdle model produces coefficients used for interpreting the degree of change induced by a variable. For positive coefficients, the degree of change is calculated by exponentiating a coefficient and subtracting one from the value. For negative coefficients, subtract the exponentiated coefficient from 1. For example, the coefficient for the variable WC is .20. Its exponentiated value is 1.22, indicating a 22% increase (1.22-1).

The exponential coefficients in the zero portion of the hurdle model indicate how the independent predictors relate to crossing the 'hurdle' of getting at least one retweet. A unit of change in the number of followers increases the odds of obtaining a retweet by 782046% while holding all other predictors in the model constant. A unit of change in Word Count increases the odds of retweeting by 22%. A unit of change in Clout score is associated with the increase in the odds by 7%. A unit of change in Authenticity score, in Anxiety, in mentions of home-related topics and of work-related topics, increases the odds by 4~5%, respectively. A unit of change in image/video count increases the odds by 26%. The change in Anger, however, decreases the odds by 5%. And a unit of change in URL count decreases the odds by 16%. Tweets sent on the days of key events are associated with a 38% increase.

The exponential coefficients in the count portion of the hurdle model focus on tweets that have received at least one retweet. With a unit of change in the number of followers, the number of retweets is expected to increase by a factor of 30333. A unit of change in Word Count is associated with an increase in retweet count by a factor of 1.36. A unit of change in Clout score is expected to lead an increase in retweet count by a factor of 1.16. A unit of change in Anxiety score is expected to lead a decrease in retweet count by a factor of .93. A unit of change in positive emotion increases retweet count by a factor of 1.09. A unit of change in image/video count increases retweet count by a factor of 1.62. Using URLs decreases retweet count by a factor of .90. And using informal language increases retweet count by a factor of 0.77. Mentioning work-related topics increases retweet count by a factor of 1.17. Tweets sent on days of key events are associated with a 22% increase.

5. Discussion

We explored a group of factors associated with information sharing during crises. The literature on information processing reveals four salient factors in a crisis context: sentiment, richness, authority, and relevance. We then operationalized and measured the four factors based on the peripheral cues in tweets and in user profiles. In doing so, the study makes several notable contributions. First, it is an innovative attempt to measure the four factors using computational linguistics. While the approach has been gradually adopted by communication scholars, it remains a novelty in the crisis communication literature. Second, the study joins a small but growing body of research that applies dual process theories to predict message diffusion on social media (Liu, Burton-Jones, & Xu, 2014, Shi, Hu, Lai, & Chen, 2018). Departing from the literature's focus on rumor retransmission and general social media sharing, our study uses the framework to understand the sharing of crisis information. In short, the study is the first that applies both dual process theories and computational linguistics to understand crisis-time social media sharing.

Findings suggest that after controlling for the effect of the timing of key events, content with a confident, self-revealing (authentic) and emotional language style, and with multimedia cues, and cues of source popularity, have higher retweetability (a key metric in online information diffusion). Contrarily, content featuring some degrees of negativity and informality seems to discourage the retweeting behavior. The findings are consistent with Elaboration Likelihood Model, suggesting that content and source-level cues do matter in users' perception and subsequent evaluation of online content. However, the findings note that such effect varies by the types of peripheral cues.

First, Sentiment — both positive and negative emotions — generally predicts more sharing. The finding supports both a key proposition in dual process theories that audiences are attentive to emotional cues (Petty & Cacioppo, 1986) and previous evidence that emotionality matters in online information sharing (Liu, et al., 2017). But, the study also confirms the varying effects of different sentiments in the online environment (Song, Dai, & Wang, 2016). The public is not necessarily drawn to negative content even though a crisis is tragic and unnerving. Cues of anger discourage retweeting. The effect of anxiety cues is mixed: it helps tweets get a retweet, but not many retweets. Cues of positive emotions, on the other hand, are linked with more sharing. Perhaps, the public avoids negativity to seek a rare sense of hope and comfort. With the negative nature of the event, anything positive-sounding might stand out as new. It is also possible that sharing negative content (especially that which indicates anger) might be deemed unattractive according to online social norms. Taken together, the findings show multiple psychological motivations at work during social crises: as individuals, sharing tweets conveying an anxious mindset could be an effective way to release their own pressure, since people are not immune to the psychological disequilibrium brought by such events (Paton & Violanti, 2006); however, as group members, the public may prefer prosocial and positive information, because positivity such as love, generosity, and altruism could produce more spontaneous, sympathetic and sentimental reactions than negative information with hostility and blame (Fritz & Williams, 1957).

Second, concerning Richness, it is debatable in the general persuasion literature regarding whether informational appeals outweigh emotional appeals. Some claim that informational appeals enhance credibility while others assert that emotional appeals make content more memorable (see Liu, et al., 2017 for a review). Our findings show that the public values informationality as well as emotionality. A tweet's informational value is two-fold, reflecting both the completeness of cues and the objective, consistent and analytic language style. Twitter's

140-character limit creates a norm that to attract attention, a message should be short and sweet. However, the message should not be so short as to lose the completeness of cues being conveyed. The finding that the most retweeted tweets are those with long sentences, images and videos supports the salient role of media richness identified in the literature (Liu, Ji, North, & Yang, 2017; Martey et al., 2015). However, an analytical language style is not appealing. Its lesser appeal could mean that the public desires something more stimulating and speculative, rather than merely informational (Ziegele, Breiner, Quiring, 2014).

Third, Authority cues is a key driver of virality. Its largest effect is seen in cues of source popularity. This finding is supportive of prior research on opinion leadership that shows a clear link between highly followed opinion leaders and word-of-mouth diffusion (Song, Dai, & Wang, 2016; Xu, Sang, Blasiola, & Park, 2014; Zhang el al., 2014). A post-hoc analysis of the most followed #MH370 users reveals that virtually all of them are legacy news outlets such as *BBC World*, *CNN*, *Strait Times*, and *The Wall Street Journal*. This echoes previous studies that reveal traditional media's prominent and remaining influence in the social media age (Li, J., Vishwanath, A., & Rao, 2014). Several reasons may account for the phenomenon. First, the highly followed accounts produced more high-quality content during the crisis, which drives virality. Second, the accounts already had a large audience base to diffuse messages. Third, as predicted by dual process theories, the Twitter users evaluated source credentials by their follower count, deeming more follower size as more trustworthy. Another aspect of Authority, reflected in the use of powerful, dominant and confident language style, also consistently predicts a better diffusion outcome. Tweets with such a linguistic cue demonstrate more we-ness and are less self-focused, a sign of leadership in the time of crisis. A post-hoc review shows that such tweets usually call for collective and united actions, such as "#nowplaying: birdy - people help the people. virtual hugs to families of passengers of #mh370." Arguably, users might be drawn to such content because it induces a sense of collectiveness and unity.

Lastly, concerning the cues of relevance, self-revealing and personal tweets are effective in generating retweets, alluding to the social and personalized nature of digital media. But, informality in tweets is also related to lower retweetability. The LIWC measures of informality are in part based on the presence of swear language. Our post-hoc analysis shows that swear language is to a large extent associated with fewer retweets, which adds further evidence to the previous point that anger drives down the sharing behavior.

In summary, findings reveal multiple peripheral cues at work in influencing social media sharing. While factors identified in the existent dual process theories are generally predictive of sharing, the effect of cues is contextualized to reveal distinct spontaneous reactions by the public to the negativity of tragic events and the informational and emotional gaps created by the crises.

6. Limitations and future directions

We remind readers of several limitations in the present study. First and foremost, the model is based on the assumption that online publics use linguistic and user profile cues to evaluate tweets. That assumption, derived theoretically, is not empirically tested in the current context. Arguably, without a controlled environment, there is no practical way of tracing how Twitter users actually process the tweets. The findings from the study afford at best partial evidence or a promising lead to unlock the link between peripheral cues and social media sharing. Second, we are aware that social media texts are unstructured and noisy. We did due diligence in cleaning the data to remove irrelevant tweets and in running an iteration of post-hoc analyses to test the validity of the measures. However, our text cleaning and validity check were rather crude. And the current version of LIWC is not able to fully capture internet slang and coded languages in tweets, as well as various types of sentiment expressed using emoji. Our approach could have overlooked unconventional message and source-related cues that are just as revealing as the cues studied. Third, the present study only examines English-language tweets, thus it overlooks how Chinese publics reacted to the accident via Chinese social media and how Malay-speaking users commented on the issue. Lastly, this study is based on a single case, which limits the generalizability of the findings.

Seeing the current study as the first step, our research team plans to apply the Sentiment-Richness-Authority-Relevance model to other contexts (such as hurricanes and earthquakes). Future studies can also test the validity of the model, particularly, in terms of whether an audience perceives and processes the studied peripheral cues. To that end, researchers can invite study participants to rate connotations of different words in a specific crisis context as applied to the four factors. They can also use manual content analysis to extract contextual meanings words as they appear in a certain context.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

References

Adamic, L. A., & Huberman, B. A. (2000). Power-law distribution of the world wide web. *Science*, *287*(5461), 2115-2115.

Allport, G. W., & Postman, L. (1947). *The psychology of rumor*, New York: Holt, Rinehart & Winston.

Back, M. D., Küfner, A. C., & Egloff, B. (2010). The emotional timeline of September 11, 2001. *Psychological Science*. http://journals.sagepub.com/doi/pdf/10.1177/0956797610382124

Boulianne, S., Minaker, J., & Haney, T. J. (2018). Does compassion go viral? Social media, caring, and the Fort McMurray wildfire. *Information, Communication & Society*, *21*(5), 697-71.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911.

Chen, R., & Sakamoto, Y. (2013, January). Perspective matters: Sharing of crisis information in social media. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on* (pp. 2033-2041). IEEE.

Choi, S. (2014). Flow, diversity, form, and influence of political talk in social-media-based public forums. *Human Communication Research*, *40*(2), 209-237.

Choi, W., & Stvilia, B. (2015). Web credibility assessment: Conceptualization, operationalization, variability, and models. *Journal of the Association for Information Science and Technology*, *66*(12), 2399-2414.

Chua, A. Y., Tee, C. Y., Pang, A., & Lim, E. P. (2016, July). The Retransmission of Rumor-related Tweets: Characteristics of Source and Message. In *Proceedings of the 7th 2016 International Conference on Social Media & Society* (p. 22). ACM.

Chung, J. E. (2017). Retweeting in health promotion: Analysis of tweets about Breast Cancer Awareness Month. *Computers in Human Behavior*, *74*, 112-119.

Davenport, T. H., & Beck, J. C. (2001). *The attention economy: Understanding the new currency of business*. Cambridge, MA: Harvard Business Press.

Doré, B., Ort, L., Braverman, O., & Ochsner, K. N. (2015). Sadness shifts to anxiety over time and distance from the national tragedy in Newtown, Connecticut. *Psychological Science*, *26*(4), 363-373.

EMarketer. (2014, June 6). Asia-Pacific Grabs Largest Twitter User Share Worldwide [Web log post]. Retrieved Feb 9, 2018, from https://www.emarketer.com/Article/Asia-Pacific-Grabs-Largest-Twitter-User-Share-Worldwide/1010905

Fritz E.C. & Williams B.H. (1957). The Human Being in Disasters: A Research Perspective. *The Annals of The American Academy*.

Getchell, M. C., & Sellnow, T. L. (2016). A network analysis of official Twitter accounts during the West Virginia water crisis. *Computers in Human Behavior*, *54*, 597–606.

Gortner, E. M., & Pennebaker, J. W. (2003). The archival anatomy of a disaster: Media coverage and community-wide health effects of the Texas A&M bonfire tragedy. *Journal of Social and Clinical Psychology*, *22*(5), 580-603.

Hagen, L., Harrison, T. M., Uzuner, Ö., May, W., Fake, T., & Katragadda, S. (2016). E-petition popularity: Do linguistic and semantic factors matter?. *Government Information Quarterly*, *33*(4), 783-795.

Hwong, Y. L., Oliver, C., Van Kranendonk, M., Sammut, C., & Seroussi, Y. (2017). What makes you tick? The psychology of social media engagement in space science communication. *Computers in Human Behavior*, *68*, 480-492.

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2013). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33, 125-143.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kaufman, D. Q., Stasson, M. F., & Hart, J. W. (1999). Are the tabloids always wrong or is that just what we think? Need for cognition and perceptions of articles in print media. *Journal of Applied Social Psychology*, *29*(9), 1984-2000.

Kim, E., Hou, J., Han, J. Y., & Himelboim, I. (2016). Predicting retweeting behavior on breast cancer social networks: network and content characteristics. *Journal of Health Communication*, *21*(4), 479-486.

Kwon, K. H., Bang, C. C., Egnoto, M., & Raghav Rao, H. (2016). Social media rumors as improvised public opinion: semantic network analyses of twitter discourses during Korean saber rattling 2013. *Asian Journal of Communication*, *26*(3), 201-222.

Li, J., Vishwanath, A., & Rao, H. R. (2014). Retweeting the Fukushima nuclear radiation disaster. *Communications of the ACM*, *57*(1), 78-85.

Liu, F., Burton-Jones, A., & Xu, D. (2014). Rumors on Social Media in disasters: Extending Transmission to Retransmission. In *PACIS* (p. 49).

Liu, J., Li, C., Ji, Y. G., North, M., & Yang, F. (2017). Like it or not: The Fortune 500's Facebook strategies to generate users' electronic word-of-mouth. *Computers in Human Behavior*, *73*, 605–613.

Martey M.R., Galley S.J., Consalvo M., Wu J., & Banks J. (2015). Communicating age in second life: the contributions of textual and visual factors. *New Media & Society*, *17*(1), 41-61.

Meraz, S., & Papacharissi, Z. (2013). Networked gatekeeping and networked framing on# Egypt. *The International Journal of Press/politics*, *18*(2), 138-166.

Mills, A., Chen, R., Lee, J., & Raghav Rao, H. (2009). Web 2.0 emergency applications: How useful can Twitter be for emergency response?. *Journal of Information Privacy and Security*, *5*(3), 3-26.

Nabi, R. (2002). Anger, fear, uncertainty, and attitudes: A test of the cognitive-functional model. *Communication Monographs*, *69*(3), 204-216.

Neubaum, G., Rösner, L., Rosenthal-von der Pütten, A. M., & Krämer, N. C. (2014). Psychosocial functions of social media usage in a disaster situation: A multi-methodological approach. *Computers in Human Behavior*, *34*, 28-38.

Oh, O., Agrawal, M., & Rao, H. R. (2013). Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. *MIS Quarterly*, *37*(2).

Oh, O., Gupta, P., Agrawal, M., & Rao, H. R. (2018). ICT mediated rumor beliefs and resulting user actions during a community crisis. *Government Information Quarterly*, *35*(2), 243-258.

Pang, N., & Ng, J. (2016). Twittering the Little India Riot: Audience responses, information behavior and the use of emotive cues. *Computers in Human Behavior*, *54*, 607-619.

Papacharissi, Z., & de Fatima Oliveira, M. (2012). Affective news and networked publics: The rhythms of news storytelling on #Egypt. *Journal of Communication*, *62*(2), 266-282.

Paton Douglas & Violanti M. Jone. (2006). Policing in the Context of Terrorism: Managing Traumatic Stress Risk. *Traumatology*, *12*(3), 236-247.

Payne, C. R. (2007). The elaboration likelihood model of persuasion: Implications for trial advocacy. *International Journal of Speech Language and the Law*, *14*(2), 309-312.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*.

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. *PloS one*, *9*(12).

Petty, R. E., & Cacioppo, J. T. (1986). Message elaboration versus peripheral cues. In *Communication and Persuasion* (pp. 141-172). Springer New York.

Reuter, C., Hughes, A. L., & Kaufhold, M. A. (2018). Social media in crisis management: An evaluation and analysis of crisis informatics research. *International Journal of Human–Computer Interaction*, *34*(4), 280-294.

Rice, R. G., & Spence, P. R. (2016). Thor visits Lexington: Exploration of the knowledge-sharing gap and risk management learning in social media during multiple winter storms. *Computers in Human Behavior*, *65*, 612–618.

Rowley, J., & Johnson, F. (2013). Understanding trust formation in digital information sources: The case of Wikipedia. *Journal of Information Science*, *39*(4), 494-508.

Shi, J., Hu, P., Lai, K. K., & Chen, G. (2018). Determinants of users' information dissemination behavior on social networking sites: An elaboration likelihood model perspective. *Internet Research*, *28*(2), 393-418.

Shklovski, I., Palen, L., & Sutton, J. (2008, November). Finding community through information and communication technology in disaster response. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 127-136). ACM.

Slater, M. D., & Rouner, D. (1996). How message evaluation and source attributes may influence credibility assessment and belief change. *Journalism & Mass Communication Quarterly*, *73*(4), 974-991.

Song, Y., Dai, X. Y., & Wang, J. (2016). Not all emotions are created equal: Expressive behavior of the networked public on China's social media site. *Computers in Human Behavior*, *60*, 525-533.

Sparks, J. R., & Areni, C. S. (2008). Style versus substance: Multiple roles of language power in persuasion. *Journal of Applied Social Psychology*, *38*(1), 37-60.

Stieglitz, S., Bunker, D., Mirbabaie, M., & Ehnis, C. (2018). Sense-making in social media during extreme events. *Journal of Contingencies and Crisis Management*, *26*(1), 4-15.

Stieglitz, S., & Dang-Xuan, L. (2012, January). Political communication and influence through microblogging--An empirical analysis of sentiment in Twitter messages and retweet behavior. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 3500-3509). IEEE.

Sutton, J. N., Palen, L., & Shklovski, I. (2008). *Backchannels on the front lines: Emergency uses of social media in the 2007 Southern California Wildfires* (pp. 624-632). University of Colorado.

Sussman, S. W., & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, *14*(1), 47-65.

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. *Social Networks*, *34*(1), 73-81.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24-54.

Teng, S., Wei Khong, K., Wei Goh, W., & Yee Loong Chong, A. (2014). Examining the antecedents of persuasive eWOM messages in social media. *Online Information Review*, *38*(6), 746-768.

Toma, C. L., & D'Angelo, J. D. (2015). Tell-tale words: Linguistic cues used to infer the expertise of online medical advice. *Journal of Language and Social Psychology*, *34*(1), 25-45.

Veltri, G. A., & Atanasova, D. (2015). Climate change on Twitter: Content, media ecology and information sharing behaviour. *Public Understanding of Science*, 0963662515613702.

Xu, W. W., Sang, Y., Blasiola, S., & Park, H. W. (2014). Predicting opinion leaders in Twitter activism networks: The case of the Wisconsin recall election. *American Behavioral Scientist*, *58*(10), 1278-1293.

Xu, Q. (2013). Social recommendation, source credibility, and recency: Effects of news cues in a social bookmarking website. *Journalism & Mass Communication Quarterly*, *90*(4), 757-775.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of statistical software*, *27*(8), 1-25.

Zhang, L., Xu, L., & Zhang, W. (2017). Social media as amplification station: factors that influence the speed of online public response to health emergencies. *Asian Journal of Communication*, *27*(3), 322-338.

Zhang, L., Peng, T. Q., Zhang, Y. P., Wang, X. H., & Zhu, J. J. (2014). Content or context: Which matters more in information processing on microblogging sites. *Computers in Human Behavior*, *31*, 242-249.

Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research*, *15*(4), 267.

Ziegele, M., Breiner, T., & Quiring, O. (2014). What creates interactivity in online news discussions? An exploratory analysis of discussion factors in user comments on news items. *Journal of Communication*, *64(6)*, 1111-1138.