Emotions in Online Content Diffusion

Yifan Yu*,Shan Huang*†,Yuchen Liu,Yong Tan
Michael G. Foster School of Business, University of Washington, Seattle, Washington 98195

Social media-transmitted online information, particularly content that is emotionally charged, shapes our thoughts and actions. In this study, we incorporate social network theories and analyses to investigate how emotions shape online content diffusion, using a computational approach. We rigorously quantify and characterize the structural properties of diffusion cascades, in which more than six million unique individuals transmitted 387,486 articles in a massive-scale online social network, WeChat. We detected the degree of eight discrete emotions (i.e., surprise, joy, anticipation, love, anxiety, sadness, anger, and disgust) embedded in these articles, using a newly generated domain-specific and up-to-date emotion lexicon. We found that articles with a higher degree of anxiety and love reached a larger number of individuals and diffused more deeply, broadly, and virally, whereas sadness had the opposite effect. Age and network degree of the individuals who transmitted an article and, in particular, the social ties between senders and receivers, significantly mediated how emotions affect article diffusion. These findings offer valuable insight into how emotions facilitate or hinder information spread through social networks and how people receive and transmit online content that induces various emotions.

Key words: Information Diffusion, Online Content, Emotion Detection, Social Networks, Social Media

---

1. Introduction

Emotions are commonly expressed in natural language and shape our daily communications. Neuroscientists have found that humans often pay special attention to and establish enhanced memory for emotional events, which activate subsequent actions (Dolan 2002). Social media, including Twitter, Facebook, and WeChat, has facilitated rapid information sharing and large-scale information cascades. Online information transmitted in social media networks, particularly content that is emotionally charged, shapes our thoughts and actions. The emotional information that spreads on social media and reaches individuals through social interactions affects our notions of such concerns as morality, ideology (Brady et al. 2017), politics, terrorism (Clarke et al. 2006, Vosoughi et al. 2018), and financial investments (Bollen et al. 2011, Nguyen et al. 2020) through the emotional states induced by the content. Notably, when misinformation is incorporated with the emotional

* These authors contributed equally to this work.
† To whom correspondence should be addressed. E-mail: shanh18@uw.edu

content that facilitates network diffusion, it can be harmful, especially in the case of COVID-19. Thus, it is timely and vital to understand how emotions affect online information diffusion.

It is empirically challenging to rigorously characterize the process of peer-to-peer spreading of online content in social media networks and, thus, to investigate the relationship between emotions and the structural properties of diffusion cascades. Understanding the characteristics of diffusion process is critical in informing the strategies to maximize online content spreading, such as by seeding in different individuals and structural elements of networks and utilizing different models: viral or broadcast (Goel et al. 2016). Prior studies, such as Berger and Milkman (2012) and Brady et al. (2017), focus mainly on the impact of emotions in initial social transmission (i.e., the number of seeds in the first-level cascades) or the general popularity of tweets or online articles. The same number of seeds, however, can lead to extremely different results in regard to diffusion. Further, the affective dimensions of content, such as embedded emotions, have received much less attention than have their cognitive counterparts (Vosoughi et al. 2018, Aral and Dhillon 2016), due mainly to the difficulty of emotion detection, particularly in a massive amount of content. To the best of our knowledge, few comprehensive empirical investigations have examined the relationship between emotions and the diffusion process of online content. We therefore conduct a very large-scale field test of a random sample that includes 387,486 online articles and their diffusion cascades in China's largest social networking site, WeChat. We incorporate social network theories and analyses to understand whether and how emotions lead to larger, deeper, faster, broader, and more viral cascades of otherwise similar content in a massive-scale online social network, using a computational approach.

We focus on eight *discrete emotions* (i.e., surprise, joy, anticipation, love, anxiety, sadness, anger, and disgust) to delineate complex emotions embedded in online content. According to discrete emotion theory (Tomkins 1962), there is a small number of core emotions that constitute other emotions. For example, awe may be viewed as a blend of anxiety and love. Once we understand discrete emotions, we can understand more complex emotions (Tomkins 1962, Plutchik and Kellerman 1980, Lerner et al. 2004, 2015). Empirical evidence also shows that discrete emotions are relatively independent of each other and can better delineate human emotions than can valence (i.e., positive and negative polarity) and arousal (also "activation" or "intensity", the extent to which a person is energized by an experience) (Lerner et al. 2004, Yin et al. 2017, Yu et al. 2019, Dorison et al. 2020). Further, as Quan and Ren (2010) found, these eight discrete emotions are the most commonly seen in articles in China and cover the most common emotions in our context. To accurately detect these discrete emotions, embedded in the content of our sampled articles, we construct a new domain-specific and up-to-date emotion lexicon, using a state-of-the-art lexicon-generation approach (Xue et al. 2014, Yu et al. 2019), and validate it with human annotation.

We find that the correlations between the eight discrete emotions embedded in each article in our sample are generally below 0.20, confirming the independence of these emotions.

We apply a two-level hierarchical regression model (Hox et al. 2017) to analyze the effects of emotions in information diffusion. This model enables us to comprehensively and effectively control for heterogeneity at both the article and publisher levels. We find that articles with a higher degree of emotion generally reach a significantly larger number of individuals and diffuse significantly more deeply, broadly, and virally but also more slowly. Among the eight discrete emotions, anxiety and love significantly increased cascade size, depth, breadth, and structural virality, whereas sadness significantly decreased these factors but accelerated article diffusion. Further, to understand how emotions shape peer-to-peer spreading that leads to differential diffusion, we analyzed the demographic and network characteristics of more than six million users involved in the cascades and the social ties between them. We find that the average age and network degree of the individuals and, in particular, the proportion of weak ties involved in a cascade significantly mediate the effects of emotions. First, articles with a greater degree of anxiety were spread more among older users, which partially explains why they spread more deeply in the network. In contrast, articles with greater embedded sadness were shared more by younger users, which partially explains the lower levels of size, depth, breadth, structural virality, and speed associated with these articles. Second, articles with more embedded anxiety and love tend to be shared by more central users, which partially explains the positive effects of these two discrete emotions on the articles' cascade depth. Third, articles with a greater degree of sadness were transmitted more through strong ties, which completely explains the negative effects of sadness on cascade depth and structural virality and partially explains its negative effects on cascade size, breadth, and speed. Articles with a greater degree of love transmitted more through weak ties, which completely explains its positive effect on cascade depth and partially explains its effect on structural virality.

Our work makes several significant contributions. First, it adds to the emerging literature on peer-to-peer spreading of online content in social media networks and its economic and social implications. Our study is timely and speaks to this expanding area of inquiry. Second, this research provides some of the first large-scale and comprehensive empirical evidence on how emotions are associated with the network diffusion of online information. Previous studies have focused mainly on psychological processes and take a dyadic perspective, such as readers' perceived arousal that drives dyadic sharing of emotional content (Berger and Milkman 2012, Zhang and Qu 2018). Our work goes beyond the dyads and investigates how online content with various emotions transmits through different individuals and network elements (e.g., social ties), leading to differential diffusion in a large-scale social network. Third, our results provide direct evidence of the critical role of weak ties and their strong explanatory power for the effects of emotions in online content diffusion.

Despite a broad consensus on the large impact of weak-tie theory (Granovetter 1977), identifying weak ties and their effects, particularly in large-scale cascades, is still an empirical challenge. Fourth, we present a methodological framework that rigorously analyzes the emotions embedded in content, accurately maps information cascades, and identifies the impact of emotions on cascades' structural properties. Fifth, our work contributes to emotion theory by comparing discrete emotion theory (Tomkins 1962, Plutchik and Kellerman 1980) with the perspective of valence and arousal (the dimensional theory) (Russell 1980). By finding that emotion pairs, which are similar in valence and arousal (i.e., anxiety and anger, love and joy), exhibit disparate effects on article diffusion, our work suggests that dimensional emotion theory is not sufficient to explain the pattern that we observed. Our results highlight the necessity of using discrete emotions to delineate human emotions. Finally, our results imply that practitioners (e.g., marketers, platform managers) can rely on the emotions embedded in content, together with seeding different individual characteristics and social ties, to promote online content, such as articles, products, and elections in networks.

2. Related Literature

2.1. Discrete Emotions

Current management research which examines the impact of emotions embedded in online content on user behavior (e.g., Rui et al. 2013, Hennig-Thurau et al. 2015, Yin et al. 2017, Song et al. 2019) is primarily driven by the dimensional emotion theory (Russell 1980), which explains the effect of emotions using dimensions such as valence and arousal. Prior research shows that valence in online reviews affects individuals' product evaluations (Doh and Hwang 2009), intentions to recommend products (Lee and Youn 2009), perceived helpfulness of reviews (Hong et al. 2017), product early adoptions (Hennig-Thurau et al. 2015) and daily sales (Song et al. 2019), whereas arousal is associated with review helpfulness (Yin et al. 2017, Ren and Nickerson 2019). Stieglitz and Dang-Xuan (2013) finds that positive and negative (valence) tweets tend to be retweeted more often and more quickly than is neutral content. Berger and Milkman (2012) illustrates that news with more positive and high-arousal emotional expressions is more likely to be shared by readers. Recent psychological research, however, finds that emotions with similar valence and arousal can lead to divergent effects on people's decision making and judgment (Lerner et al. 2015). For example, anger and anxiety are similar in both valence and arousal, but anxiety embedded in online reviews induces higher perceived helpfulness than does anger (Yin et al. 2014). These findings suggest that the dimensions of valence and arousal cannot fully explain the variations in the effects of emotions (Plutchik 2001, Lerner et al. 2015).

Recent work turns to discrete emotion theory to measure emotions embedded in online content (e.g., Malik and Hussain 2017, Yu et al. 2019, Nguyen et al. 2020). Discrete emotion theory, a competing theory for the dimensional emotion model (Watson and Spence 2007), identifies specific basic and independent emotions, i.e., discrete emotions, that constitute other human emotions (Tomkins 1962, Plutchik and Kellerman 1980). Discrete emotions are rooted in human evolution, with expression and recognition fundamentally the same across all individuals, regardless of ethnic or cultural differences (Plutchik 2001). The theory originated from the evolutionary theory, in the late 19th century, of Charles Darwin, who argued that certain basic emotions evolved from natural selection (Darwin and Prodger 1998). Neuroimaging analyses have shown that these discrete emotions are linked to discrete neural signatures and certain structures of human brains (Vytal and Hamann 2010, Saarimäki et al. 2016). Prior management research shows that discrete emotions embedded in online word-of-mouth are of larger predictive power than is valence on perceived review helpfulness (Yin et al. 2014, Malik and Hussain 2017), consumers' purchase decisions and sales performance (Yu et al. 2019), and stock market return (Nguyen et al. 2020). Thus, discrete emotions may provide a means for us to understand the effects of emotion embedded in online content. The eight discrete emotions (i.e., surprise, joy, anticipation, love, anxiety, sadness, anger, and disgust) used in our work are commonly expressed in online content. (Quan and Ren 2010, Yu et al. 2019). These emotions also are considered to be the most basic emotions in discrete emotion research (Lazarus 1991, Lewis et al. 2010, Lerner et al. 2004, 2015, Yin et al. 2014). To understand the effects of these emotions is of greater theoretical significance than to understand others, as various emotions can be considered mixtures of these discrete emotions.

2.2. Emotions and Online Content Diffusion

Understanding what drives online content diffusion has attracted the increasing efforts of researchers from various fields (e.g., Banerjee et al. 2013, Ransbotham and Mitra 2013, Shi et al. 2014, Mitra and Ransbotham 2015, Vosoughi et al. 2018). Notably, emotions are critical in the social transmission of online information. Individuals are motivated to share emotionally charged content to strengthen social connections, coordinate actions (Peters and Kashima 2007), build persona (Berger and Milkman 2012), reduce emotion-related ambiguous sensations and rationalize their emotional experiences (Rime et al. 1991). The emotions embedded in content also can affect the audience's emotional state and their subsequent decisions (Rime et al. 1991, Heath et al. 2001, Berger and Milkman 2012). Individuals who are exposed to high-arousal emotions tend to be activated and are more likely to take action, such as sharing with peers (Berger and Milkman 2012). In this regard, laboratory experiments show that urban legends that evoke a disgusting experience are more likely to be socially transmitted (Heath et al. 2001).

Our paper is related most closely to that of Berger and Milkman (2012), who take a mainly psychological approach and show that valence and arousal of emotions in news articles can affect users' intention to share and predict content virality. Stories with positive and high-arousal emotions, such as awe, have a high probability to be shared by readers, whereas those with negative and low-arousal emotions, such as sadness, are less likely to go viral. Our paper differs from that of Berger and Milkman (2012) in at least three aspects. First, they measured diffusion by determining simply whether a news article is on the "most emailed list" of the New York Times without looking into the subsequent social transmission of the article. Instead, we investigate how social networks transmit online content with different emotions and focus on the diffusion processes and the cascades' structures, using a much larger-scale and richer sample. Second, Berger and Milkman (2012) take a mainly psychological approach to understanding the role of emotions in the intention of sharing online content. The diffusion of online content in social networks, however, is more than a psychological issue between a dyad. We thus use social network theories and network analyses to examine how articles with different emotions spread differently. We provide novel empirical evidence on the role of individual characteristics, weak ties, and social reinforcement in enabling or disabling the effects of emotions on online content diffusion in a massive-scale social network. Third, the main focus of their work is the valence and arousal of emotion. Our focus is the effects of discrete emotions, and we advance their work by finding that discrete emotions with similar valence and arousal (e.g., anger and anxiety, love and joy) still have distinct impacts on content diffusion.

It is valuable to understand the process and the structure of a cascade beyond a general sense of popularity, e.g., number of retweets, as different structural dimensions of a cascade provide distinct and meaningful implications to enable or disable information diffusion. Extant research on information diffusion measures the cascades in five dimensions: size (the total number of users involved in sharing the article), depth (the maximum number of sharing hops from the origin article, where a hop is a sharing action by a new unique user), maximum breadth (the maximum number of users involved in the cascade at any depth), time (the average number of hours that the cascading process takes to finish one level of depth), and structural virality (the average length of the shortest paths between all pairs of nodes in a diffusion tree) (Liben-Nowell and Kleinberg 2008, Goel et al. 2015, Vosoughi et al. 2018). Cascade time indicates the diffusion speed. Size indicates the general popularity of an article, without specifying its diffusion structure. Depth and maximum breadth are standard summary statistics of the structure of a cascade (Liben-Nowell and Kleinberg 2008). Intuitively, depth is the generations of a cascade, indicating the maximum degree of contacts that the information can reach from the root node. Information with a cascade of a higher level of depth is more likely to break into different social communities. Maximum breadth provides an intuition about how "broad" or "wide" the cascade is, which often results from broadcast structures. A "broad" but not "deep" cascade implies that the content is spread within the same but large social community. Structure virality quantifies the distinction between single broadcast and viral diffusion (Goel et al. 2015). Higher structure virality indicates that the cascade is driven more by decentralized and peer-to-peer sharing than by broadcasting. Broadcasting and viral diffusion are two typical alternative ways that enable information to reach large audiences (Van den Bulte et al. 2018). The cascade structures can be influenced by diffusion strategies, such as whether to utilize mass media and high-centrality users to broadcast content or use rewards to encourage peer-to-peer diffusion. Finally, how to promote an article or product in networks also correlates with potential cascade structures.

3. Data

To analyze emotions embedded in online articles and their diffusion cascades, we first randomly sampled 100,000 official accounts, the publishers on WeChat[1], then filtered out the inactive publishers who posted fewer than 10 articles during our observation period (August 31 to November 30, 2018), and, finally, sampled 38,839 publishers. We recorded the average number of followers of each publisher during our observation period (publisher's popularity), the average number of articles that the publishers posted per day (publisher's proactivity), and publisher type (i.e. individual, media[2], business enterprises, government or other organizations).

For each of these publishers, we randomly selected 10% of the articles that they posted during the study period and collected 387,486 online articles in total for analysis. This random sample well represents the articles distributed and diffused on WeChat and covers the topics as diverse as politics, economics, business, society, sports, technology, etc. The sample includes both short texts (fewer than 100 characters, similar to tweets) and long articles (more than 100,000 characters, similar to newspaper and magazine articles and chapters in a novel), with a mean article length of 1164.60 characters and a standard deviation of 2060.98 characters. Although the WeChat platform mainly publishes text, it allows articles to contain images and videos[3]. We also recorded the number of images and the number of videos in each article.

6,823,576 unique individuals shared these articles with their first-degree friends (i.e., strong-tie contacts), or their acquaintances (i.e., non-first-degree friends, weak-tie contacts)[4] on WeChat. We collected data on the demographic (i.e., age and gender) and network characteristics (i.e., network degree) of all the users involved in the cascades as well as the tie strength between every pair of sender and receiver (whether they were first-degree friends or not) in each cascade.

[1] 3.5 million active publishers generate, on average, more than 4.9 million articles per day on the WeChat official account platform, a feature of WeChat platform.

[2] Organizations with government permission to publish the content of newspaper, radio, and television.

[3] Our results are robust to the sample that excluded the articles containing mainly video but few words, which we will detail in Robustness Checks.

[4] Acquaintances (weak-tie contacts) are non-first-degree friends but are in the same group chats.

3.1. Mapping Cascades

Our dataset uniquely overcomes the cascade mapping problem, which commonly occurs in Twitter retweet cascades mapping. Retweet cascades are essentially inferred based on timing and relations (Shi et al. 2014, Vosoughi et al. 2018). Our context, however, enables us to precisely catalog the diffusion cascades of each article in our sample by recording the *exact* account ID of all individuals who shared the article as well as from whom they received access to the article and the timestamp of each sharing[5]. An article cascade starts when an article is posted to a publisher's followers, who then share the article in their local social networks. Users at the end of each cascade are those who shared the article, which was not shared by anyone again during the following week. On WeChat, if an article is not shared by anyone again during the following seven days, it is improbable (less than 1% probability) that the article would be shared again. Figure 1 shows an example of a large article cascade in our sample.

3.2. Measuring Cascades

As noted in Section 2.2, we measured the cascades in five dimensions: size, depth, maximum breadth, time, and structural virality (Goel et al. 2015, Vosoughi et al. 2018). These five measures are succinct representations of the shape and dynamics of a cascade process. Specifically, the structural virality is defined as Equation 1:

$$\text{Structural Virality} = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}, \tag{1}$$

where n represents the number of individuals involved in the cascades, and d_{ij} represents the shortest distance between individual i and individual j. Provided a same level of cascade size, the higher structural virality indicates that the cascade is driven more by decentralized and peer-to-peer sharing than by broadcasting.

Figure 2 (A-E) describes the empirical distributions of cascade size, depth, maximum breadth, structural virality and time of the articles in our sample, indicating that a very small fraction of the articles exhibit high values in every cascade dimension. To see the relationships among these cascades' dimensions, we present the correlation matrix in Table 1. We find a strong relationship between cascades' size and maximum breadth (0.974), as 41% of articles' cascades ended with the first level. Another exception is the correlation between cascades' depth and structural virality (0.921). To explain the correlation, first, as can be seen in Equation 1, structural virality is the average shortest distance between each pair of the nodes in a cascade. Second, the larger depth

[5] When the sender sends an article to a receiver, the link of the article (including article title and a short description) is displayed to the receiver. Such sharing behaviors are recorded in the platform's database. When the receiver opens the link and again shares the article, we catalog both the sender's and the receiver's ID.

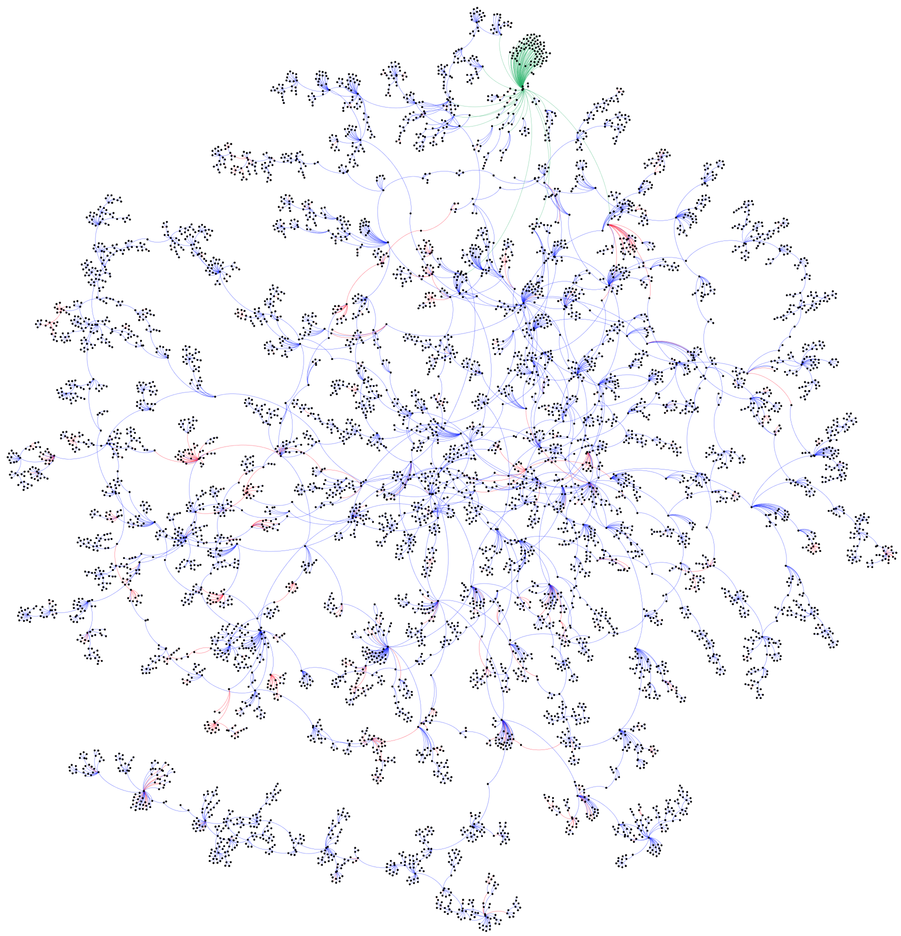

Figure 1 An example of a large cascade.

Note: The cascade involves 7,225 unique individuals (represented by nodes). Green edges represent sharing from the public account. Red and blue edges represent sharing from acquaintances (weak ties) or sharing from first-degree friends (strong ties). The article documented a scandal that involved certain individual bloggers and journalists who illegally blackmailed a number of real estate firms. The bloggers and journalists threatened to spread negative online articles about the firms unless they received a large payment.

indicates that the corresponding diagram of the cascade has a larger diameter (because depth is an inferior of diameter). Thus, we can expect a larger diameter to be positively correlated with the average shortest distance between each pair of nodes in the cascade. Moreover, due to the positive correlation between cascades' depth and size, structural virality is also positively correlated with cascades' size and maximum breadth. Despite the relatively high correlations, it is not trivial to include all the dimensions as our outcome variables, because not only different dimensions have distinct implications for us to understand diffusion, but also emotions can have disparate effects on these correlated dimensions, as we detail later.

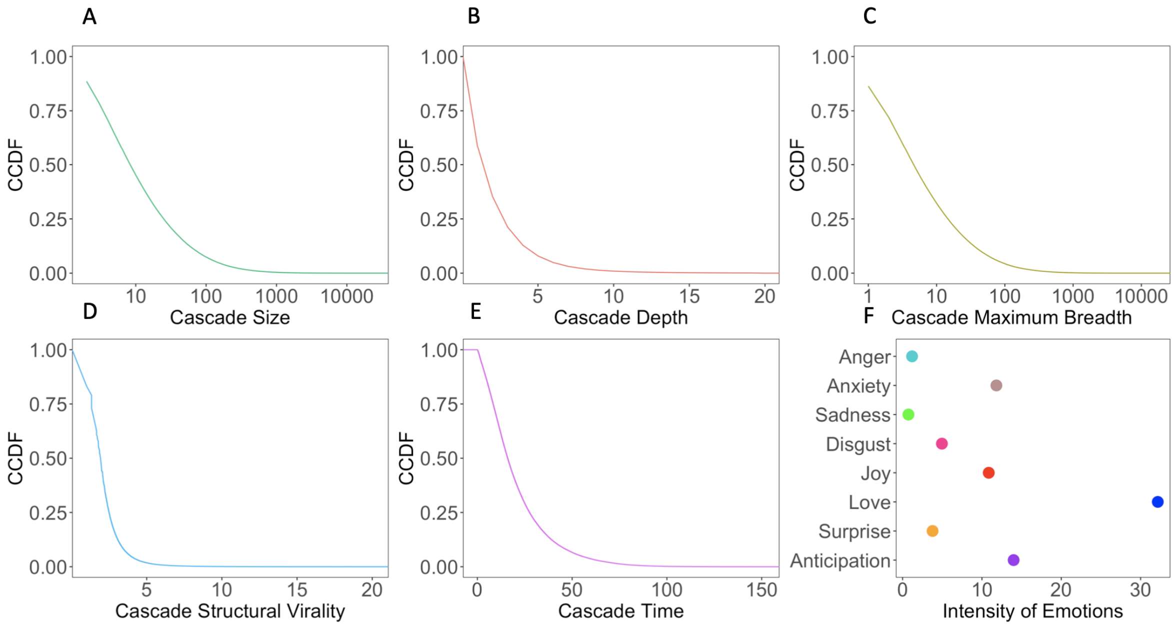

Figure 2 The complementary cumulative distribution functions (CCDFs) of cascade scales and the level of intensity of eight emotions for an average article.

Note: (A-E) The CCDF of cascade size, depth, maximum breadth, structural virality, and time. The CCDF is the complement of the cumulative distribution function (CDF). Take cascade size as an example: for a given level of cascade size in the x axis, the corresponding value of the CCDF shows the fraction of articles whose cascade size is above or equal to this level. (F) The level of intensity of eight emotions for an average article. The values are directly from our emotion detection analysis, and have not been normalized to a standard distribution (so that we can compare the mean values of different types of emotions). Differences between all pairs of the eight emotions are statistically significant ($ps < 0.001$).

Table 1 Correlations among cascade Dimensions[1]

	Depth	Size	Breadth	Time	Structural Virality
Depth	1.000				
Size	0.720	1.000			
Maximum Breadth	0.588	0.974	1.000		
Time	0.157	0.175	0.173	1.000	
Structural Virality	0.921	0.743	0.616	0.146	1.000

[1] *Cascades' size positively correlated with maximum breadth and depth. Structural virality positively correlated with cascades' depth, size, and maximum breadth.*

4. Emotion Detection

Emotion is inherently ambiguous, subjective, and dynamic over time. It is often difficult to reach an agreement on document-level emotion annotation (Quan and Ren 2010), which hinders training a document-level supervised learning model. We therefore adopted a state-of-art lexicon-generation approach (Xue et al. 2014, Yu et al. 2019) and constructed a new domain-specific and up-to-date emotion lexicon based on a general emotion lexicon with 16,017 unique emotion words (Quan and

Ren 2010) and 8 million word-embedding vectors (Song et al. 2018). Because we start from the word-level, our approach mitigates the document-level ambiguity.

4.1. Emotion Lexicon Construction

First, we use an existing emotion lexicon, Ren-CECps (Quan and Ren 2010), as a basic lexicon. Each word w_i in the lexicon is mapped to an eight-dimension vector $v_i = (I_1^i, I_2^i, ..., I_8^i)$, where $I_k^i \in [0,1]$ is manually annotated and represents the intensity of the kth discrete emotions expressed by w_i. Second, we retrieve word vectors that contain words' semantic information. The word vectors can be derived by statistical language modeling (e.g. Word2Vec by Mikolov et al. (2013)). We used pre-trained word vectors by Song et al. (2018), who provides 200-dimension word vectors for over 8 million common Chinese words and phrases. These word vectors are pretrained on up-to-date, large-scale, and high-quality Chinese online content, and have been validated through various natural language processing tasks (Song et al. 2018). Then, the similarity between the two words can be measured by the cosine distance of the two corresponding word vectors (Mikolov et al. 2013). Third, we follow the approach proposed by Xue et al. (2014) and Yu et al. (2019) to extend the basic lexicon to a domain-specific and up-to-date lexicon. For every emotion word in the basic lexicon, we use 8 million word vectors to mine 100 most similar words[6] as potential emotion words. For each potential emotion word, we use its semantically nearest words in the basic lexicon to determine whether it is an emotion word and its intensities for eight types of discrete emotions. Fourth, we combine the newly mined emotion words with words in the basic lexicon to be an extended lexicon. We then treat the extended lexicon as a new "basic lexicon," which is used to repeat the third step until the total word amount in the combined lexicon converged (See Appendix A for more details). 16,921 new words were found after this process, and the extended lexicon contains totally 28,969 words. This result confirms the necessity of constructing a domain-specific lexicon, without which 58.4% of unique emotion words (16,921 out of 28,969) would be ignored if only the basic lexicon were used. Five raters were recruited to annotate the emotional intensities of the newly mined words for the eight discrete emotions. There is no statistically significant difference between the results generated by the algorithms and by the raters, confirming the validity of our new lexicon (See Appendix B).

4.2. Article-level Emotion Detection

We analyzed article-level discrete emotions by detecting these emotion words in the text, and also considered the negation and degree words (if any) associated with the emotion words. In accordance with the document-level emotion expression space model (Quan and Ren 2010), we

[6] This parameter is suggested by Xue et al. (2014).

map each online article into an eight-dimensional vector. Each element of the vector represented the emotional intensity of its corresponding discrete emotion expressed in the article d:

$$d = (e_1, ..., e_k, ..., e_8), \tag{2}$$

where e_k is determined by emotion, negation, and degree words contained in the article. Negation and degree words are frequently used in Chinese and thus are helpful for accurate emotion analysis (Quan and Ren 2010). We adopted a negation word dictionary provided by TextMind, a Chinese language psychological analysis system developed by the Chinese Academy of Sciences. The dictionary contains 31 frequently used negation Chinese words. We used 60 degree words provided by Ren-CECps as our degree word dictionary and annotated their degree values. For example, the degree value of the word meaning "the most" in Chinese is annotated as 1.5, and that of the word meaning "kind of" is annotated as 0.8.

We use a forward-sliding window (with the window size set to three words) to capture these negations and degree words. If our algorithm finds an emotion word in the article, it checks the three words before the emotion words and captures any negation and degree words that appear within these three words. We choose three as sliding-window size because normally in Chinese, negative and degree words are within three words before the corresponding emotion words. The kth discrete emotion intensity of the article d, denoted as $e_k(d)$, is determined as follows:

$$e_k(d) = \Sigma_{i=1}^{n} (-1)^{m_i} \times DegV_i \times I_k(w_i), \tag{3}$$

where $\{w_i\}_{i=1}^{n}$ are the emotion words both in article d and in our lexicon. It expresses the kth discrete emotion. If $n = 0$, then $e_k(d)$ will be set to zero. $I_k(w_i)$ refers to the kth discrete emotion intensity of w_i ($k \in \{1, 2, \ldots, 8\}$). m_i is the total number of negative words that appeared in the sliding window of w_i. Finally, $DegV_i$ is the average degree value of all degree words that appeared in the sliding window of w_i. The emotions of each article are then summarized into an eight-dimensional vector, as shown in Equation 2. We used the same procedure and analyzed eight discrete emotions in the comments.

4.3. Descriptive Results

We present descriptive results on the article-level emotions. First, we find that love is the most expressed emotion, whereas sadness is the least expressed. Among all the negative emotions, anxiety has the highest level of emotion intensity (See Figure 2 (F)). These raw discrete emotion intensities are in different scales (i.e., with different population means and variances). To enable the comparison of different discrete emotions under the same scale, we standardize these raw discrete emotion variables to z-scores. A z-score is calculated by subtracting the population mean from

Table 2 Correlations among Emotions[1]

	Anger	Anxiety	Anticipation	Disgust	Joy	Love	Surprise	Sadness
Anger	1.000							
Anxiety	0.190	1.000						
Anticipation	−0.010	0.109	1.000					
Disgust	0.435	0.358	−0.015	1.000				
Joy	−0.065	−0.001	0.081	−0.092	1.000			
Love	−0.138	−0.153	0.070	−0.167	0.294	1.000		
Surprise	0.139	0.454	−0.011	0.247	0.007	−0.103	1.000	
Sadness	0.056	0.178	−0.009	0.058	0.018	−0.035	0.108	1.000

[1] *The absolute values of most correlations are below 0.20, with only a few exceptions that exist between 0.20 and 0.46. The results indicate the independence of these eight emotions.*

an individual raw score and then dividing the difference by the population standard deviation. The standardization makes the intensities of discrete emotions fluctuate around a zero mean and measured on a scale of 1 standard deviation (close to the standard normal distribution) (Bollen et al. 2011). Thus, if an article has a 0.0 intensity score in anxiety, it would contain more anxious expressions (emotion words and relevant degree words) than about 50% of the articles in our sample; 1.0 means more than about 84%; 2.0 means more than about 97.5%, and so forth. For short, we refer the normalized intensity scores as intensity score hereafter. Second, we present the correlation matrix of the eight discrete emotions in Table 2 to check their independence. As can be seen, the absolute values of most correlations are below 0.20, with only a few exceptions that are between 0.20 and 0.46. The results indicate the independence of these eight emotions.

5. Multilevel Model

5.1. Main Analysis

To analyze the effects of emotions in information diffusion, we applied a two-level hierarchical regression model (Hox et al. 2017), commonly used in estimating the effects of online content characteristics (Brady et al. 2017, Gan et al. 2017). Articles are nested within publishers. Articles by the same publishers may share common characteristics (e.g., topics, writing styles) and may not be independent. Thus, the independence assumption of using an ordinary least squares regression model would be violated. In contrast, a two-level hierarchical regression model addresses this issue by considering both between- and within-group (i.e., publisher) variances (Hox et al. 2017). The two-level hierarchical regression model allows us to comprehensively and effectively control for heterogeneity at both the article and publisher levels.

To control for the article characteristics, we first estimated a latent Dirichlet allocation (LDA) model (Blei et al. 2003) based on all sampled articles and controlled for the latent topic distribution of an article by a 30-dimension vector. More specifically, first, we used Chinese word segmentation repository by Python and cut titles and content into words. Second, we filtered out all emotion words (in our constructed lexicon) from these titles and content. By doing so, topic variables and emotion variables are constructed by different linguistic features, which reduces the possibility that the two sets of variables would confound each other. We also filtered out low-frequency words (i.e. words that appear in only less than 0.1% of the articles) to reduce noise in model estimation. Third, we trained an LDA model and specified the optimal number of topics that minimizes the model perplexity (Blei et al. 2003). When the number of topics is 30, the perplexity is minimized (Figure C.1 in Appendix C). Finally, we used the trained LDA model to map the titles and content of each article to a 30-dimension vector as topic distribution, with each element representing the probability that the article belongs to the corresponding latent topic category.

We also included article length (log-transformed number of characters), number of images and videos embedded in the article (media richness), whether the article was posted during a weekend (time effect), and the number of comments in our model. To control for publisher characteristics, we included the variables, such as the average number of followers of each publisher during our observation period (publisher's popularity), the average number of articles that the publishers posted per day (publisher's proactivity), and publisher type (i.e. individual, media, business enterprises, government or other organizations). The model is shown as follows:

$$\begin{aligned} &\text{Level 1: } y_{ij} = \beta_{0i} + \beta_1^{'}\mathbf{emotion}_{ij} + \boldsymbol{\beta}_2^{'}\mathbf{article}_{ij} + \epsilon_{ij}; \\ &\text{Level 2: } \beta_{0i} = \gamma_0 + \boldsymbol{\gamma}_1^{'}\mathbf{pubacc}_i + \mu_i, \end{aligned} \tag{4}$$

where level 1 indicates the article level and level 2 indicates the publisher level. y_{ij} indicates one of the five cascades' dimensions of the jth article published by the ith publisher, $\mathbf{emotion}_{ij}$ indicates the intensity scores of (discrete) emotions of the jth article published by the ith publisher, $\mathbf{article}_{ij}$ indicates the article-level control variables of the jth article published by the ith publisher, and $\mathbf{pubacc}_i$ indicates the publisher-level control variables of the ith publisher. We use β_1 to capture the effect of emotion, $\boldsymbol{\beta}_2$ to capture the effects of article-level control variables, and $\boldsymbol{\gamma}_1$ to capture the effects of publisher-level control variables. All of the error terms ϵ_{ij} and μ_i are assumed to be independent and identically normally distributed with a mean of zero.

5.2. Mediation Analysis

Based on the main analysis, we further perform mediation analysis to unveil the mechanisms behind the main effects, following the widely-used procedures established by Baron and Kenny (1986). As we detail in Section 6.2, we tested the mediating effects of the average age, gender composition, average network centrality (measured by users' average friend number), and social relations (measured by the proportion of weak ties) among the users in a cascade, respectively, on the relationships between each discrete emotion and each cascade dimension.

Our mediation analysis consists of three regression equations. The first step is to investigate the effects of eight discrete emotions on each mediator by regressing each mediator (average age, average friend number and the proportion of weak ties among all the users involved in a cascade) on eight discrete emotions with control variables to test which emotions reveal significant relationships with the mediator. We utilize the two-level hierarchical regression model shown in equation 5.

$$\begin{aligned} \text{Level 1: } Mediator_{ij} &= \beta_{0i} + \boldsymbol{\beta}_1^{'}\textbf{emotion}_{ij} + \boldsymbol{\beta}_2^{'}\textbf{article}_{ij} + \epsilon_{ij}; \\ \text{Level 2: } \beta_{0i} &= \gamma_0 + \boldsymbol{\gamma}_1^{'}\textbf{pubacc}_i + \mu_i, \end{aligned} \tag{5}$$

where $Mediator_{ij}$ indicates the mediator of the cascade of the jth article published by the ith publisher.

The second step is to measure the indirect effects[7] of emotions that exhibit significant relationships with the mediator in the first step by regressing five cascade dimensions respectively on all of the independent variables, including degrees of eight discrete emotions and control variables, which is exactly the same procedure as in our main model (Equantion 4) to measure the effects of emotions on cascades. The third step is to measure the direct effects[8] of emotions that are significantly correlated with the mediator by regressing each of the five cascade dimensions on all of the independent variables together with the mediator (users' average age, users' network centrality, and users' social relations).

$$\begin{aligned} &\text{Level 1: } y_{ij} = \beta_{0i} + \boldsymbol{\beta}_1^{'}\textbf{emotion}_{ij} + \boldsymbol{\beta}_2^{'}\textbf{article}_{ij} + \beta_3 Mediator_{ij} + \epsilon_{ij}; \\ &\text{Level 2: } \beta_{0i} = \gamma_0 + \boldsymbol{\gamma}_1^{'}\textbf{pubacc}_i + \mu_i, \end{aligned} \tag{6}$$

where we use β_3 to capture the effect of the mediator.

We then compared the coefficients of the emotions significantly associated with both the cascade dimensions and the mediator measured in the third step with those measured in the second step to investigate the mediating effect of the mediator. Partial mediation effect is exhibited if the coefficients of emotions measured in the third step are smaller in absolute value than are the coefficients of emotions measured in the second step, whereas a complete mediation effect is exhibited if the coefficients of emotions measured in the third step become insignificant. We show the results of these three mediators (users' average age, average network centrality, and the proportion of weak ties, in a cascade) in Section 6.2. We also investigated the mediation effects of gender, which, however, are not significant.

[7] Here indirect effect refers to the effect of the emotion before separating the effect of the mediator.

[8] Here direct effect refers to the effect of emotion after separating the effect of mediator.

Table 3 Multilevel Analysis Results: Degree of Emotion [1]

	Depth	Size[2]	Breadth[2]	Time	SV[6]
Degree of Emotion	0.037∗∗∗ (0.005)[5]	0.022∗∗∗ (0.002)	0.020∗∗∗ (0.002)	0.116∗∗∗ (0.003)	0.017∗∗∗ (0.003)
$\boldsymbol{\sigma_{\mu_i}}$[3]	1.087∗∗∗ (0.050)	0.807∗∗∗ (0.031)	0.834∗∗∗ (0.031)	9.906∗∗∗ (1.072)	0.658∗∗∗ (0.012)
Controlled[4]					

[1] *Significance level:* * $p<0.05$, ** $p<0.01$, *** $p<0.001$

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] $\boldsymbol{\sigma_{\mu_i}}$ *indicates between-publisher standard deviation.*

[4] *All of the article-level variables and publisher-level variables are controlled for.*

[5] *Standard error of estimation.*

[6] *SV indicates structural virality.*

6. Results

6.1. Effects of Emotions on Online Article Diffusion

We first fit the regression model to investigate the relationship between an article's degree of emotion in general and its cascade process. An article's degree of emotion is operationalized as the summation of the intensity of the eight discrete emotions, which is then standardized to z-scores. Articles with a higher level of emotion are associated with statistically significantly greater expected cascades' size ($\beta^{size}_{emotion} = 0.022$, $p < 0.001$), depth ($\beta^{depth}_{emotion} = 0.037$, $p < 0.001$), maximum breadth ($\beta^{breadth}_{emotion} = 0.020$, $p < 0.001$), and structural virality ($\beta^{sv}_{emotion} = 0.017$, $p < 0.001$), but lower expected cascades' speed ($\beta^{time}_{emotion} = 0.116$, $p < 0.001$) (See Figure 3 and Table 3). These suggest that a 1-standard-deviation (or 1.0 unit of intensity score) increase in the degree of emotion leads to, on average, 2.2% increase in size[9], a 0.037-level increase in cascade depth, 2.0% increase in maximum breadth, a 0.017-unit increase in structural virality (indicating that the cascade contains more peer-to-peer sharing and fewer broadcasting structures), and 0.116 more hours to reach the next level (i.e., lower cascade speed). In our sample, the maximum emotion intensity of an article is 266 units, which can make the article's diffusion cascade 585.2% larger, 9.842 levels deeper, 532.0% broader, and 4.522 more viral but 30.856 hours slower to reach the next level based on our model prediction. This confirms the significantly positive effects of emotion (embedded in content) in information cascading of online articles, although the higher degree of emotion is associated with slower diffusion speed. In addition, the between-publisher standard deviation σ_{μ_i} is statistically significantly larger than zero, indicating that the hierarchical model is preferable to a non-hierarchical model, in which σ_{μ_i} will be assumed as zero and the results will be biased.

[9] The coefficients of size and maximum breadth are interpreted as percentages, as we used log-transformation on the two dependent variables due to their right-skewed distributions.

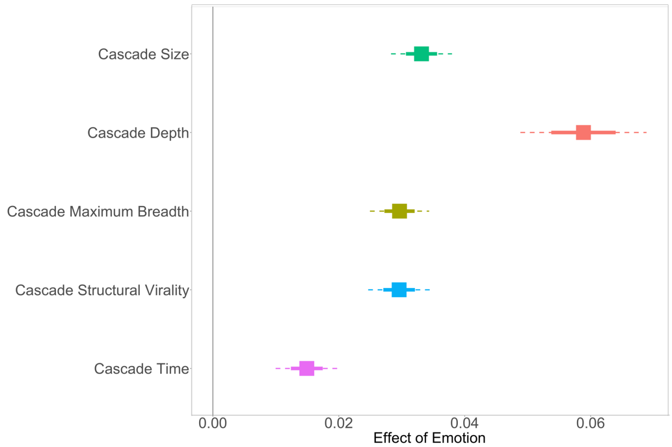

Figure 3 **The relationships between degree of emotion embedded in content and article cascade Size, Depth, Maximum breadth, Structural virality, Time.**

Note: The x axis displays the size of model coefficients. The y axis displays five cascades' dimensions. The squares in the middle of the intervals represent mean values. The intervals covered by dashed lines represent 95% confidence intervals. The intervals covered by solid lines represent mean values plus or minus 1 unit of their standard deviations.

We then followed the same method to investigate the relationship between eight discrete emotions embedded in content and the articles' cascade process (See Table 4). Figure 4 shows that three discrete emotions (i.e., anxiety, love and sadness) play significant roles in article diffusion. Articles with a higher degree of anxiety and love embedded in content exhibit significantly greater expected cascade size ($\beta^{size}_{anxiety} = 0.042$, $p < 0.001$; $\beta^{size}_{love} = 0.019$, $p < 0.001$), depth ($\beta^{depth}_{anxiety} = 0.046$, $p < 0.001$; $\beta^{depth}_{love} = 0.028$, $p < 0.01$), breadth ($\beta^{breadth}_{anxiety} = 0.042$, $p < 0.001$; $\beta^{breadth}_{love} = 0.017$, $p < 0.001$), and structural virality ($\beta^{sv}_{anxiety} = 0.022$, $p < 0.001$; $\beta^{sv}_{love} = 0.018$, $p < 0.01$). In contrast, those with a higher degree of sadness are associated with smaller expected cascades' size ($\beta^{size}_{sadness} = -0.031$, $p < 0.001$), depth ($\beta^{depth}_{sadness} = -0.024$, $p < 0.001$), breadth ($\beta^{breadth}_{sadness} = -0.033$, $p < 0.001$), and structural virality ($\beta^{sv}_{sadness} = -0.013$, $p < 0.001$). Articles with a higher degree of sadness and surprise exhibit a faster expected cascade speed ($\beta^{time}_{sadness} = -0.175$, $p < 0.001$; $\beta^{time}_{surprise} = -0.176$, $p < 0.01$), whereas those with a higher degree of love are associated with a slower expected cascade speed ($\beta^{time}_{love} = 0.347$, $p < 0.001$). Articles with a higher degree of anticipation are associated with a larger expected maximum breadth ($\beta^{breadth}_{anticipation} = 0.012$, $p < 0.01$), whereas those with a larger degree of joy exhibit smaller expected maximum breadth ($\beta^{breadth}_{joy} = -0.013$, $p < 0.01$). Therefore, anxiety, love and sadness are significantly associated with content spread in almost all cascade dimensions (except anxiety and cascade speed), after holding constant the heterogeneity of article-level and publisher-level characteristics.

Our results also suggest that the emotion pairs, similar in valence and arousal (i.e., anxiety and anger, love and joy), exhibit disparate effects on article diffusion. This result empirically

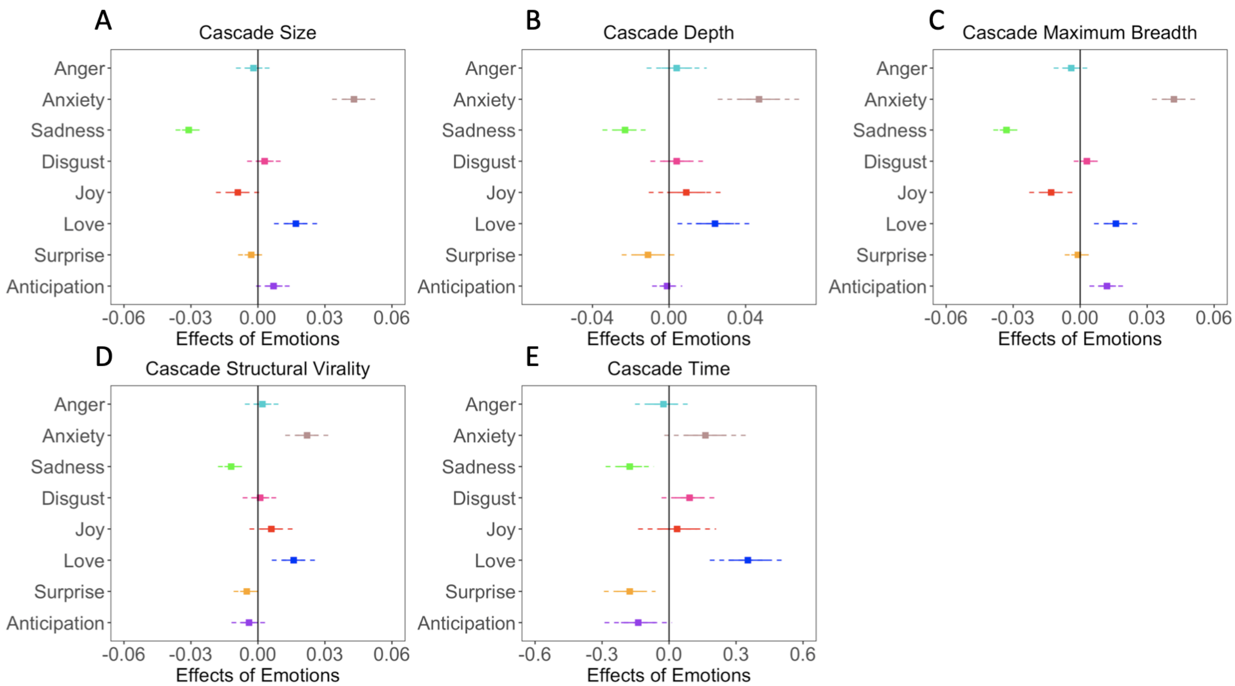

Figure 4 The relationships between eight discrete emotions embedded in content and article cascade (A) Size, (B) Depth, (C) Maximum breadth, (D) Structural virality, (E) Time.

Note: The x axis displays the size of model coefficients. The y axis displays eight discrete emotions. The squares in the middle of the intervals represent mean values. The intervals covered by dashed lines represent 95% confidence intervals. The intervals covered by solid lines represent mean values plus or minus 1 unit of their standard deviations.

pits discrete emotions against valence and arousal perspectives (dimensional theory of emotion), confirming that discrete emotions better delineate human emotions. This finding is not surprising in theory as well. Anxiety and anger can induce different cognitive appraisals of online content, which is associated with readers' tendency to share. Online content with anxiety is often perceived as more helpful and valuable, because anxiety indicates a higher level of writer's cognitive efforts (Yin et al. 2014). In contrast, anger leads to a lower level of perceived rationality of writers (Xiao et al. 2018). The perceived helpfulness of content may increase readers' intentions to share (Cheung et al. 2013). Compared to joy, love is associated with greater trust and acceptance (Plutchik and Kellerman 1980), and thus inspires individuals' intention to help and share with others (Henderson 2012). Anxiety and love therefore can be more likely to increase the spreading of online content in users' local social networks than anger and joy.

Further, in our hierarchical regression model, the publisher-level control variables may be insufficient to fully solve the endogeneity problem caused by omitted publisher characteristics. We thus run a publisher-level fixed effects panel regression model which accounts for potential correlations of $\mathbf{emotion}_{ij}$ and $\mathbf{article}_{ij}$ with unobserved publisher-specific features. The results of the fixed

Table 4 Multilevel Analysis Results:Discrete Emotions[1]

	Depth	Size[2]	Breadth[2]	Time	SV[5]
Emotions					
Anger	0.002	−0.002	−0.005	−0.024	0.002
	(0.008)[6]	(0.004)	(0.004)	(0.065)	(0.004)
Anxiety	0.046∗∗∗	0.042∗∗∗	0.042∗∗∗	0.162	0.022∗∗∗
	(0.011)	(0.005)	(0.005)	(0.094)	(0.005)
Sadness	−0.024∗∗∗	−0.031∗∗∗	−0.033∗∗∗	−0.175∗∗	−0.013∗∗∗
	(0.006)	(0.003)	(0.003)	(0.055)	(0.003)
Disgust	0.006	0.003	0.004	0.090	0.001
	(0.007)	(0.004)	(0.003)	(0.064)	(0.004)
Joy	0.011	−0.008	−0.013∗∗	0.034	0.007
	(0.010)	(0.005)	(0.005)	(0.089)	(0.005)
Love	0.028∗∗	0.019∗∗∗	0.017∗∗∗	0.347∗∗∗	0.018∗∗∗
	(0.010)	(0.005)	(0.005)	(0.087)	(0.005)
Surprise	−0.012	−0.003	−0.001	−0.176∗∗	−0.006
	(0.007)	(0.003)	(0.003)	(0.059)	(0.003)
Anticipation	0.000	0.008	0.012∗∗	−0.139	−0.003
	(0.009)	(0.004)	(0.004)	(0.077)	(0.004)
(Intercept)	1.879∗∗∗	1.821∗∗∗	1.249∗∗∗	14.442∗∗∗	1.707∗∗∗
	(0.051)	(0.026)	(0.025)	(0.442)	(0.025)
$\boldsymbol{\sigma_{\mu_i}}$[4]	1.088∗∗∗	0.809∗∗∗	0.836∗∗∗	9.898∗∗∗	0.660∗∗∗
	(0.051)	(0.031)	(0.032)	(1.051)	(0.012)
Controlled[3]					

[1] *Significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] *All of the article-level variables and publisher-level variables are controlled for.*

[4] $\boldsymbol{\sigma_{\mu_i}}$ *indicates between-publisher standard deviation.*

[5] *SV indicates structural virality.*

[6] *Standard error of estimation.*

effects model are qualitatively consistent with the hierarchical model that uses random effects at the publisher level (See Table D.1 in Appendix D). This confirms that our hierarchical model effectively controls the publisher characteristics. We further run the Durbin-Wu-Hausman Test (Hausman 1978), and and the results (See Table D.2 in Appendix D) show that our random-effects models are consistent. Given that they are both consistent and efficient, the hierarchical models based on random effects are the choice of our analyses.

6.2. Unveiling Mechanisms

One may hypothesize that the characteristics of the individuals involved in a cascade, weak ties or the mechanism of social reinforcement may explain why articles with a higher level of anxiety, love, and sadness diffuse significantly differently in the online social network. To test this hypothesis, we perform mediation analysis (Baron and Kenny 1986), by analyzing the demographic and network characteristics of more than 6 million users involved in the cascades and the social ties between

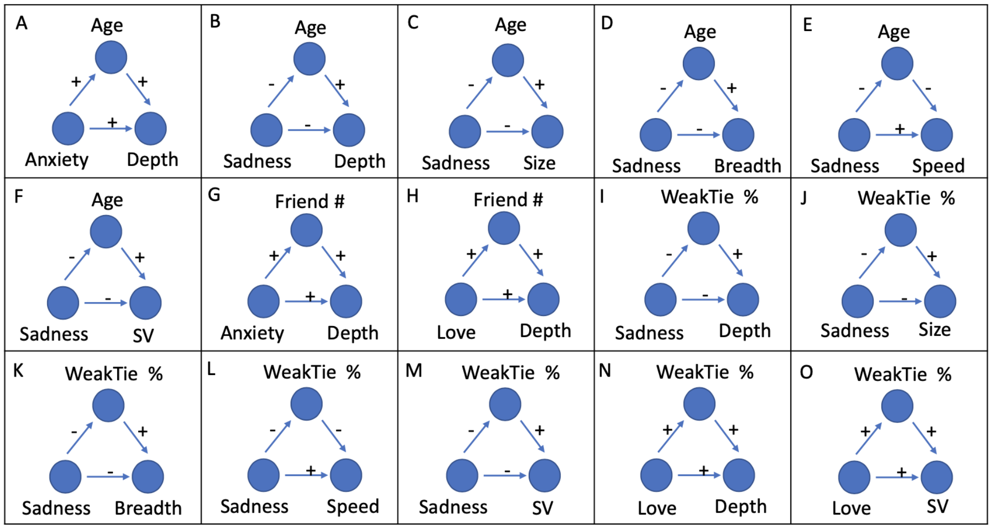

Figure 5 **Mediating effects of individuals' age and network centrality (friend #), and weak tie proportion in a cascade on how discrete emotions affect the diffusion of online articles.**

Note: (A-F) Mediating effect of age. (G-H) Mediating effect of network centrality. (I-O) Mediating effect of weak tie proportion. They show that how age and network centrality of users involved in a cascade, and the weak tie proportion in the cascade, significantly mediate the effects of anxiety, love, and sadness on cascade size, depth, breadth, structural virality, and speed. We only include the significant mediating effects. Note that '-' indicates a significant negative effect, and '+' indicates a significant positive effect. SV represents structural virality.

them. We find significant mediating effects of the average age, average network centrality (measured by users' average friend number), and social relations (measured by the proportion of weak ties) among the users in a cascade, respectively, on the relationships between each discrete emotion and each cascade dimension. The correlations among these three mediators are relatively small (less than 0.13), which guarantees the independence of these mediators. See Figure 5 for the summary of findings.

We first investigated whether articles with different embedded emotions were shared by individuals with different ages, leading to differential information cascades. Age difference can result in different attitudes toward different emotions. For example, younger individuals can be more susceptible to negative low-arousal emotions, such as sadness (Kensinger 2008). Prior research also suggests that older users are more influential than are younger users in decision making (Aral and Walker 2012, Leatherdale et al. 2005). Thus, the age of the users, involved in the cascades, may affect how the information transmitted. Through regression analysis, we find that anger, anxiety, disgust, joy, love, and surprise are significantly and positively associated with the age of the users in the cascade (a 1-unit increase in the intensity score of anger, anxiety, disgust, joy, love and surprise is associated with, respectively, a 0.17-, 0.21-, 0.28-, 0.18-, 0.11-, 0.14-year increase in the users' average age, $ps < 0.001$), whereas the effects of sadness and anticipation are significantly

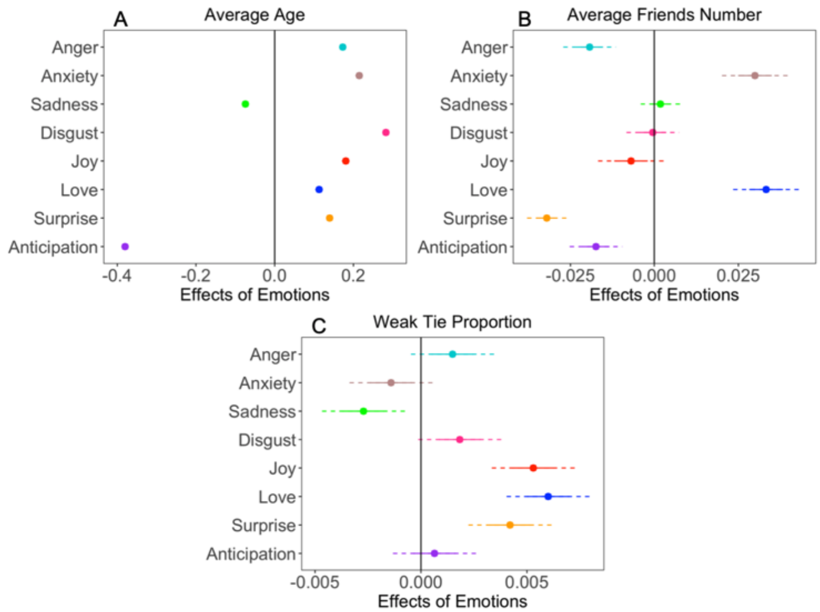

Figure 6 **The relationships between eight discrete emotions embedded in content and mediators : (A) Average age of users, (B) Average friends number of users, and (C) Weak tie proportion of users in the cascades.**

Note: The x axis displays the size of model coefficients. The y axis displays eight discrete emotions. The squares in the middle of the intervals represent mean values. The intervals covered by dashed lines represent 95% confidence intervals. The intervals covered by solid lines represent mean values plus or minus 1 unit of their standard deviations.

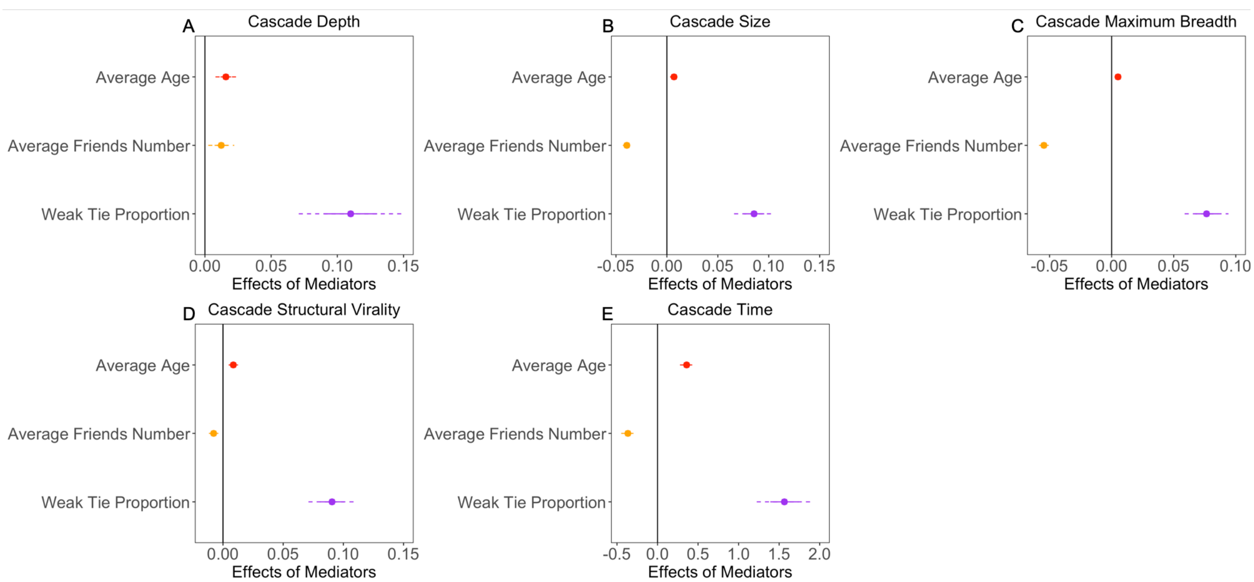

Figure 7 **The relationships between three mediators and article cascade (A) Size, (B) Depth, (C) Maximum breadth, (D) Structural Virality, (E) Time.**

Note: The x axis displays the size of model coefficients. The y axis displays three mediators. The squares in the middle of the intervals represent mean values. The intervals covered by dashed lines represent 95% confidence intervals. The intervals covered by solid lines represent mean values plus or minus 1 unit of their standard deviations.

negative (a 1-unit increase in the intensity score of sadness and anticipation is associated with, respectively, a 0.07- and 0.38-year decrease in the users' average age, $ps < 0.001$) (See Figure 6A, and Table E.1 in Appendix E). We further find that older users are associated with deeper, larger, broader, and more viral but slower cascades (i.e., an increase of one year in users' average age in a cascade leads to a 0.016-level increase in cascade depth, 0.7% increase in size, 0.5% increase in maximum breadth, and a 0.009-unit increase in structural virality but 0.036 more hours to reach the next level. See Figure 7, and Table E.1 in Appendix E.). Our results of mediation analysis suggest that articles with a greater degree of anxiety were spread more among older users, which partially explains why they spread more deeply in the network. In contrast, articles with greater embedded sadness were shared more by younger users, which partially explains the lower levels of size, depth, breadth, structural virality, and speed associated with these articles. Further, although prior research suggests that emotional reactivity can work differently across individuals with different genders (Birditt and Fingerman 2003), we did not find statistically significant mediation effects of gender in how the emotions affect the articles' cascades.

We then explored the role of individuals' network centrality, measured by the average friend number in the observation period (Jackson 2008), in the relationship between emotions and article cascades. Empirical evidence suggests that users' network centrality can affect the decisions of whether to share online content with different characteristic (Figueiredo et al. 2015) and lead to larger information cascades (Banerjee et al. 2013). Individuals with greater network centrality often have higher social status and exert increased social influence in adoption decisions (Hu and Van den Bulte 2014, Huang et al. 2020). Users with more friends can broadcast information to more contacts, which results in larger but less viral cascades, by adding broadcast structures (Goel et al. 2016). Through a regression analysis, we find that anxiety and love significantly and positively associate with the average number of friends of the users in the cascade (a 1-unit increase in the intensity score of anxiety and love is associated with, respectively, a 0.030- and 0.033-unit increase in the users' average number of friends, $ps < 0.001$), whereas the effects of anger, surprise and anticipation are significantly negative (a 1-unit increase in the intensity score of anger, surprise and anticipation is associated with, respectively, a 0.02-, 0.03- and 0.02-unit decrease in the users' average friend number, $ps < 0.001$) (See Figure 6B, and Table E.2 in Appendix E.). Our results suggest that a larger average friend number of the users involved in a cascade is associated with deeper, and faster but less viral cascades (i.e., one unit increase in the users' average friend number can link to 0.01 levels increase in depth, 0.37 hours less for the cascade to reach the next depth, and 0.008 unit decrease in structural virality, $ps < 0.001$). Further, our mediation analysis suggests that the articles with more embedded anxiety and love tend to be shared by more central users, which partially explains the effects of these two discrete emotions on the articles' cascade depth (See Figure 7, and Table E.2 in Appendix E).

We further investigated whether articles with different embedded emotions spread through different social ties (i.e., strong or weak ties), which enable or disable information diffusion. Weak ties are more likely to be local bridges in a social network than are strong ties and, therefore, can facilitate information transmission across communities and enable large information cascades by bringing novel information to more communities (Granovetter 1977). However, the smaller bandwidth of weak ties may indicate slower information processing, which can reduce transmission speed through weak ties (Aral and Van Alstyne 2011, Bruggeman 2016). Indeed, we find that the cascades with a larger percentage of weak ties (i.e., articles are transmitted between acquaintances instead of between first-degree friends) are associated with higher expected size, depth, maximum breadth and structural virality but smaller expected speed (i.e., a 1-unit increase in weak tie proportion in the cascade is associated with a 8.56% increase in size, 0.11-level increase in depth, 7.65% increase in maximum breadth, a 0.09-unit increase in structural virality and 1.56 hours more to reach the next level of depth, $ps < 0.001$) (See Figure 7, and Table E.3 in Appendix E). Further, individuals often express different emotions to close friends, acquaintances, and strangers. For example, we may share our negative emotions, such as sadness and anxiety, only with those we trust. Using a regression analysis, we find that joy, love, and surprise are statistically positively associated with the percentage of weak ties in the cascade (a 1-unit increase in intensity score in joy, love, and surprise is associated with respectively a 0.53%, 0.60% and 0.42% increase in weak tie proportions in the cascade, $ps < 0.001$), whereas the effect of sadness is negative (a 1-unit intensity score increase in sadness is associated with a 0.27% decrease in weak ties in the cascade, $p < 0.01$) (See Figure 6C). Articles with a greater degree of sadness were transmitted more through strong ties, which completely explains its negative effects on cascade depth and structural virality, and partially explains its negative effects on cascade size, breadth and speed. Articles with a greater degree of love transmitted more through weak ties, which completely explains its effect on cascade depth, and partially explains its effect on structural virality. These results also suggest that weak (strong) ties play a more important role in shaping the differential diffusion patterns of articles with sadness and love than do the age and network centrality of the users in the cascade.

Finally, we tested whether social reinforcement (or complex contagion, the idea that "people usually require contact with multiple sources of infection before being convinced to adopt a behavior.") (Centola 2010) mediates the effects of emotions on online artcile diffusion. We first test whether the number of transmission sources, which is proxied by the clusterness of the seed users of an article, affect article diffusion. The clusterness of the seed users is defined as ratio of the number of friend pairs to the number of all possible friend pairs in seed users[10]. The seed users are those who initially shared the articles from the publishers to their friends. Two individuals connected by a strong tie tend to have more common friends than do those connected by a weak tie or those who disconnected[11] (Granovetter 1977). Thus, the more friend pairs that are among the seed users of an article, the more users in the next level tend to be the common friends of at least two seed users and therefore be exposed to the article multiple times. As a result, the clusterness of the seed users should positively correlate with the number of contacts from whom a user in the next level can receive an article. We find, however, that the clusterness of seed users did not affect any cascade scale or speed at the next level, and thus it does not mediate the effect of the emotions on how articles cascade. Our results thus suggest no significant difference between one and multiple transmission sources in enabling information diffusion with different emotions, and it is likely that online article diffusion follows a simple contagion.

7. Robustness Checks

A battery of tests validated our results and confirm their robustness. First, although emotions were expressed and detected in the content of articles, it is still uncertain whether and how readers received these emotions. We therefore analyzed the emotions expressed in the articles' comments. We first selected out the articles that expressed an extreme emotion. In particular, for each of the eight discrete emotions, we selected articles with only one of the emotion intensity scores as higher than 1.96 and others as lower than 1.96. Because all emotion intensity scores were standardized as z-scores, the threshold of 1.96 indicates that the articles selected out have only one of the emotion intensity scores which is statistically higher than has an average article at a 95% significance level, whereas other emotion intensity scores are not statistically significantly higher than that of an average article. Our results show that the most expressed emotion in comments is consistent with the emotion most expressed in the associated article (See Figure 8). For example, if anxiety is the most expressed emotion in an article, anxiety would also be the most expressed emotion in the comments. Our results demonstrate that the audience can well receive the emotions in an article. This shows that the audience can receive the same emotion expressed in the article, which is consistent with the prior research (Berger and Milkman 2012, Ferrara and Yang 2015).

[10] Clusterness $C = m/\binom{n}{2}$, where n is the total number of seed users, and m is the number of friend pairs in seed users. The measure resembles the clustering coefficient by Watts and Strogatz (1998), but from a dyad perspective, to fit our context.

[11] According to Granovetter (1977), the stronger the tie between two individuals, the more extensive overlap in their friendship circles. This is due to at least two reasons. First, two individuals with a stronger tie will (by definition) spend more time together. Thus, they have a greater chance to meet and build connections with each other's friends. Second, Granovetter (1977) deduces that strong ties are seldom local bridges that connect different communities. In other words, the stronger the tie that connects two individuals, the higher the probability that they are in the same communities. Thus, they tend to have common friends from these communities.

Second, we ran the hierarchical model with random slopes (Hox et al. 2017), which further considers the publisher-level heterogeneity in the effects of emotions on articles' cascades. In our main model, we fix the coefficients of eight discrete emotions' intensity scores and allow the intercept to vary across publishers in order to capture the consistent effects of the emotions. Now we also allow the coefficients of eight discrete emotions to vary across different publishers, as emotions embedded in articles published by different publishers may have different effects on articles' diffusion. The model is shown in equation 7.

$$\begin{aligned} \text{Level 1 (article): } & y_{ij} = \beta_{0i} + \boldsymbol{\beta}_{1i}^{'}\mathbf{emotion}_{ij} + \boldsymbol{\beta}_{2i}^{'}\mathbf{article}_{ij} + \epsilon_{ij} \\ \text{Level 2 (publisher): } & \beta_{0i} = \gamma_{00} + \boldsymbol{\gamma}_{01}^{'}\mathbf{pubacc}_i + \mu_{0i} \\ & \boldsymbol{\beta}_{1i}^{'} = \boldsymbol{\gamma}_{\mathbf{10}}^{'} + \boldsymbol{\gamma}_{11}^{'}\mathbf{pubacc}_i + \boldsymbol{\mu}_{\mathbf{1i}}^{'} \\ & \boldsymbol{\beta}_{2i}^{'} = \boldsymbol{\beta}_{20}^{'} \end{aligned} \tag{7}$$

where we use $\boldsymbol{\gamma}_{11}$ to capture the effect of different publishers on the effects of eight discrete emotions. Here ϵ_{ij} is assumed to be independent of μ_{0i} and $\boldsymbol{\mu}_{\mathbf{1i}}$ to ensure the independence of different levels, but μ_{0i} and $\boldsymbol{\mu}_{\mathbf{1i}}$ are allowed to be correlated. All of the error terms are assumed to be identically normally distributed with a mean of zero. We selected out the cascades of the articles published by the active publishers who published over 30 articles during our observation period (from August 31 to November 30 in 2018) as the dataset for estimation utilizing the above extended model. We filtered out inactive publishers, because we found that the within-publisher variance of these samples is too small for the model to produce feasible identification. The new dataset includes 118,794 articles. The results are consistent between fix-slope and random-slope models, except that the effect of love on maximum breadth changes (See Table 5), which indicates that there are heterogeneous effects of love on the maximum breadth across different publishers. Specifically, articles with a higher degree of anxiety embedded in the content lead to significantly greater cascades' depth ($\beta_{anxiety}^{depth} = 0.092,\ p < 0.001$), size ($\beta_{anxiety}^{size} = 0.054,\ p < 0.001$), maximum breadth ($\beta_{anxiety}^{breadth} = 0.049,\ p < 0.001$), and structural virality ($\beta_{anxiety}^{sv} = 0.045,\ p < 0.001$) but lower speed ($\beta_{anxiety}^{time} = 0.206,\ p < 0.05$). Articles embedded with a higher degree of love are associated with lower cascades' speed ($\beta_{love}^{time} = 0.128,\ p < 0.05$) and cascades' maximum breadth ($\beta_{love}^{breadth} = -0.055,\ p < 0.001$) but higher structural virality ($\beta_{love}^{sv} = 0.012,\ p < 0.01$). In addition, articles with a higher degree of sadness lead to cascades with significantly lower depth ($\beta_{sadness}^{depth} = -0.061,\ p < 0.001$), size ($\beta_{sadness}^{size} = -0.050,\ p < 0.001$), maximum breadth ($\beta_{sadness}^{breadth} = -0.049,\ p < 0.001$), and structural virality ($\beta_{sadness}^{sv} = -0.026,\ p < 0.001$), but higher speed ($\beta_{sadness}^{time} = -0.207,\ p < 0.001$). Further, articles with a higher degree of anger are associated with higher cascades' speed ($\beta_{anger}^{time} = -0.216,\ p < 0.05$), whereas those with a higher degree of surprise and anticipation are associated with higher cascades' size ($\beta_{surprise}^{size} = 0.021,\ p < 0.001;\ \beta_{anticipation}^{size} = 0.030,\ p < 0.001$) and maximum breadth ($\beta_{surprise}^{breadth} = 0.020,\ p < 0.05;\ \beta_{anticipation}^{breadth} = 0.039,\ p < 0.001$).

Table 5 Multilevel Analysis Results: with Random Slopes[1]

	Depth	**Size[2]**	**Breadth[2]**	**Time**	**SV[5]**
Anger	−0.015	0.001	0.001	−0.216∗	−0.007
	(0.015)[6]	(0.001)	(0.007)	(0.094)	(0.007)
Anxiety	0.092∗∗∗	0.054∗∗∗	0.049∗∗∗	0.206∗	0.045∗∗∗
	(0.026)	(0.014)	(0.014)	(0.151)	(0.011)
Sadness	−0.061∗∗∗	−0.050∗∗∗	−0.049∗∗∗	−0.207∗∗∗	−0.026∗∗∗
	(0.012)	(0.007)	(0.006)	(0.080)	(0.005)
Disgust	0.019	−0.006	−0.006	0.143	0.008
	(0.016)	(0.009)	(0.009)	(0.108)	(0.008)
Joy	0.023	0.015	0.011	0.146	0.007
	(0.025)	(0.013)	(0.013)	(0.136)	(0.011)
Love	0.009	−0.045	−0.055∗∗∗	0.128∗	0.012∗∗
	(0.027)	(0.015)	(0.015)	(0.098)	(0.007)
Surprise	0.004	0.021∗∗∗	0.020∗	−0.123	−0.002
	(0.016)	(0.008)	(0.008)	(0.092)	(0.007)
Anticipation	−0.007	0.030∗∗∗	0.039∗	−0.121	−0.005
	(0.020)	(0.011)	(0.011)	(0.119)	(0.009)
(Intercept)	2.062∗∗∗	2.260∗∗∗	1.759∗∗∗	13.771∗∗∗	1.923∗∗∗
	(0.078)	(0.046)	(0.046)	(0.595)	(0.037)
$\boldsymbol{\sigma_{\mu_{0i}}}$[4]	0.944∗∗∗	0.861∗∗∗	0.917∗∗∗	5.339∗∗∗	0.470∗∗∗
	(0.031)	(0.031)	(0.031)	(0.070)	(0.030)
$\boldsymbol{\sigma_{\mu_{1i}}^{Anger}}$[4]	0.189∗∗∗	0.784∗∗∗	0.077∗∗∗	0.150∗∗∗	0.076∗∗∗
	(0.011)	(0.034)	(0.031)	(0.010)	(0.030)
$\boldsymbol{\sigma_{\mu_{1i}}^{Anxiety}}$[4]	0.354∗∗∗	0.250∗∗∗	0.246∗∗∗	1.040∗∗∗	0.132∗∗∗
	(0.020)	(0.030)	(0.041)	(0.012)	(0.041)
$\boldsymbol{\sigma_{\mu_{1i}}^{Sadness}}$[4]	0.118∗∗∗	0.098∗∗∗	0.094∗∗∗	0.486∗∗∗	0.035∗∗∗
	(0.011)	(0.010)	(0.010)	(0.051)	(0.007)
$\boldsymbol{\sigma_{\mu_{1i}}^{Disgust}}$[4]	0.158∗∗∗	0.117∗∗∗	0.121∗∗∗	0.350∗∗∗	0.117∗∗∗
	(0.022)	(0.031)	(0.031)	(0.040)	(0.034)
$\boldsymbol{\sigma_{\mu_{1i}}^{Joy}}$[4]	0.324∗∗∗	0.209∗∗∗	0.221∗∗∗	0.644∗∗∗	0.116∗∗∗
	(0.021)	(0.031)	(0.031)	(0.042)	(0.030)
$\boldsymbol{\sigma_{\mu_{1i}}^{Love}}$[4]	0.522∗∗∗	0.313∗∗∗	0.318∗∗∗	1.425∗∗∗	0.218∗∗∗
	(0.035)	(0.032)	(0.061)	(0.071)	(0.030)
$\boldsymbol{\sigma_{\mu_{1i}}^{Surprise}}$[4]	0.234∗∗∗	0.123∗∗∗	0.113∗∗∗	0.662∗∗∗	0.102∗∗∗
	(0.021)	(0.031)	(0.044)	(0.060)	(0.031)
$\boldsymbol{\sigma_{\mu_{1i}}^{Anticipation}}$[4]	0.341∗∗∗	0.218∗∗∗	0.215∗∗∗	1.173∗∗∗	0.128∗∗∗
	(0.031)	(0.031)	(0.030)	(0.074)	(0.032)
Controlled[3]					

[1] *Significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] *All of the article-level variables and publisher-level variables are controlled.*

[4] $\boldsymbol{\sigma_{\mu_{0i}}}$ *indicates between-publisher standard deviation of publishers' effect on cascades' dimensions, and* $\boldsymbol{\sigma_{\mu_{1i}}}$ *indicates between-publisher standard deviation of publishers' effect on the relationship between each emotion and cascades' dimensions.*

[5] *SV indicates structural virality.*

[6] *Standard error of estimation.*

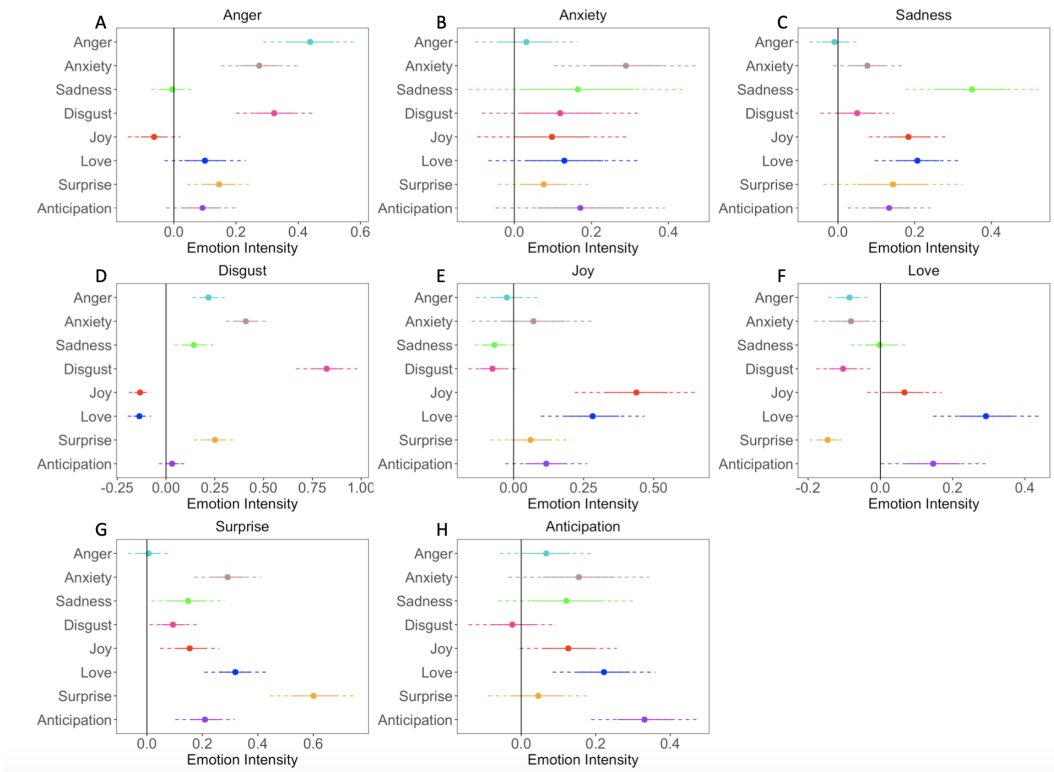

Figure 8 **Discrete emotions in the comments of articles that contain extreme emotions: (A) Anger, (B) Anxiety, (C) Sadness, (D) Disgust, (E) Joy, (F) Love, (G) Surprise, and (H) Anticipation.**

Note: The x axis indicates emotion intensity scores (after normalized to a standard normal distribution). The y axis indicates the category of emotions in comments. The titles of the subplots indicate the type of the extreme emotion that the article contains. The articles that express an extreme emotion are those with only one of the emotion intensity scores higher than 1.96, whereas the others are lower than 1.96. The intervals covered by dashed lines represent 95% confidence intervals. The intervals covered by solid lines represent mean values plus or minus 1 unit of their standard deviations. The figure indicates that the most-expressed emotion embedded in the comments is consistent with the mainly expressed emotion embedded in the articles.

Third, one may concern that there are interactive effects between emotions and topics of articles, which we did not include in the main analysis. We find that including all the interaction terms will result in identification difficulty due to the hierarchical model's high complexity. To address this issue, alternatively, we investigate the relationship between publishers and the number of unique topics involved in their articles (See Figure F.1 in Appendix F). For each article, we identify the topic with the highest probability in its topic distribution vector estimated by the LDA model, and categorize it to that topic. We find that a publisher's articles are consistently on the same topics. The interactive effects between topics and emotions can thus be represented by the interactive effects between publishers and emotions, which we have already considered in the random slope model. As such, the concern of omitting the interactive effects of emotions and topics can be mitigated in this way.

Fourth, our sample includes articles with only a few words and with videos. The emotions in these articles may be expressed mainly in videos instead of texts, although in most of the cases, consistent emotions are expressed between texts and video. Our emotion analysis approach focused exclusively on emotional expressions embedded in text, we thus dropped those articles with videos and that were shorter than 90% of the articles in length, and replicated our analysis with this new sample (See Table G.1 in Appendix G). The results are qualitatively consistent with our main results. Articles with a higher degree of anxiety and love embedded in the content lead to significantly greater cascades' depth ($\beta^{depth}_{anxiety} = 0.037,\ p < 0.001;\ \beta^{depth}_{love} = 0.033,\ p < 0.01$), size ($\beta^{size}_{anxiety} = 0.036,\ p < 0.001;\ \beta^{size}_{love} = 0.021,\ p < 0.001$), maximum breadth ($\beta^{breadth}_{anxiety} = 0.036,\ p < 0.001;\ \beta^{breadth}_{love} = 0.018,\ p < 0.001$), and structural virality ($\beta^{sv}_{anxiety} = 0.017,\ p < 0.001;\ \beta^{sv}_{love} = 0.021,\ p < 0.001$). Articles embedded with a higher degree of love also are associated with lower cascades' speed ($\beta^{time}_{love} = 0.329,\ p < 0.001$). In addition, articles with a higher degree of sadness lead to cascades with significantly lower depth ($\beta^{depth}_{sadness} = -0.023,\ p < 0.001$), size ($\beta^{size}_{sadness} = -0.031,\ p < 0.001$), maximum breadth ($\beta^{breadth}_{sadness} = -0.033,\ p < 0.001$), and structural virality ($\beta^{sv}_{sadness} = -0.012,\ p < 0.001$), but higher speed ($\beta^{time}_{sadness} = -0.172,\ p < 0.001$). Further, articles with a higher degree of anger and joy are associated with lower cascades' maximum breadth ($\beta^{breadth}_{anger} = -0.008,\ p < 0.05;\ \beta^{breadth}_{joy} = -0.012,\ p < 0.05$), whereas those with a higher degree of anticipation are associated with higher cascades' maximum breadth ($\beta^{breadth}_{anticipation} = 0.008,\ p < 0.05$). Articles with a higher degree of surprise exhibit faster cascades ($\beta^{time}_{surprise} = -0.161,\ p < 0.01$).

Finally, information novelty is arguably an influential factor in information diffusion (Berger and Milkman 2012, Vosoughi et al. 2018), and novel information may appear with emotional expressions (e.g., surprise, disgust) (Vosoughi et al. 2018). We add *article originality* (a verification by the WeChat platform that indicated officially whether an article is original) into our main model to control for information novelty. We use a dummy variable for which an article is coded as 1 if it is original and as 0 if it is reproduced, verified by the WeChat platform. The results are shown in Table G.2 in Appendix G, and the effects of anxiety, love, and sadness are qualitatively consistent with the previous results. In particular, articles with a higher degree of anxiety and love embedded in the content lead to significantly greater cascades' depth ($\beta^{depth}_{anxiety} = 0.047,\ p < 0.001;\ \beta^{depth}_{love} = 0.029,\ p < 0.01$), size ($\beta^{size}_{anxiety} = 0.043,\ p < 0.001;\ \beta^{size}_{love} = 0.019,\ p < 0.001$), maximum breadth ($\beta^{breadth}_{anxiety} = 0.042,\ p < 0.001;\ \beta^{breadth}_{love} = 0.017,\ p < 0.001$), and structural virality ($\beta^{sv}_{anxiety} = 0.022,\ p < 0.001;\ \beta^{sv}_{love} = 0.018,\ p < 0.001$). Articles embedded with a higher degree of love, however, also are associated with lower cascades' speed ($\beta^{time}_{love} = 0.351,\ p < 0.001$). In addition, articles with a higher degree of sadness lead to cascades with significantly lower depth ($\beta^{depth}_{sadness} = -0.024,\ p < 0.001$), size ($\beta^{size}_{sadness} = -0.031,\ p < 0.001$), maximum breadth ($\beta^{breadth}_{sadness} = -0.033,\ p < 0.001$), and structural virality ($\beta^{sv}_{sadness} = -0.013,\ p < 0.001$), but higher speed ($\beta^{time}_{sadness} = -0.174,\ p < 0.001$). Further, articles with a higher degree of joy are associated with lower cascades' maximum breadth ($\beta^{breadth}_{joy} = -0.013,\ p < 0.01$), whereas those with a higher degree of anticipation are associated with higher cascades' maximum breadth ($\beta^{breadth}_{anticipation} = 0.012,\ p < 0.01$). Articles with a higher degree of surprise exhibit faster cascades ($\beta^{time}_{surprise} = -0.177,\ p < 0.01$) with higher structural virality ($\beta^{sv}_{surprise} = -0.006,\ p < 0.05$).

8. Discussion

Our work identifies the critical role of emotions in online content diffusion through an analysis of the cascading process of 387,486 articles, randomly sampled from a massive-scale online social network, and an investigation of how over 6 million users transmitted these articles in their local social networks. We find that emotions, especially anxiety, love, and sadness, embedded in content can lead to differential diffusion patterns (i.e., cascade size, depth, maximum breadth, and speed) of otherwise highly similar online articles. Anxiety and love exhibit significant positive impacts in terms of article transmission, whereas sadness generally hinders transmission in online social networks. We further explore the underlying mechanisms and show that individual characteristics (i.e., age and network centrality) and, in particular, the proportion of weak (strong) ties involved in the cascades explain how emotions shape online article diffusion. Our results validate the significance of weak ties in information diffusion and demonstrate that individuals react differently to online content embedded with different emotions. For example, older people are more likely to transmit articles embedded with anger, anxiety, disgust, joy, love, and surprise, but their intentions to share the content with sadness or anticipation are lower than that of younger people. Individuals with larger network centrality tend to spread articles embedded with love and anxiety. Articles with a higher degree of joy, love, and surprise were transmitted more between weak-tie contacts, whereas articles with a higher degree of sadness were transmitted more through strong-tie contacts. Our work further suggests that emotion pairs, which are similar in valence and arousal (e.g., anxiety and anger, love and joy) exhibit disparate effects on article diffusion. This confirms that discrete emotions are unique and relatively independent. The results imply that discrete emotion theory is more useful than the dimensional theory of emotion because our observed pattern cannot be explained by the latter. We also provide empirical evidence of emotion contagion by showing that consistent emotions were detected between the content of articles and the associated readers' comments.

Our work provides valuable insights for practitioners, such as social networking platform managers, content producers, and marketers. First, people can rely on emotions, such as anxiety and love, embedded in content to promote the transmission of online information, such as news articles, product ads, and election campaigns. Second, different strategies can be used to promote the online information with different emotions. For example, a broadcasting strategy (e.g., broadcasting content by mass media or by users with high network centrality) is especially useful in spreading content embedded with anxiety and love. Our results show that anxiety and love are more likely to be broadcast by users with higher network centrality, leading to larger cascades. Love and anxiety can be used to reinforce a viral strategy (e.g., using rewards to encourage peer-to-peer diffusion), as content with these emotions can have cascades with higher expected depth and structural virality. Infusing surprise into the content may hasten its diffusion without significantly affecting the cascade scale. Although sadness also may speed up content diffusion, it significantly shrinks the cascade scale. We find that content with more sadness transmits less through weak ties, implying that sadness is more likely to be diffused within rather than across social communities. Third, to reduce negative emotional content, the platform should pay more attention to anxious information, as it spreads more broadly and more virally. Online communities, especially older Internet users who have more friends than do younger users, are more vulnerable and reactive to anxious content.

Our study is not without limitations. First, existing research in psychology has not reached an agreement on the basic discrete emotion categories for human beings. Although we focused on eight discrete emotions that are well defined in the existing literature and commonly expressed in online content, future research could investigate whether other discrete emotions, for example, contempt, self-hostility, shyness, fear, shame, and guilt, can affect the information-cascade process as well and how they interact with social or psychological processes. Second, future work may investigate the psychological processes involved in each relationship that we find. For example, it would be valuable to answer the question regarding why content with certain emotions tends to spread through various individuals and social ties. Third, future research can also examine how the effects of emotions in content diffusion vary across different topics. Understanding this will better inform content producers to write and promote their work in social media in specific topics. We did not explore the interaction effects of emotions and topics in article diffusion, since, in our context, articles with several topic tags do not have a single clear classification.

References

Aral S, Dhillon P (2016) Unpacking novelty: The anatomy of vision advantages. *Available at SSRN 2388254* .

Aral S, Van Alstyne M (2011) The diversity-bandwidth trade-off. *American Journal of Sociology* 117(1):90–171.

Aral S, Walker D (2012) Identifying influential and susceptible members of social networks. *Science* 337(6092):337–341.

Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO (2013) The diffusion of microfinance. *Science* 341(6144):1236498.

Baron RM, Kenny DA (1986) The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* 51(6):1173.

Berger J, Milkman KL (2012) What makes online content viral? *Journal of marketing research* 49(2):192–205.

Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. *Journal of machine learning research* 13(Feb):281–305.

Birditt KS, Fingerman KL (2003) Age and gender differences in adults' descriptions of emotional reactions to interpersonal problems. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 58(4):P237–P245.

Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. *Journal of machine Learning research* 3(Jan):993–1022.

Bollen J, Mao H, Zeng X (2011) Twitter mood predicts the stock market. *Journal of computational science* 2(1):1–8.

Brady WJ, Wills JA, Jost JT, Tucker JA, Van Bavel JJ (2017) Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* 114(28):7313–7318.

Bruggeman J (2016) The strength of varying tie strength: Comment on aral and van alstyne. *American Journal of Sociology* 121(6):1919–1930.

Centola D (2010) The spread of behavior in an online social network experiment. *Science* 329(5996):1194–1197.

Cheung CM, Lee MK, Lee ZW (2013) Understanding the continuance intention of knowledge sharing in online communities of practice through the post-knowledge-sharing evaluation processes. *Journal of the American Society for Information Science and Technology* 64(7):1357–1374.

Clarke S, Hoggett P, Thompson S (2006) *Emotion, politics and society* (Springer).

Darwin C, Prodger P (1998) *The expression of the emotions in man and animals* (Oxford University Press, USA).

Doh SJ, Hwang JS (2009) How consumers evaluate ewom (electronic word-of-mouth) messages. *CyberPsychology & Behavior* 12(2):193–197.

Dolan RJ (2002) Emotion, cognition, and behavior. *science* 298(5596):1191–1194.

Dorison CA, Wang K, Rees VW, Kawachi I, Ericson KMM, Lerner JS (2020) Sadness, but not all negative emotions, heightens addictive substance use. *Proceedings of the National Academy of Sciences* 117(2):943–949, ISSN 0027-8424, URL `http://dx.doi.org/10.1073/pnas.1909888116`.

Ferrara E, Yang Z (2015) Measuring emotional contagion in social media. *PloS one* 10(11):e0142390.

Figueiredo C, Chen W, Azevedo J (2015) Central nodes and surprise in content selection in social networks. *Computers in Human Behavior* 51:382–392.

Gan Q, Ferns BH, Yu Y, Jin L (2017) A text mining and multidimensional sentiment analysis of online restaurant reviews. *Journal of Quality Assurance in Hospitality & Tourism* 18(4):465–492.

Goel S, Anderson A, Hofman J, Watts DJ (2015) The structural virality of online diffusion. *Management Science* 62(1):180–196.

Goel S, Anderson A, Hofman J, Watts DJ (2016) The structural virality of online diffusion. *Management Science* 62(1):180–196.

Granovetter MS (1977) The strength of weak ties. *Social networks*, 347–367 (Elsevier).

Hausman JA (1978) Specification tests in econometrics. *Econometrica: Journal of the econometric society* 1251–1271.

Heath C, Bell C, Sternberg E (2001) Emotional selection in memes: the case of urban legends. *Journal of personality and social psychology* 81(6):1028.

Henderson PG (2012) *Empirical study of the healing nature of artistic expression: Using mandalas with the positive emotions of love and joy* (Texas A&M University).

Hennig-Thurau T, Wiertz C, Feldhaus F (2015) Does twitter matter? the impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marketing Science* 43(3):375–394.

Hong H, Xu D, Wang GA, Fan W (2017) Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems* 102:1–11.

Hox JJ, Moerbeek M, Van de Schoot R (2017) *Multilevel analysis: Techniques and applications* (Routledge).

Hu Y, Van den Bulte C (2014) Nonmonotonic Status Effects in New Product Adoption. *Marketing Science* 33(4):509–533.

Huang S, Aral S, Hu YJ, Brynjolfsson E (2020) Social Advertising Effectiveness Across Products: A Large-Scale Randomized Field Experiment. *Marketing Science* to appear.

Jackson MO (2008) *Social and economic networks*, volume 3 (Princeton University Press).

Kensinger EA (2008) Age differences in memory for arousing and nonarousing emotional words. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 63(1):P13–P18.

Lazarus RS (1991) Cognition and motivation in emotion. *American psychologist* 46(4):352.

Leatherdale ST, Cameron R, Brown KS, McDonald PW (2005) Senior student smoking at school, student characteristics, and smoking onset among junior students: a multilevel analysis. *Preventive medicine* 40(6):853–859.

Lee M, Youn S (2009) Electronic word of mouth (ewom) how ewom platforms influence consumer product judgement. *International Journal of Advertising* 28(3):473–499.

Lerner JS, Li Y, Valdesolo P, Kassam KS (2015) Emotion and decision making. *Annual review of psychology* 66:799–823.

Lerner JS, Small DA, Loewenstein G (2004) Heart strings and purse strings: Carryover effects of emotions on economic decisions. *Psychological science* 15(5):337–341.

Lewis M, Haviland-Jones JM, Barrett LF (2010) *Handbook of emotions* (Guilford Press).

Liben-Nowell D, Kleinberg J (2008) Tracing information flow on a global scale using internet chain-letter data. *Proceedings of the national academy of sciences* 105(12):4633–4638.

Malik M, Hussain A (2017) Helpfulness of product reviews as a function of discrete positive and negative emotions. *Computers in Human Behavior* 73:290–302.

Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781* .

Mitra S, Ransbotham S (2015) Information disclosure and the diffusion of information security attacks. *Information Systems Research* 26(3):565–584.

Nguyen H, Calantone R, Krishnan R (2020) Influence of social media emotional word of mouth on institutional investors' decisions and firm value. *Management Science* 66(2):887–910.

Peters K, Kashima Y (2007) From social talk to social action: shaping the social triad with emotion sharing. *Journal of personality and social psychology* 93(5):780.

Plutchik R (2001) The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist* 89(4):344–350.

Plutchik R, Kellerman H (1980) *Emotion, theory, research, and experience* (Academic press).

Quan C, Ren F (2010) A blog emotion corpus for emotional expression analysis in chinese. *Computer Speech & Language* 24(4):726–749.

Ransbotham S, Mitra S (2013) The impact of immediate disclosure on attack diffusion and volume. *Economics of Information Security and Privacy III*, 1–12 (Springer).

Ren J, Nickerson JV (2019) Arousal, valence, and volume: how the influence of online review characteristics differs with respect to utilitarian and hedonic products. *European Journal of Information Systems* 28(3):272–290.

Rime B, Mesquita B, Boca S, Philippot P (1991) Beyond the emotional event: Six studies on the social sharing of emotion. *Cognition & Emotion* 5(5-6):435–465.

Rui H, Liu Y, Whinston A (2013) Whose and what chatter matters? the effect of tweets on movie sales. *Decision support systems* 55(4):863–870.

Russell JA (1980) A circumplex model of affect. *Journal of personality and social psychology* 39(6):1161.

Saarimäki H, Gotsopoulos A, Jääskeläinen IP, Lampinen J, Vuilleumier P, Hari R, Sams M, Nummenmaa L (2016) Discrete neural signatures of basic emotions. *Cerebral cortex* 26(6):2563–2573.

Shi Z, Rui H, Whinston AB (2014) Content sharing in a social broadcasting environment: evidence from twitter. *MIS quarterly* 38(1):123–142.

Song T, Huang J, Tan Y, Yu Y (2019) Using user-and marketer-generated content for box office revenue prediction: Differences between microblogging and third-party platforms. *Information Systems Research* 30(1):191–203.

Song Y, Shi S, Li J, Zhang H (2018) Directional skip-gram: Explicitly distinguishing left and right context for word embeddings. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, 175–180.

Stieglitz S, Dang-Xuan L (2013) Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of management information systems* 29(4):217–248.

Tomkins SS (1962) *Affect imagery consciousness: Volume I: The positive affects*, volume 1 (Springer publishing company).

Van den Bulte C, Bayer E, Skiera B, Schmitt P (2018) How customer referral programs turn social capital into economic capital. *Journal of Marketing Research* 55(1):132–146.

Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359(6380):1146–1151.

Vytal K, Hamann S (2010) Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *Journal of cognitive neuroscience* 22(12):2864–2885.

Watson L, Spence MT (2007) Causes and consequences of emotions on consumer behaviour. *European Journal of Marketing* .

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world'networks. *nature* 393(6684):440.

Xiao Y, Zhang H, Cervone D (2018) Social functions of anger: A competitive mediation model of new product reviews. *Journal of Product Innovation Management* 35(3):367–388.

Xue B, Fu C, Shaobin Z (2014) A study on sentiment computing and classification of sina weibo with word2vec. *2014 IEEE International Congress on Big Data*, 358–363 (IEEE).

Yang X, Zhang Z, Zhang Z, Mo Y, Li L, Yu L, Zhu P (2016) Automatic construction and global optimization of a multisentiment lexicon. *Computational intelligence and neuroscience* 2016.

Yin D, Bond S, Zhang H (2014) Anxious or angry? effects of discrete emotions on the perceived helpfulness of online reviews. *Mis Quarterly* 38(2):539–560.

Yin D, Bond S, Zhang H (2017) Keep your cool or let it out: Nonlinear effects of expressed arousal on perceptions of consumer reviews. *Journal of Marketing Research* 54(3):447–463.

Yu Y, Yang Y, Huang J, Tan Y (2019) Emotions in online reviews and product sales: Unifying empirical and theoretical perspectives. *Available at SSRN 3497884* .

Zhang H, Qu C (2018) Emotional, especially negative microblogs are more popular on the web: evidence from an fmri study. *Brain imaging and behavior* 1–11.

Appendix A: Construction of a Domain-Specific Emotion Lexicon

Our approach requires an existing emotion lexicon as a basic lexicon. Ren-CECps is an emotion lexicon based on 1,487 Chinese blog texts (Quan and Ren 2010). Each word in Ren-CECps has eight basic emotion types (i.e., surprise, joy, anticipation, love, anxiety, sadness, anger, and disgust). Each emotion class for each word is manually annotated from 0.0 to 1.0, which represents the emotion intensity expressed by the word. We select Ren-CECps for the following three reasons. First, as Quan and Ren (2010) noted, these eight types of emotions are most commonly expressed in Chinese blog texts; using these emotions decreases confusion in the emotion category selection. Second, the texts are annotated based on Chinese blogs. Blogs and WeChat articles are both typical Chinese online contents. Hence, we believe that Ren-CECps is more suitable for our context than are lexicons based on general Chinese texts (e.g., NTUSD, HowNet, DUT). Finally, Ren-CECps is manually annotated and statistically validated (Quan and Ren 2010)[12]. Compared with automatically constructed emotion lexicons (e.g., Yang et al. 2016), the results of Ren-CECps are more precise and reliable. 12,048 emotion words and their intensities were obtained from Ren-CECps as our basic emotion lexicon.

Next, our approach requires word vectors that contain their semantic information. The word vectors can be derived by statistical language modeling (e.g. Word2Vec by Mikolov et al. (2013)). We used pre-trained word vectors by Song et al. (2018), who provides 200-dimension word vectors for over 8 million common Chinese words and phrases. These word vectors are pretrained on up-to-date, large-scale, and high-quality Chinese online content (Song et al. 2018).

Then, the similarity between the two words can be measured by the cosine distance of the two corresponding word vectors (Mikolov et al. 2013). The similarity distance (SD) between words can be calculated as follows:

$$SD(w_1, w_2) = \frac{\Sigma_{k=1}^{N} x_{1k} x_{2k}}{\sqrt{\Sigma_{k=1}^{N} x_{1k}^2}\sqrt{\Sigma_{k=1}^{N} x_{2k}^2}} \tag{8}$$

where the word vector of w_1 is $(x_{11}, x_{12}, \ldots, x_{1N})$; that of w_2 is $(x_{21}, x_{22}, \ldots, x_{2N})$, assuming that the dimension of the word vector in Word2Vec is N; and $SD(w_1, w_2)$ refers to the similarity distance between w_1 and w_2. Then, a certain word w can be used as a seed word to mine a list of n words that are semantically closest to it by calculating the similarity distance. We named this word list $NearstWordList(w, n)$. For example, if w is a known emotion word, then these similar words mined by Word2Vec can be considered potential emotion words.

We follow three steps by Yu et al. (2019) to extend the basic lexicon to a domain-specific and up-to-date lexicon. First, for every emotion word in the basic lexicon, we use the Word2Vec model to mine its N most similar words as potential emotion words. We selected N as 100, following Xue et al. (2014). Second, we determine the emotion classes contained by each of these potential emotion words. The emotion orientation similarity distance (EO-SD) algorithm is formulated to determine the emotion classes of the potential emotion

[12] Eleven annotators participated in the annotation work. According to Quan and Ren (2010), the authors spent two months on the joint training of annotators and made annotation instructions. The authors also used a Kappa statistic to measure the pairwise agreement among 11 annotators. The Kappa coefficient of the agreement is a statistic adopted by the computational linguistics community as a standard measure for such a purpose. The agreement for emotion words is 0.785. Given the complexity of this annotation task, we believe that the annotations are reliable and valid.

word w, which is adopted and adjusted from the sentiment orientation similarity distance (SO-SD) algorithm by Xue et al. (2014):

$$SDI_k(w) = \sum_{w_i \in NearstWordList(w,n)} SD(w, w_i) I_k(w_i) \tag{9}$$

where w_i belongs to the $NearstWordList(w,n)$; and $I_k(w_i)$ refers to the intensity of the kth emotion ($k \in 1, 2, \ldots, 8$) for w_i, which is from our lexicon. A higher $SDI_k(w)$ indicates a higher probability that w belongs to the kth emotion. To decrease noise, we set a threshold α for $SDI_k(w)$. If $SDI_k(w)$ is less than α, we set it as zero.

Third, we mined the m nearest words[13] of w as $NearestWordList(w,m)$ and used the average emotion intensity level to determine the emotional intensity of w under the kth emotion. The calculation formula is shown below:

$$I_k(w) = Average(I_k(w_i)), \ s.t. \ I_k(w_i) > 0 \quad w_i \in NearestWordList(w,m) \tag{10}$$

Now we have three hyperparameters to be selected (n, m, and α). We determined the parameters ($n = 12, m = 10, \alpha = 1.2$) and tested the validity of our methodology through repeated experiments. More specifically, we randomly selected 2,000 words from the basic lexicon, 1,000 for validation and 1,000 for testing, respectively. The remaining words in the basic lexicon were used as a training set. We followed the above-mentioned three steps to use words in the training set to estimate the emotional intensities of the emotion words in the testing set. The prediction error is defined as the mean absolute error (MAE) between predicted value and true value:

$$MAE = \frac{\Sigma_{i=1}^{s} \Sigma_{k=1}^{8} |I_k(w_i) - I_k^{'}(w_i)|}{8s} \tag{11}$$

where w_i belongs to the validation or the test set, and s is the size of the testing set, which is 1,000 in our case. $I_k(w_i)$ and $I_k^{'}(w_i)$, respectively, represent the true and predicted emotional intensity of the kth emotion of the word w_i. We used a random research approach (Bergstra and Bengio 2012) to determine the hyperparameters. Using the validation set, we find that the best result was achieved by n = 12, m = 10, and α =1.2. Using the test set, we find the corresponding out-of-sample testing error is 0.056. Given that the intensities can vary from 0.0 to 1.0, this result demonstrates the validity of our methodology.

Next, we used an iterative process to repeatedly mine potential emotion words until the number of words in the lexicon converged. To be specific, first, based on the basic lexicon, we followed the above-noted three steps to get potential emotion words and estimated their emotional intensities. Second, we combined the newly mined words with words in the basic lexicon to be an extended lexicon. We then treated the extended lexicon as a new "basic lexicon," which was used to repeat the first and second steps until the total word amount in the combined lexicon converged. During the iteration, the prediction error defined in the previous

[13] We use different parameters n and m to generate two word lists separately, as these two wordlists are used for different tasks. $NearestWordList(w,n)$ is used to determine the emotion classes contained by w, which is the response to questions such as, "Does w belong to the category of joy?" $NearestWordList(w,m)$ is used to determine the emotional intensity of w, which answers questions such as, "Now that we know that w belongs to the category of joy, how much joy has been expressed by the word w (i.e., emotional intensity, measured from 0.0 to 1.0)?" m should generally be smaller than n because intensity needs to be determined by similar words to reduce noise, meaning that we need a smaller set.

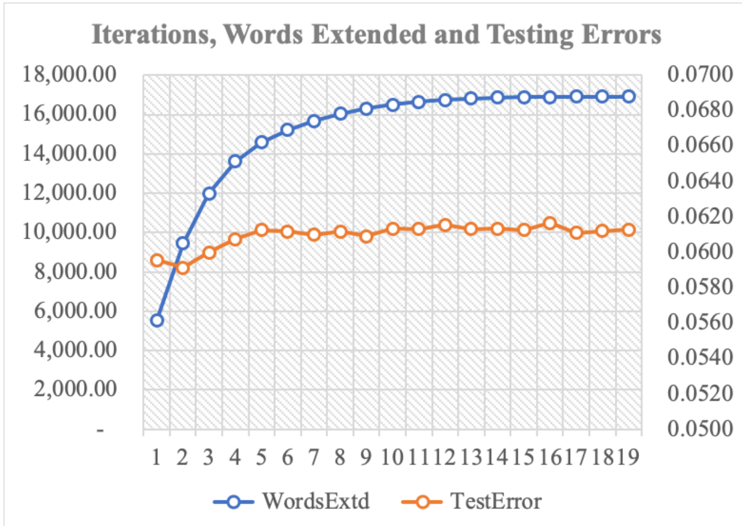

Figure A.1 Iterations, words extended and testing errors.
The figure illustrates the number of extended words and prediction errors for each iteration. When the number of iterations was 20, the number of emotion words converged. During the iteration, the prediction error varied slightly, from 0.0580 to 0.0620. The low and stable prediction error level ensured the validity of the entire iteration process and the accuracy of the lexicon. A total of 16,921 new words were found through this process.

equation was continuously monitored to test the validity of the entire process. Figure A.1 illustrates the number of extended words and prediction error for each iteration. When the number of iterations is 20, the number of emotion words converged. During the iteration, the prediction error varied slightly, from 0.0580 to 0.0620. The low and stable prediction error level ensured the validity of the entire iteration process and the accuracy of the lexicon. A total of 16,921 new words were found after this process. This result confirms the necessity of constructing a domain-specific lexicon, without which 58.4% of unique emotion words (16,921 out of 28,969) would be ignored if only the basic lexicon were used.

Appendix B: Evaluation of the Constructed Lexicon

Our out-of-sample testing, presented in the previous section, guarantees the estimation of the algorithm is consistent with the basic lexicon, annotated by Quan and Ren (2010). Nevertheless, one could be concerned that the basic lexicon might not be valid in our context. A more convincing evaluation would involve a comparison of our algorithm's estimation to human annotations.

To do this, first, 1,200 newly mined emotion words were randomly selected as a testing set. Second, five research assistants were asked to independently annotate the emotional intensities for each of the 1,200 words, and for each of the eight emotion categories. Third, we averaged the results provided by the annotators, so that the results would not be biased due to any one of the annotators' subjectivity. Then, for the eight emotion categories of the 1,200 words, we obtained 9,600 annotations by humans and 9,600 estimations by the algorithm. Finally, we conducted a Student's t-test to compare the two results. We found that the difference between human annotation and algorithm estimation is insignificant ($p = 0.05$, $N = 9{,}600$).

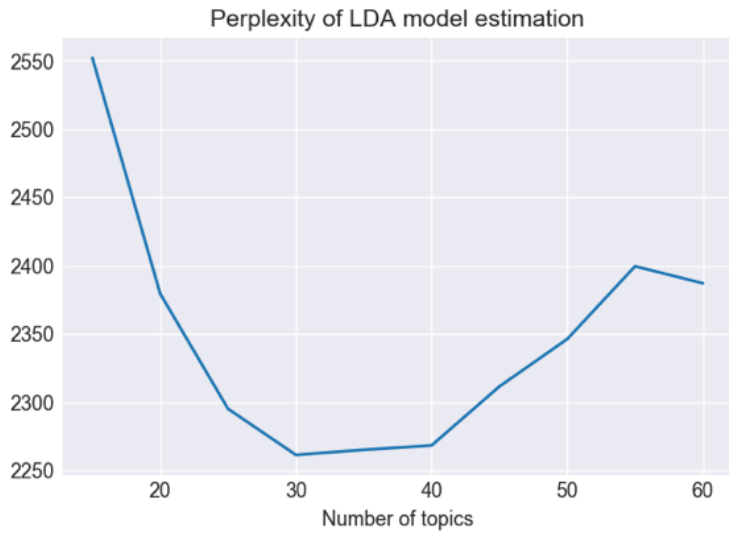

Figure C.1 **Perplexity of Latent Dirichlet Allocation (LDA) model estimation across number of topics.**

Appendix C: Parameter Selection for Topic Modeling

We present the result of parameter selection for topic modeling in Figure C.1. The perplexity is a commonly used metric to select the best parameter (i.e., number of topics) for an LDA model (Blei et al. 2003). The perplexity is defined as exp(-1 $\times$ log-likelihood per word). The optimal number of topics should minimize the perplexity of the model. Thus, we choose number of topics as 30.

Appendix D: Publisher Fixed Effects Model

We run a publisher-level fixed effects panel regression model with clustered standard errors at the publisher level to better control publisher characteristics. The model is shown as follows:

$$y_{ij} = \gamma_i + \beta_1' \mathbf{emotion}_{ij} + \boldsymbol{\beta}_2' \mathbf{article}_{ij} + \epsilon_{ij}, \tag{12}$$

where y_{ij} indicates one of the five cascades' dimensions of the jth article published by the ith publisher, $\mathbf{emotion}_{ij}$ indicates the intensity scores of (discrete) emotions of the jth article published by the ith publisher, and $\mathbf{article}_{ij}$ indicates the article-level control variables of the jth article published by the ith publisher. We use β_1 to capture the effect of emotion, $\boldsymbol{\beta}_2$ to capture the effects of article-level control variables, and γ_i to capture the fixed effect of the ith publisher. γ_i is allowed to be correlated with $\mathbf{emotion}_{ij}$ and $\mathbf{article}_{ij}$. The error term ϵ_{ij} is assumed to be exogenous with $E(\epsilon_{ij}|\gamma_i, \mathbf{emotion}_{ij}, \mathbf{article}_{ij}) = 0$ and with a block-diagonal covariance matrix.

We present the estimation results of the effects of discrete emotions on different cascade dimensions in Table D.1. We note that except that the effects of anger on maximum breadth, love on cascade depth and anticipation on cascade time become less significant, the results are qualitatively consistent with the hierarchical model, where we model publishers' effects as random effects. To further test the consistency of the estimates of the hierarchical models, we run the Durbin-Wu-Hausman test (Hausman 1978). The results are presented in Table D.2, indicating that our random effects models are indeed consistent.

Thus, we conclude that the hierarchical models do not suffer from the endogeneity due to publisher-level omitted variables and the estimates are consistent. Therefore, given that they are both consistent and efficient, the hierarchical models based on random effects are the choice of our analyses.

Table D.1 Fixed Effect Model: Discrete Emotions[1]

	Depth	Size[2]	Breadth[2]	Time	SV[4]
Emotions					
Anger	0.003	−0.002	−0.004	−0.072	0.002
	(0.009)[5]	(0.005)	(0.004)	(0.073)	(0.004)
Anxiety	0.039∗	0.039∗∗∗	0.039∗∗∗	0.101	0.019∗∗
	(0.011)	(0.009)	(0.009)	(0.098)	(0.007)
Sadness	−0.021∗∗	−0.030∗∗∗	−0.032∗∗∗	−0.144∗∗	−0.011∗∗
	(0.007)	(0.004)	(0.004)	(0.055)	(0.004)
Disgust	0.005	0.003	0.003	0.096	0.001
	(0.010)	(0.005)	(0.005)	(0.068)	(0.005)
Joy	0.013	−0.006	−0.011	0.147	0.008
	(0.014)	(0.008)	(0.008)	(0.095)	(0.008)
Love	0.018	0.015∗	0.014∗	0.179∗	0.014∗
	(0.014)	(0.008)	(0.007)	(0.087)	(0.007)
Surprise	−0.009	−0.002	0.000	−0.149∗	−0.005
	(0.010)	(0.005)	(0.005)	(0.060)	(0.005)
Anticipation	0.006	0.008	0.012∗	−0.165∗	−0.002
	(0.012)	(0.007)	(0.007)	(0.082)	(0.006)
Controlled[3]					

[1] *Significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] *All of the article-level variables and publisher fixed effects are controlled for.*

[4] *SV indicates structural virality.*

[5] *Standard error of estimation.*

Table D.2 Durbin–Wu–Hausman Test Results

	Depth	Size	Breadth	Time	SV[2]
Wu–Hausman Statistic	3.175	2.067	1.946	3.081	4.411
P-value	0.923	0.979	0.983	0.929	0.818

[1] *Test of the consistency of the two-level hierarchical regression models' estimates.*
[2] *SV indicates structural virality.*

Appendix E: Tables of Mediation Analysis

Table E.1 Mediation Analysis: Average Age[1]

	Average Age	Depth		Size[2]		Breadth[2]	
	A1	*A2*	*A3*	*A4*	*A5*	*A6*	*A7*
Anger	0.17287***	0.00178	0.00102	−0.00239	−0.00211	−0.00512	−0.00532
	(0.004)[4]	(0.008)	(0.008)	(0.004)	(0.004)	(0.004)	(0.004)
Anxiety	0.21481***	0.04595***	0.04592***	0.04158***	0.04164***	0.04115***	0.04120***
	(0.005)	(0.011)	(0.011)	(0.005)	(0.005)	(0.005)	(0.005)
Sadness	−0.07425**	−0.02430***	−0.02304***	−0.03060***	−0.03015***	−0.03340***	−0.03307***
	(0.003)	(0.006)	(0.007)	(0.003)	(0.003)	(0.003)	(0.003)
Disgust	0.28239***	0.00597	0.00251	0.00328	0.00183	0.00381	0.00249
	(0.004)	(0.007)	(0.007)	(0.004)	(0.004)	(0.004)	(0.003)
Joy	0.18064***	0.01084	0.00629	−0.00876	−0.00990**	−0.01329**	−0.01413***
	(0.005)	(0.01)	(0.01)	(0.005)	(0.005)	(0.005)	(0.005)
Love	0.11291**	0.02813**	0.02855**	0.01857***	0.01883***	0.01707***	0.01727***
	(0.005)	(0.01)	(0.01)	(0.005)	(0.005)	(0.005)	(0.005)
Surprise	0.13950***	−0.01140	−0.01227*	−0.00284	−0.003090	−0.00065	−0.00082
	(0.003)	(0.007)	(0.006)	(0.003)	(0.003)	(0.003)	(0.003)
Anticipation	−0.38030***	0.00039	0.00263	0.00803	0.01101***	0.01190***	0.01236***
	(0.004)	(0.009)	(0.009)	(0.004)	(0.004)	(0.004)	(0.004)
AvgAge			0.01581***		0.00700***		0.00512***
			(0.004)		(0.002)		(0.001)
(Intercept)	38.65426***	1.87923***	1.49988***	1.82061***	1.82337***	1.24905***	1.34019***
	(0.167)	(0.051)	(0.041)	(0.026)	(0.021)	(0.025)	(0.021)
$\boldsymbol{\sigma}_{\mu_i}$[5]		1.088***	1.012***	0.809***	0.811***	0.836***	0.812***
		(0.051)	(0.040)	(0.031)	(0.031)	(0.032)	(0.031)

		Time		Structural Virality	
		A8	*A9*	*A10*	*A11*
Anger		−0.02393	−0.02432	0.00167	0.00237***
		(0.065)	(0.065)	(0.004)	(0.004)
Anxiety		0.16236	0.16322	0.02163***	0.02165***
		(0.094)	(0.094)	(0.005)	(0.005)
Sadness		−0.17513***	−0.17164***	−0.01297***	−0.01234***
		(0.055)	(0.055)	(0.003)	(0.003)
Disgust		0.08678	0.09856	0.00136	0.00498
		(0.064)	(0.064)	(0.004)	(0.004)
Joy		0.03461	0.02856	0.00736	0.00498
		(0.089)	(0.089)	(0.005)	(0.005)
Love		0.34650***	0.34679***	0.01827***	0.01854***
		(0.087)	(0.087)	(0.005)	(0.005)
Surprise		−0.17643***	−0.17841***	−0.00552	−0.00590
		(0.059)	(0.059)	(0.003)	(0.003)
Anticipation		−0.13941	−0.10247	−0.00344	−0.00161
		(0.077)	(0.077)	(0.004)	(0.004)
AvgAge			0.35900***		0.00858***
			(0.042)		(0.002)
(Intercept)		14.44150***	15.66931***	1.70679***	1.51720***
	(0.442)	(0.36)	(0.025)	(0.021)	
$\boldsymbol{\sigma}_{\mu_i}$		9.898***	9.782***	0.660***	0.751
		(1.051)	(1.049)	(0.012)	(0.011)
Controlled[3]					

[1] *Significance level:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] *All of the article-level variables and publisher-level variables are controlled for.*

[4] *Standard error of estimation.*

[5] $\boldsymbol{\sigma}_{\mu_i}$ *indicates between-publisher standard deviation.*

Table E.2 Mediation Analysis: Average Friends Number[1]

	Average Friends Num	Depth		Size[2]		Breadth[2]	
	A1	*A2*	*A3*	*A4*	*A5*	*A6*	*A7*
Anger	−0.01928 ***	0.00178	0.00182	−0.00239	−0.00224	−0.00512	−0.00532
	(0.004)[4]	(0.008)	(0.008)	(0.004)	(0.004)	(0.004)	(0.004)
Anxiety	0.02993 ***	0.04595 ***	0.04594 ***	0.04158 ***	0.04166 ***	0.04115 ***	0.04123 ***
	(0.005)	(0.011)	(0.011)	(0.005)	(0.005)	(0.005)	(0.005)
Sadness	0.00181	−0.02430 ***	−0.02439 ***	−0.03060 ***	−0.03133 ***	−0.03340 ***	−0.03401 ***
	(0.003)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)	(0.003)
Disgust	−0.00046	0.00597	0.00562	0.00328	0.00270	0.00381	0.00304
	(0.004)	(0.007)	(0.007)	(0.004)	(0.004)	(0.003)	(0.003)
Joy	−0.00695	0.01084	0.01054	−0.00876	−0.00881	−0.01329 **	−0.01336 ***
	(0.005)	(0.01)	(0.01)	(0.005)	(0.005)	(0.005)	(0.005)
Love	0.03325 ***	0.02813 **	0.028 **	0.01857 ***	0.01968 ***	0.01707 ***	0.01724 ***
	(0.005)	(0.01)	(0.01)	(0.005)	(0.005)	(0.005)	(0.005)
Surprise	−0.03203 ***	−0.0114	−0.01142*	−0.00284	−0.00301	−0.00065	−0.00087
	(0.003)	(0.007)	(0.007)	(0.003)	(0.003)	(0.003)	(0.003)
Anticipation	−0.01738 ***	0.00039	0.00111	0.00803	0.01044 **	0.0119 **	0.01198 ***
	(0.004)	(0.009)	(0.009)	(0.004)	(0.004)	(0.004)	(0.004)
AveFriendNum			0.01231 **		−0.03931 ***		−0.05456 ***
			(0.005)		(0.002)		(0.002)
(Intercept)	0.17416 ***	1.87923 ***	2.10739 ***	1.82061 ***	2.09239 ***	1.24905 ***	1.52412 ***
	(0.018)	(0.051)	(0.041)	(0.026)	(0.021)	(0.025)	(0.021)
$\boldsymbol{\sigma}_{\mu_i}$[5]		1.088 ***	1.010 ***	0.809 ***	0.821 ***	0.836 ***	0.822 ***
		(0.051)	(0.050)	(0.031)	(0.029)	(0.032)	(0.031)

	Time		Structural Virality	
	A8	*A9*	*A10*	*A11*
Anger	−0.02393	−0.02344	0.00167	0.00163
	(0.065)	(0.065)	(0.004)	(0.004)
Anxiety	0.16236	0.15605*	0.02163 ***	0.02167 ***
	(0.094)	(0.094)	(0.005)	(0.005)
Sadness	−0.17513 ***	−0.17565 ***	−0.01297 ***	−0.01393 ***
	(0.055)	(0.055)	(0.003)	(0.003)
Disgust	0.08678	0.08598	0.00136	−0.00141
	(0.064)	(0.064)	(0.004)	(0.004)
Joy	0.03461	0.03302	0.00736	0.00730
	(0.089)	(0.089)	(0.005)	(0.005)
Love	0.34650 ***	0.35018 ***	0.01827 ***	0.01927 ***
	(0.087)	(0.087)	(0.005)	(0.005)
Surprise	−0.17643 **	−0.17935 ***	−0.00552	−0.00556
	(0.059)	(0.059)	(0.003)	(0.003)
Anticipation	−0.13941	−0.13114	−0.00344	−0.00347
	(0.077)	(0.077)	(0.004)	(0.004)
AveFriendNum		−0.36558 ***		−0.00778 ***
		(0.042)		(0.002)
(Intercept)	14.44150 ***	17.07388 ***	1.70679 ***	1.84627 ***
	(0.442)	(0.36)	(0.025)	(0.021)
$\boldsymbol{\sigma}_{\mu_i}$	9.898 ***	9.882 ***	0.660 ***	0.645
	(1.051)	(1.037)	(0.012)	(0.012)
Controlled[3]				

[1] *Significance level:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] *All of the article-level variables and publisher-level variables are controlled for.*

[4] *Standard error of estimation.*

[5] $\boldsymbol{\sigma}_{\mu_i}$ *indicates between-publisher standard deviation.*

Table E.3 Mediation Analysis: Weak Tie Proportion[1]

	Weak Tie Proportion	Depth		Size[2]		Breadth[2]	
	A1	*A2*	*A3*	*A4*	*A5*	*A6*	*A7*
Anger	0.00149	0.00178	0.01056	−0.00239	−0.00258∗	−0.00512	0.00501
	(0.001)[4]	(0.008)	(0.008)	(0.004)	(0.004)	(0.004)	(0.004)
Anxiety	−0.00141	0.04595∗∗∗	0.04638	0.04158∗∗∗	0.04399∗∗∗	0.04115∗∗∗	0.04506∗∗∗
	(0.001)	(0.011)	(0.041)	(0.005)	(0.005)	(0.005)	(0.005)
Sadness	−0.00271∗∗	−0.02430∗∗∗	0.01579	−0.03060∗∗∗	−0.01870∗∗∗	−0.03340∗∗∗	−0.02630∗∗∗
	(0.001)	(0.006)	(0.064)	(0.003)	(0.003)	(0.003)	(0.003)
Disgust	0.00183	0.00597	0.00948	0.00328	0.00147	0.00381	0.00112
	(0.001)	(0.007)	(0.007)	(0.004)	(0.004)	(0.003)	(0.003)
Joy	0.00530∗∗∗	0.01084	0.02140	−0.00776	−0.01006	−0.01329∗∗	−0.01520∗∗
	(0.001)	(0.01)	(0.01)	(0.005)	(0.005)	(0.005)	(0.005)
Love	0.00600∗∗∗	0.02813∗∗	0.01865	0.01857∗∗∗	0.02029∗∗∗	0.01707∗∗∗	0.02161∗∗∗
	(0.001)	(0.01)	(0.019)	(0.005)	(0.005)	(0.005)	(0.005)
Surprise	0.00420∗∗∗	−0.01140	−0.02352∗∗	−0.00284	−0.01094∗∗∗	−0.00065	−0.00791∗∗
	(0.001)	(0.007)	(0.007)	(0.003)	(0.003)	(0.003)	(0.003)
Anticipation	0.00064	0.00039	0.01970	0.00803	0.02196∗∗∗	0.01190∗∗	0.02499∗∗∗
	(0.001)	(0.009)	(0.009)	(0.004)	(0.004)	(0.004)	(0.004)
WeaktiePro			0.11007∗∗		0.08558∗∗∗		0.07646∗∗∗
			(0.02)		(0.01)		(0.009)
(Intercept)	0.17421∗∗∗	1.87923∗∗∗	3.17957∗∗∗	1.82061∗∗∗	2.73332∗∗∗	1.24905∗∗∗	2.06332∗∗∗
	(0.004)	(0.051)	(0.042)	(0.026)	(0.021)	(0.025)	(0.021)
$\boldsymbol{\sigma}_{\mu_i}$[5]		1.088∗∗∗	1.091∗∗∗	0.809∗∗∗	0.792∗∗∗	0.836∗∗∗	0.852∗∗∗
		(0.051)	(0.051)	(0.031)	(0.040)	(0.032)	(0.028)

	Time		Structural Virality	
	A8	*A9*	*A10*	*A11*
Anger	−0.02393	0.11817	0.00167	0.00628
	(0.065)	(0.065)	(0.004)	(0.004)
Anxiety	0.16236	0.14825	0.02163∗∗∗	0.02903
	(0.094)	(0.094)	(0.005)	(0.051)
Sadness	−0.17513∗∗∗	−0.04405∗∗∗	−0.01297∗∗∗	0.00840
	(0.055)	(0.055)	(0.003)	(0.04)
Disgust	0.08678	0.09331	0.00136	0.00225
	(0.064)	(0.064)	(0.004)	(0.004)
Joy	0.03361	−0.11777	0.00736	0.01411∗
	(0.089)	(0.089)	(0.005)	(0.004)
Love	0.34650∗∗∗	0.53856∗∗∗	0.01827∗∗∗	0.01754∗∗
	(0.087)	(0.087)	(0.005)	(0.007)
Surprise	−0.17643∗∗	−0.19061∗∗∗	−0.00552	−0.01227∗∗∗
	(0.059)	(0.059)	(0.003)	(0.003)
Anticipation	−0.13941	−0.20062∗∗	−0.00344	0.00351
	(0.077)	(0.077)	(0.004)	(0.007)
WeaktiePro		1.56484∗∗∗		0.09059∗∗∗
		(0.174)		(0.01)
(Intercept)	14.44150∗∗∗	16.57288∗∗∗	1.70679∗∗∗	2.43135∗∗∗
	(0.442)	(0.361)	(0.025)	(0.021)
$\boldsymbol{\sigma}_{\mu_i}$	9.898∗∗∗	9.934∗∗∗	0.660∗∗∗	0.637
	(1.051)	(1.030)	(0.012)	(0.015)
Controlled[3]				

[1] *Significance level:* $^{*}\,p<0.05$, $^{**}\,p<0.01$, $^{***}\,p<0.001$

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] *All of the article-level variables and publisher-level variables are controlled for.*

[4] *Standard error of estimation.*

[5] $\boldsymbol{\sigma}_{\mu_i}$ *indicates between-publisher standard deviation.*

Appendix F: Figures of Robustness Checks

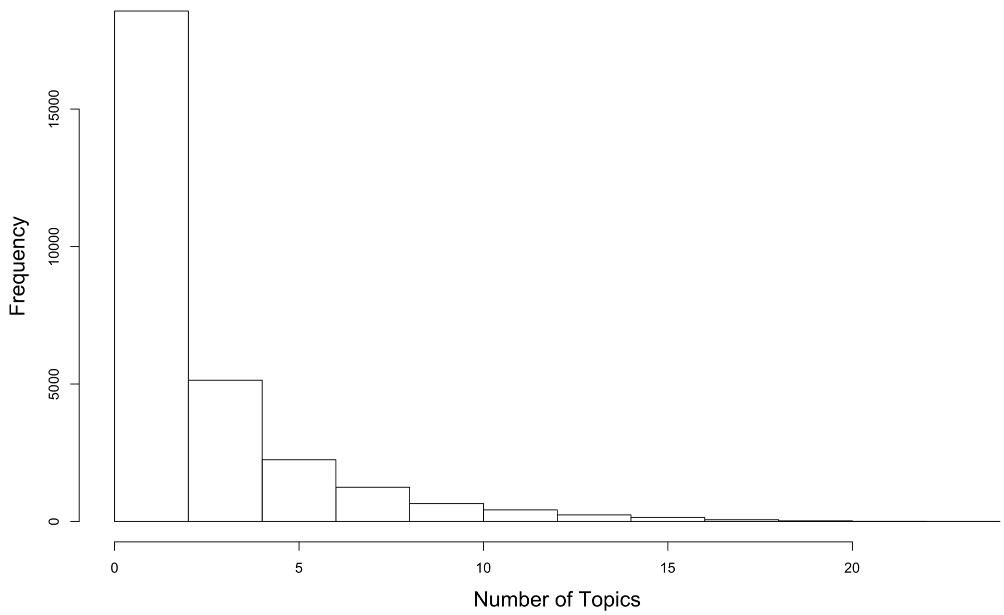

Figure F.1 Distribution of the number of different topics a publisher wrote about.

Note: The figure shows the distribution of the number of unique topics that were involved in a publisher's articles. The results indicate that publishers mainly publish articles with consistent topics.

Appendix G: Tables of Robustness Checks

Table G.1 Multilevel Analysis Results: Excluding Articles with Videos and Few Words[1]

	Depth	**Size[2]**	**Breadth[2]**	**Time**	**SV[5]**
Anger	−0.004	−0.005	−0.008*	−0.023	−0.001
	(0.008)[6]	(0.004)	(0.004)	(0.065)	(0.004)
Anxiety	0.037***	0.036***	0.036***	0.144	0.017**
	(0.011)	(0.005)	(0.005)	(0.094)	(0.005)
Sadness	−0.023***	−0.031***	−0.033***	−0.172**	−0.012***
	(0.006)	(0.003)	(0.003)	(0.056)	(0.003)
Disgust	0.002	0.001	0.002	0.084	−0.001
	(0.008)	(0.004)	(0.003)	(0.064)	(0.004)
Joy	0.013	−0.007	−0.012*	0.031	0.009
	(0.010)	(0.005)	(0.005)	(0.089)	(0.005)
Love	0.033**	0.021***	0.018***	0.329***	0.021***
	(0.010)	(0.005)	(0.005)	(0.087)	(0.005)
Surprise	−0.004	0.002	0.004	−0.161**	−0.002
	(0.007)	(0.003)	(0.003)	(0.059)	(0.003)
Anticipation	−0.009	0.004	0.008*	−0.139	−0.008
	(0.009)	(0.004)	(0.004)	(0.077)	(0.004)
(Intercept)	1.914***	1.818***	1.241***	14.208***	1.727***
	(0.052)	(0.026)	(0.025)	(0.449)	(0.026)
$\boldsymbol{\sigma_{\mu_i}}$[4]	1.081***	0.816***	0.834***	9.983***	0.647***
	(0.030)	(0.010)	(0.010)	(0.072)	(0.040)
Controlled[3]					

[1] *Significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

[2] *Given the right-skewed distributions of size and maximum Breadth, we did log-transformation on these two variables.*

[3] *All of the article-level variables and publisher-level variables are controlled for.*

[4] $\boldsymbol{\sigma_{\mu_{0i}}}$ *indicates between-publisher standard deviation.*

[5] *SV indicates structural virality.*

[6] *Standard error of estimation.*

Table G.2 Multilevel Analysis Results: Controlling for Article Originality[1]

	Depth	Size[2]	Breadth[2]	Time	SV[5]
Emotions					
Anger	0.003	−0.002	−0.004	−0.024	0.002
	(0.008)[6]	(0.004)	(0.004)	(0.065)	(0.004)
Anxiety	0.047∗∗∗	0.043∗∗∗	0.042∗∗∗	0.162	0.022∗∗∗
	(0.011)	(0.005)	(0.005)	(0.094)	(0.005)
Sadness	−0.024∗∗∗	−0.031∗∗∗	−0.033∗∗∗	−0.174∗∗	−0.013∗∗∗
	(0.006)	(0.003)	(0.003)	(0.055)	(0.003)
Disgust	0.006	0.003	0.004	0.087	0.001
	(0.007)	(0.004)	(0.003)	(0.065)	(0.004)
Joy	0.010	−0.008	−0.013∗∗	0.034	0.007
	(0.010)	(0.005)	(0.005)	(0.089)	(0.005)
Love	0.029∗∗	0.019∗∗∗	0.017∗∗∗	0.351∗∗∗	0.018∗∗∗
	(0.010)	(0.005)	(0.005)	(0.089)	(0.005)
Surprise	−0.011	−0.003	−0.001	−0.177∗∗	−0.006∗
	(0.007)	(0.003)	(0.003)	(0.059)	(0.003)
Anticipation	0.001	0.008	0.012∗∗	−0.141∗	−0.003
	(0.009)	(0.004)	(0.004)	(0.078)	(0.004)
Originality	0.399∗∗∗	0.260∗∗∗	0.230∗∗∗	−0.241	0.219∗∗∗
	(0.023)	(0.012)	(0.012)	(0.203)	(0.012)
(Intercept)	1.856∗∗∗	1.801∗∗∗	1.232∗∗∗	14.712∗∗∗	1.693∗∗∗
	(0.050)	(0.025)	(0.025)	(0.434)	(0.025)
$\boldsymbol{\sigma_{\mu_i}}$[4]	1.087∗∗∗	0.807∗∗∗	0.834∗∗∗	9.898∗∗∗	0.659∗∗∗
	(0.030)	(0.012)	(0.012)	(0.060)	(0.030)
Controlled[3]					

[1] *Significance level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

[2] *Given the right-skewed distributions of size and maximum breadth, we did a log-transformation on these two variables.*

[3] *All of the article-level variables and publisher-level variables are controlled for.*

[4] $\boldsymbol{\sigma_{\mu_{0i}}}$ *indicates between-publisher standard deviation.*

[5] *SV indicates structural virality.*

[6] *Standard error of estimation.*