



Emotions in online rumor diffusion

Nicolas Pröllochs^{1*} , Dominik Bär² and Stefan Feuerriegel^{2,3}

*Correspondence:

nicolas.proellochs@wi.jlug.de

¹University of Giessen, Licher Str. 62,
35394 Giessen, Germany

Full list of author information is
available at the end of the article

Abstract

Emotions are regarded as a dominant driver of human behavior, and yet their role in online rumor diffusion is largely unexplored. In this study, we empirically study the extent to which emotions explain the diffusion of online rumors. We analyze a large-scale sample of 107,014 online rumors from Twitter, as well as their cascades. For each rumor, the embedded emotions were measured based on eight so-called basic emotions from Plutchik's wheel of emotions (i.e., anticipation–surprise, anger–fear, trust–disgust, joy–sadness). We then estimated using a generalized linear regression model how emotions are associated with the spread of online rumors in terms of (1) cascade size, (2) cascade lifetime, and (3) structural virality. Our results suggest that rumors conveying anticipation, anger, and trust generate more reshares, spread over longer time horizons, and become more viral. In contrast, a smaller size, lifetime, and virality is found for surprise, fear, and disgust. We further study how the presence of 24 dyadic emotional interactions (i.e., feelings composed of two emotions) is associated with diffusion dynamics. Here, we find that rumors cascades with high degrees of aggressiveness are larger in size, longer-lived, and more viral. Altogether, emotions embedded in online rumors are important determinants of the spreading dynamics.

Keywords: Online rumors; Information cascade; Online diffusion; Emotions;
Regression analysis

1 Introduction

Social media platforms such as Facebook, Sina Weibo, and Twitter allow users to disseminate content through sharing (e.g., called retweeting in the case of Twitter). As a result, content can go viral and reach a large audience despite the fact that it originated from a single broadcast. To this end, understanding the diffusion of online content is relevant for a number of reasons. Marketers are interested in identifying what makes content go viral, so that marketing content can be designed accordingly [1–4]. Humanitarian organizations leverage the potential of online diffusion in social media to collect information for effective responses to natural disasters and to inform the wider public [5–7]. Public stakeholders are confronted with the diffusion of political content and, by understanding the underlying mechanics, can help prevent the spread of rumors [8–11].

Previous research has identified several drivers of online diffusion (see Additional file 1 for an overview). These drivers are primarily located in the different characteristics of senders. For instance, senders with a larger follower base (i.e., with more outgoing ties in the network) also reach, on average, a larger audience [12]. Other characteristics of

© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

senders are the number of followees (i.e., how many incoming ties a user has [13–15]) or their past engagement (i.e., the number of posts or reshares [11]). A different stream of research has examined online diffusion around specific topics (e.g., a specific election [9] or a specific disaster [5–7, 16–19]). In this work, we add by studying the role of emotions in the diffusion of online rumors.

Emotions have been established as an important determinant of human behavior in offline behavior [20–22]. Emotions typically arise as a response to environmental stimuli that are of relevance to the needs, goals, or concerns of users and, as a consequence, also guide user behavior in online settings [23]. Emotions influence what type of information users seek, what they process, how they remember it, and ultimately what judgments and decisions they derive from it. Emotions are themselves contagious and can spread among people, both offline (i.e., in person) [24] and online (i.e., via social media) [25–29].

Following the above, an important driver of online behavior are emotions embedded in online content. For instance, it was previously confirmed that emotions influence posting and liking activities [30], users' willingness-to-share [1], and actual sharing behavior [2, 31–33]. As such, embedded emotions explain, to a large extent, the propensity to share posts, as well as user response time. Here, emotional stimuli such as emotion-laden wording trigger cognitive processing [34], which in turn results in the behavioral response of information sharing [35–37]. In particular, emotions embedded in online content also explain the dynamics of online diffusion. For instance, emotions describe different properties of diffusion cascades, such as their size, branching, or lifetime [38–41]. Especially misinformation relies upon emotions in order to attract attention [11, 38, 42–46]. Given the importance of emotions in online behavior, we investigate how emotions are linked to the spread of online rumors.

Hypothesis *Emotions embedded in online rumors are associated with the size, lifetime, and structural virality of the cascade.*

In this study, we empirically analyze to what extent emotions explain the diffusion of online rumors. For this, we infer the emotions embedded in replies to online rumors through the use of affective computing (see Methods). For each rumor, the degree of emotion is rated along so-called basic emotions. Basic emotions refer to a subset of emotions that are universally recognized across cultures and through which other, more complex emotions can be derived. In this work, we adopt Plutchik's wheel of emotions [22], comprising 8 basic emotions (ANTICIPATION, SURPRISE, ANGER, FEAR, TRUST, DISGUST, JOY, SADNESS). Based on these, we infer 24 dyadic emotional interactions, each representing a more complex emotion composed of two basic emotions (e.g., AGGRESSION as a combination of ANGER and ANTICIPATION). These emotions are then linked to the spread of online rumors using regression analysis. Thereby, we estimate to what extent emotions embedded in online rumors explain: (1) cascade size, that is, how many reshares a rumor generates; (2) cascade lifetime, that is, how long a rumor is active; and (3) structural virality, that is, how effectively it spreads. The latter, structural virality, provides a quantitative metric [47] aggregating the depth-breadth variation in rumor diffusion.

One work [11] contains summary statistics reporting which emotions are present in online rumors but not how emotions affect *sharing*. Hence, any statistical claims measuring the emotion effect (= which emotions drive a faster and wider rumor spreading) are

precluded. This presents the added value of our work. We measure how emotions are associated with the diffusion dynamics (e.g., TRUST as an emotion is present in only a small portion of rumors but it has a large influence on virality). Because of this, our work is different in several ways: (i) we focus not only on basic emotions but also dyadic emotions, (ii) we infer the emotion effect on diffusion dynamics, and, because of that, (iii) we use a regression analysis as opposed to summary statistics. Therefore, this work is—to the best of our knowledge—the first comprehensive study assessing the link between emotions and the spread of online rumors.

We analyze a large-scale, representative sample of Twitter rumors and their corresponding cascades [11]. Specifically, our data cover the complete time frame from the launch of Twitter in 2006 until (and including) 2017. Altogether, this results in 2189 rumors associated with 107,014 cascades. The sample comprises approx. 3.7 million reshares that originate from almost 3 million different users. Based on the cascades, various control variables are constructed. Specifically, in our regression analysis, we capture time- and rumor-effects through the use of random effects, based on which we control for the heterogeneity among rumors (see Materials and Methods).

2 Materials and methods

2.1 Dataset

A rumor is defined as a piece of content that is propagated between users but without confirmation of its veracity. This definition is rooted in social psychology literature [43, 48]. For this study, a large-scale dataset comprising of rumor cascades from Twitter [11] was analyzed. The resulting sample comprises *all* rumors from Twitter between its founding in the year 2006 until (and including) 2017. Ethics approval was obtained from ETH Zurich (2020-N-44). Overall, our sample includes 2189 rumors with a total of $N = 107,014$ cascades (i.e., some rumor contents were shared as part of multiple but different cascades). The rumors had approx. 3.7 million reshares originating from 3 million users (see [11] for details).

2.2 Characteristics of online rumor diffusion

The cascades were then processed as follows in order to generate additional variables. These variables refer to different characteristics of online rumor diffusion and later represent the dependent variables in the regression analysis. For simplicity, we introduce the following notation. We refer to the cascades via $j = 1, \dots, N$. These belong to $i = 1, \dots, 2189$ different rumors. Each cascade is a three-tuple $T_j = (r_j, t_{j0}, R_j)$, where r_j is the root post that corresponds to the original broadcast and where t_{j0} is its timestamp and R_j the set of reshairs. A reshare k has a parent p_{jk} and a timestamp t_{jk} , i.e., $R_j = \{(p_{jk}, t_{jk})\}_k$.

- (1) *Cascade size*: The cascade size counts how many reshares a cascade generated.

Formally, it amounts to all reshares plus 1 (for the root), i.e., $|R_j| + 1$.

- (2) *Cascade lifetime*: The cascade lifetime is the timespan during which a rumor cascade was active, thus the elapsed time between the root broadcast and the last reshare. It is calculated via $\max_k t_{jk} - t_{j0}$.
- (3) *Structural virality*: Structural virality [47] provides an aggregated metric combining the depth and breadth of a cascade. A higher structural virality corresponds to a cascade that is both of great depth *and* where each reshare generated a large relative number of additional reshares (i.e., a high branching factor). As proposed in [47],

structural virality is based on the idea of the Wiener index, i.e.,

$$\nu(T_j) = \frac{1}{|R_j| \times (|R_j| + 1)} \sum_{j_1=0}^{|R_j|} \sum_{j_2=0}^{|R_j|} d_{j_1 j_2}, \quad (1)$$

where $d_{j_1 j_2}$ is the shortest path between nodes j_1 and j_2 in the tree T_j . Intuitively, structural virality reflects the average distance between all reshares in the graph.

2.3 Model variables on heterogeneity between rumor cascades

Model variables x_j , concerning the heterogeneity among rumor cascades, were computed as in earlier research [11, 12, 31, 38]. These later act as controls. In our study, controls are (1) account age; (2) a binary dummy representing whether the account is officially labeled as “verified” (= 1 if yes, i.e., Twitter displays a blue badge next to it); (3) the number of followers (outgoing ties); (4) the number of followees (incoming ties); and (5) user engagement, that is, the average number of posts, reshares, and likes relative to the account age as in [11]. These variables reflect that the senders of rumors vary in their social influence.

Note that all of the above variables were computed at the level of cascades (which is later our unit of analysis). Additional sources of heterogeneity among rumors are captured via rumor-level random effects.

2.4 Computing emotions embedded in online rumors

For all cascades, we measured the emotions embedded in replies to rumor cascades. Here, we distinguish *basic emotions*, *bipolar emotion pairs*, and *dyadic emotional interactions* comprising primary, secondary, tertiary dyads. The computation of the emotions is detailed below (see [22] for further details).

Basic emotions: Basic emotions refer to a subset of emotions that are universally recognized across cultures and through which other, more complex emotions can be derived [20, 21]. In our study, Plutchik’s wheel of emotions [22] is adopted as it is a common tool in affective computing [49]. It defines 8 basic emotions (see Fig. 1, petals): ANTICIPATION, SURPRISE, ANGER, FEAR, TRUST, DISGUST, JOY, and SADNESS.

Our computation follows a dictionary-based approach as in [11]. Dictionary-based approaches are widely used when large-scale analyses of emotions are performed with the objective of explanatory modeling and thus reliable interpretations [38, 41]. In our work, the NRC emotion lexicon was used [50], which classifies English words into the 8 basic emotions. For all cascades j , the content of the replies was tokenized and the frequency of dictionary terms per basic emotion was counted, resulting in an 8-dimensional emotion score e_j . Afterwards, the vector was normalized to sum to one across basic emotions (i.e., $e'_j = \frac{1}{\|e_j\|_1} e_j$). We omit rumor cascades that do not contain any emotional words from the NRC emotion lexicon (since, otherwise, the denominator is not defined). As a result, the 8 emotion dimensions in $e'_j \in [0, 1]^8$ range from zero to one. Owing to this fact, replies to rumors can embed a combination of multiple emotions (e.g., 40% ANGER and 60% FEAR).

Bipolar emotion pairs: In Plutchik’s wheel of emotions, the 8 basic emotions are organized according to 4 pairs of bipolar emotions (i.e., the opposite petals in Fig. 1). The 4 pairs of bipolar emotions are ANTICIPATION–SURPRISE, ANGER–FEAR, TRUST–DISGUST, JOY–SADNESS. In each case, one dimension of the pair is considered to be positive and the other

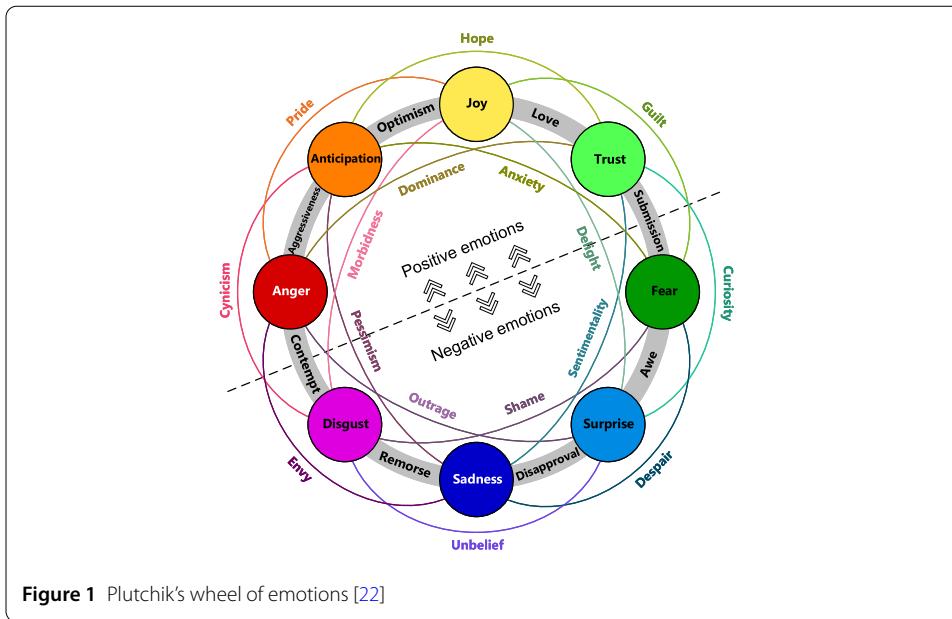


Figure 1 Plutchik's wheel of emotions [22]

negative. We calculate a 4-dimensional score ϕ_j^{pairs} that measures the difference between a specific positive emotion and its complement from the set of negative emotions. For example, ANGER–FEAR refers to the difference between ANGER and FEAR.

Dyadic emotional interactions: Plutchik's wheel of emotions further defines 24 dyadic emotional interactions, which are more complex emotions composed of two basic emotions (see Fig. 1, round lines). The dyadic emotional interactions comprise:

- 1 *Primary dyads* that are one petal apart from each other (e.g., AGGRESSION = ANGER + ANTICIPATION). The 8 primary dyadic emotional interactions are OPTIMISM, DISAPPROVAL, LOVE, REMORSE, SUBMISSION, CONTEMPT, AWE, and AGGRESSION.
- 2 *Secondary dyads* that are two petals apart from each other (e.g., HOPE = ANTICIPATION + TRUST). The 8 secondary dyadic emotional interactions are HOPE, UNBELIEF, GUILT, ENVY, CURIOSITY, CYNICISM, DESPAIR, and PRIDE.
- 3 *Tertiary dyads* that are three petals apart from each other (e.g., ANXIETY = ANTICIPATION + FEAR). The 8 tertiary dyadic emotional interactions are ANXIETY, OUTRAGE, DELIGHT, PESSIMISM, SENTIMENTALITY, MORBIDNESS, SHAME, and DOMINANCE.

Similar to the bipolar emotion pairs, the dyadic emotional interactions are arranged such that each has an opposite emotion. For example, LOVE is the opposite of REMORSE. Hence, for each pair, we again compute a score that is the difference between the opposing emotions. This yields $\phi_j^{primary}, \phi_j^{secondary}, \phi_j^{tertiary} \in [0, 1]^4$.

2.5 Regression analysis

To analyze the role of emotions in online rumor diffusion, we apply a generalized regression model. Regression models are generally regarded as an explanatory approach with the ability to document statistical relationships and, in particular, estimate effect sizes [51]. Furthermore, regression models are widely used to estimate the marginal effect of content on diffusion characteristics [11, 31, 38, 41]. This allows us to later make inferences that test our research hypothesis statistically.

Let y_j denote a characteristic of the cascade of interest, namely cascade size, cascade lifetime, or structural virality. We then model y_j of the cascade via a two-level generalized hierarchical regression:

$$\text{Level 1: } y_j = \alpha_i + \beta^T \phi_j + \gamma^T x_j + \varepsilon_j, \quad (2)$$

$$\text{Level 2: } \alpha_i = \gamma_0 + \gamma_i, \quad (3)$$

where level 1 refers to the cascade level and level 2 to the rumor level. The other variables are as follows. The coefficient β captures the marginal effect of emotions. This is later our variable of interest as it measures the contribution of emotions to rumor diffusion. The coefficient γ is used to control for other model variables at the rumor cascade level. Both γ_0 and γ_i are assumed to be independent and identically normally distributed with mean zero. Then γ_0 reflects the base diffusion in the sample, while γ_i controls for variation at rumor level. Notably, this turns α_i into a rumor-specific random effect. The error term ε_j is assumed to be independent and identically normally distributed with mean zero.

The use of regression analysis is imperative for the scope of our study. The reasons are as follows. (1) Our objective is different from predictive modeling [51], where the focus is on accurate estimates of the outcome variable. Instead, we are concerned with the model logic as it allows us to interpret the model coefficients. (2) Our objective is also different from analyzing summary statistics as in [11]. Summary statistics deal with comparisons across groups and thereby ignore other sources of heterogeneity in the sample. For instance, the summary statistics on rumor emotions in [11] only report which emotions are common but not how emotions are associated with sharing dynamics. This is especially relevant for our research as we expect that some properties of rumor diffusion are also due to the social influence of the sender. Hence, by combining emotions and further controls in a joint regression model, we can isolate the *marginal* effect of emotions on the diffusion dynamics, which would not be possible with summary statistics.

Later, a regression analysis based on basic emotions is precluded due to multicollinearity (recall that the emotion scores e_j sum to one across basic emotions). Instead, the regression analysis is performed using bipolar emotion pairs ϕ_j^{pairs} and the dyadic emotional interactions $\phi_j^{primary}$, $\phi_j^{secondary}$, $\phi_j^{tertiary}$. For the latter, we fit 12 separate models, i.e., one for each pair among the emotional dyads, due to linear dependencies between the dyads.

In our implementation, the estimator depends on the distribution of y_j as follows:

1 Cascade size is modeled via a negative binomial regression with log-transformation.

The reason is that cascade size denotes count data with overdispersion (i.e., variance larger than the mean).

2 Cascade lifetime is first log-transformed and then modeled via a normal distribution.

This is consistent with previous research assuming a log-normal distribution for response times [12].

3 Structural virality is modeled via a gamma regression with a log-link. This allows us to account for a skewed distribution of continuous, non-negative variables.

All estimations are conducted based on the R package `lme4`. Before estimation, all model variables are z -standardized. Owing to this, the regression coefficients quantify changes in the dependent variable in standard deviations. This is beneficial as it allows us to compare the estimated coefficients across emotions in a straightforward manner.

3 Results

3.1 Summary statistics

The diffusion dynamics in our data are as follows. Figure 2 compares cascade size, lifetime, and structural virality via complementary cumulative distribution functions (CCDF). On average, a rumor cascade reaches 31.95 users and has a lifetime of 123.18 hours. The mean structural virality is 1.26.

Basic emotions: Fig. 3 plots the CCDFs for each of the eight basic emotions, while Fig. 4 reports the relative proportion of emotional intensity averaged over all rumors. We find that a large proportion of rumors embed DISGUST and SURPRISE, whereas comparatively few rumors embed JOY and SADNESS. Evidently, rumors embed more ANGER (relative share of 12.34%) than FEAR (10.74%), more SURPRISE (16.44%) than ANTICIPATION (14.23%), more DISGUST (23.58%) than TRUST (9.05%), and more JOY (7.39%) than SADNESS (6.23%). Overall,

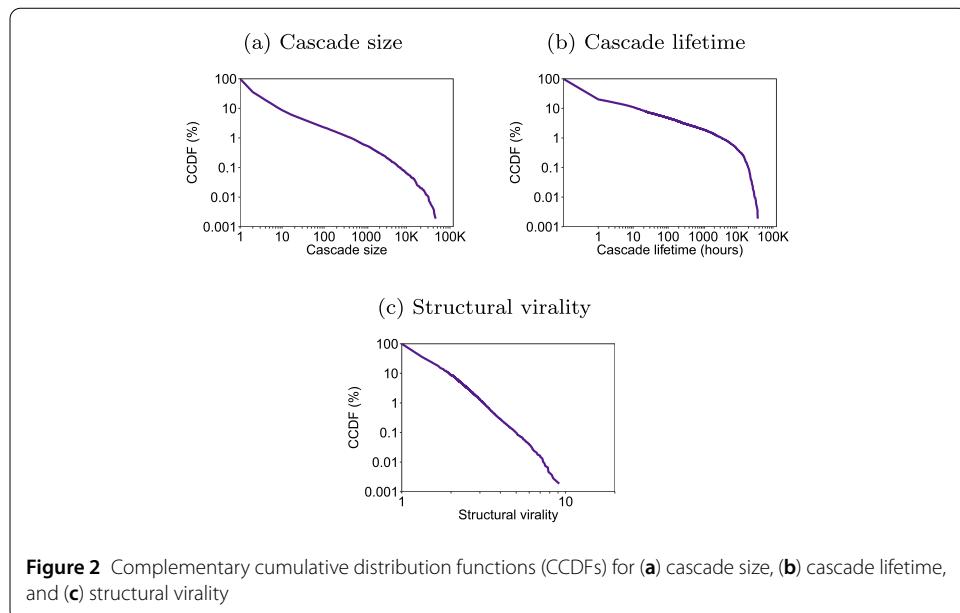


Figure 2 Complementary cumulative distribution functions (CCDFs) for (a) cascade size, (b) cascade lifetime, and (c) structural virality

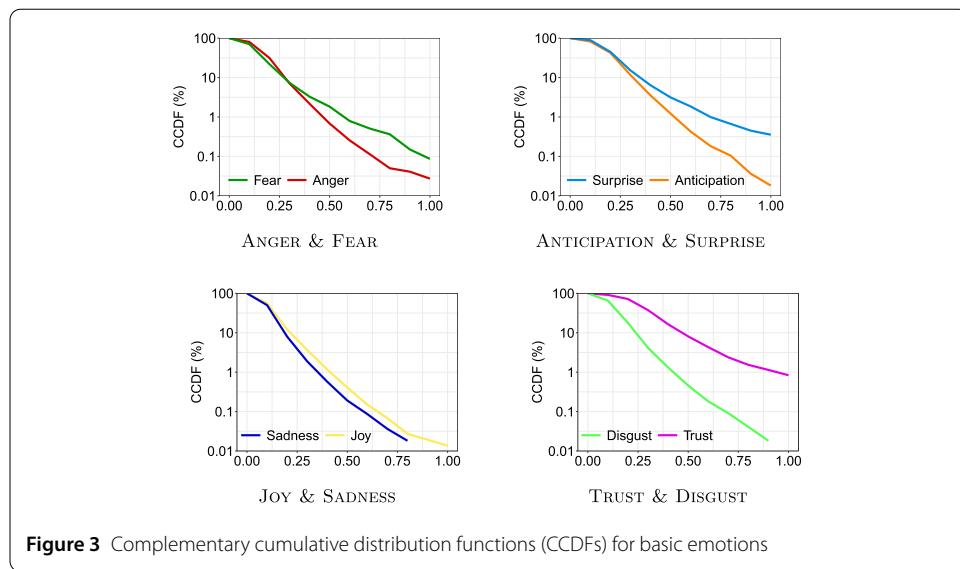
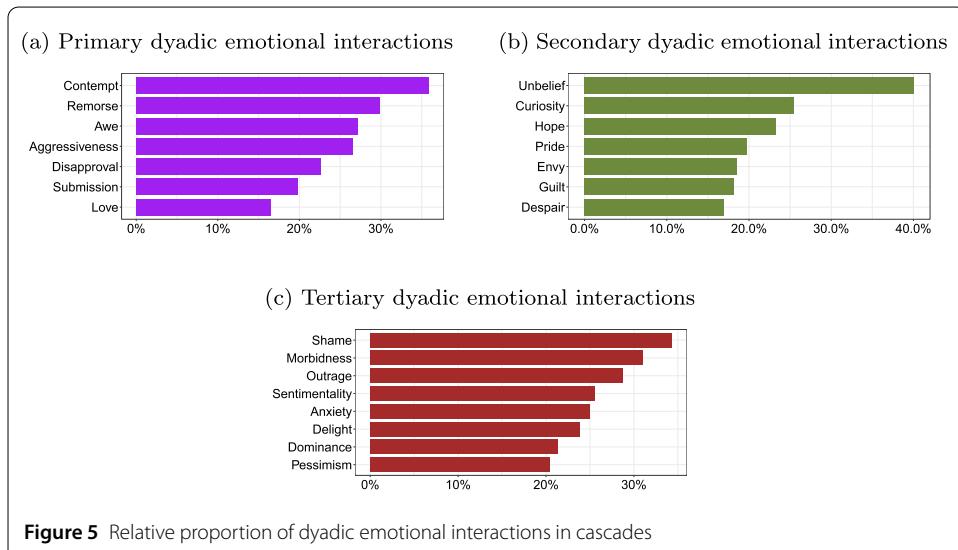
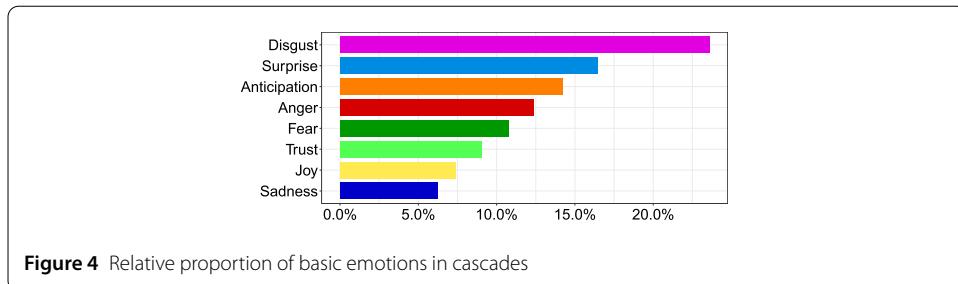


Figure 3 Complementary cumulative distribution functions (CCDFs) for basic emotions



43.01% of the embedded emotions originate from the group of positive emotions, while 56.98% belong to the group of negative emotions. Hence, rumors comprise more negative than positive emotions.

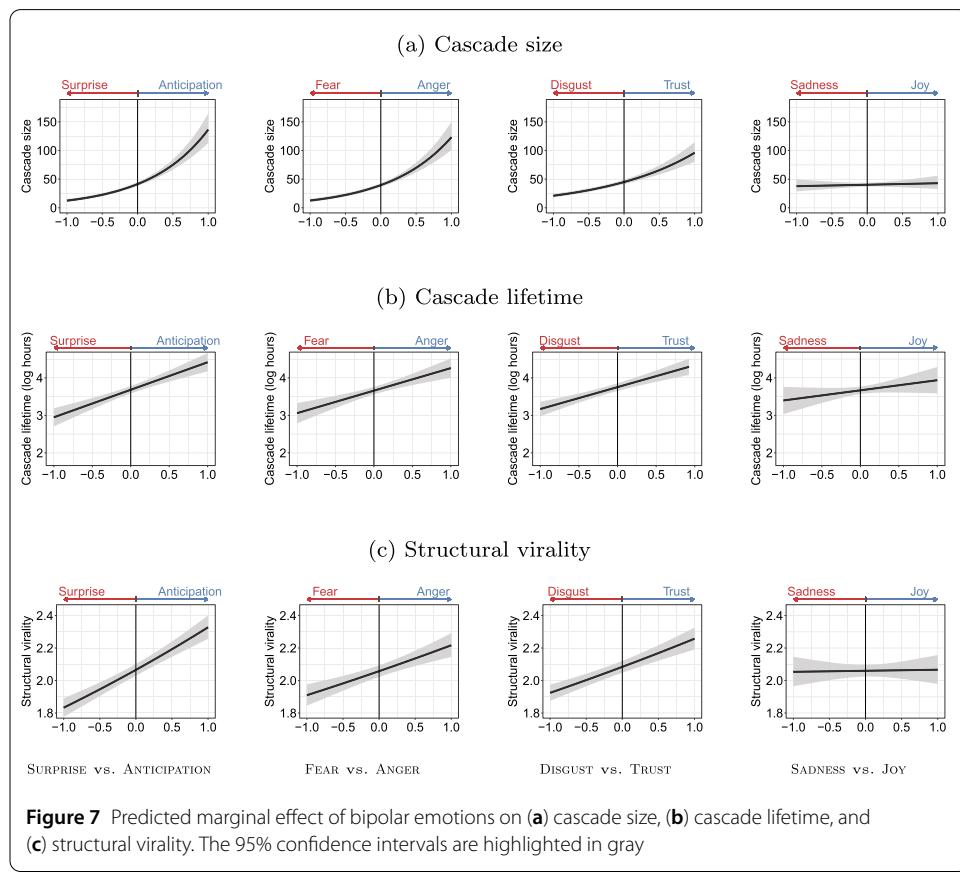
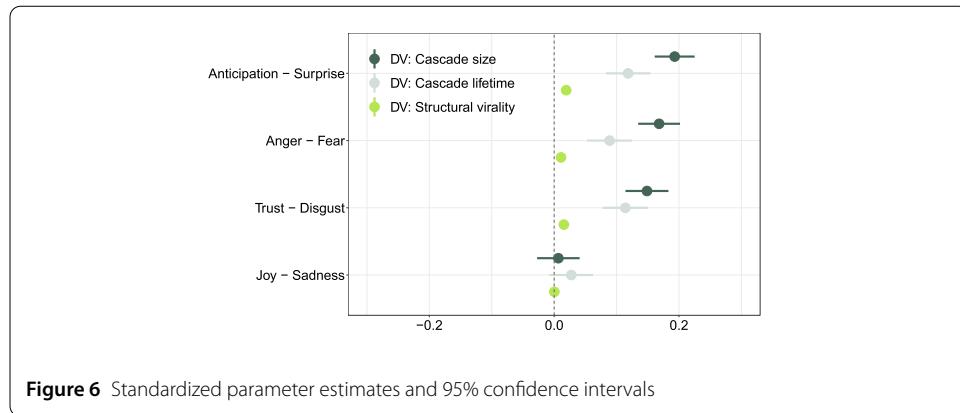
Dyadic Emotional Interactions: Fig. 5 shows the distribution of the dyadic emotional interactions. For the primary emotion dyads, we find that a large proportion of rumors embed CONTEMPT and REMORSE, whereas fewer rumors embed LOVE and SUBMISSION. For the secondary and tertiary emotion dyads, we find that many rumor cascades embed UNBELIEF and SHAME. In contrast, only a relatively small proportion of rumors embed DESPAIR and PESSIMISM.

Note that the above summary statistics only report the relative frequency of emotions but do not allow one to draw conclusions regarding how users respond to emotions. This is studied in the following regression analyses.

3.2 Regression results from bipolar emotion pairs

In the following, we report results for the bipolar emotion pairs ϕ_j^{pairs} .

We use regression analysis to explain different characteristics of cascades based on the bipolar emotion pairs. The parameter estimates in Fig. 6 show that the 8 basic emotions are important determinants of the spreading dynamics of rumors. Across all dependent variables, we find coefficients that are positive and statistically significant for the ANTICIPATION–SURPRISE, ANGER–FEAR, and TRUST–DISGUST dimensions. Hence, rumors are estimated to diffuse more pronouncedly when embedding positive emotions. For instance,



the estimated effect sizes for the ANTICIPATION–SURPRISE pair are as follows: the coefficients amount to 0.193 for cascade size (p -value < 0.001), to 0.118 for cascade lifetime (p -value < 0.001), and to 0.019 for structural virality (p -value < 0.001). Hence, a one standard deviation change in this bipolar emotion pair is linked to a 21.29% increase in the cascade size, a 12.52% increase in the cascade lifetime, and a 1.92% increase in structural virality.

The predicted marginal effects for the bipolar emotion pairs are shown in Fig. 7. Rumors embedding ANTICIPATION, ANGER, and TRUST generate more reshares, spread over a longer

time horizon, and become more viral. The coefficient for the JOY–SADNESS emotion pair is not significant.

Our regression model controls for heterogeneity in users' social influence. The corresponding estimates are omitted for the sake of brevity (their findings have been discussed elsewhere, e.g., in [31]). In short, rumor cascades initiated from accounts that are verified and younger are linked to a larger, longer, and more viral spread. Similar relationships are observed for users exhibiting a higher engagement level and a greater number of followers. In contrast, a higher number of followees is negatively associated with the size, lifetime, and structural virality of a cascade.

We calculated the pseudo- R^2 for each model, resulting in relatively high values of 0.64 for cascade size, 0.43 for cascade lifetime, and 0.31 for structural virality. Evidently, the model variables explain the variation in the dependent variables to a large extent. Furthermore, a visual inspection of the actual vs. fitted plot and goodness-of-fit tests indicate that the models are well specified. This is also supported when considering the differences between the AIC models for individual models estimated with/without emotion variables. For each dependent variable, the difference is greater than the threshold [52] of 10 (difference in cascade size: 226.16; lifetime: 52.22; structural virality: 121.03), indicating strong support for the corresponding candidate models. Therefore, the inclusion of the emotion variables in the regression model is to be preferred.

3.3 Regression results from dyadic emotional interactions

We now study how the presence of 24 dyadic emotional interactions is associated with the diffusion dynamics of online rumors. For this purpose, we employ the previous regression model, but this time include the emotion variables $\phi_j^{primary}$, $\phi_j^{secondary}$, and $\phi_j^{tertiary}$. Figure 8 shows the predicted marginal effects for the 8 primary, 8 secondary, and 8 tertiary dyadic emotional interactions.

Primary dyadic emotional interactions: Rumor cascades with higher values of AGGRESSIONESS, LOVE, OPTIMISM are larger in size, longer-lived, and more viral. We observe no statistically significant effect for the SUBMISSION–CONTEMPT pair. Overall, the largest positive association is observed for AGGRESSIONESS (i.e., the combination of ANTICIPATION and ANGER). An increase of one standard deviation in this dimension is linked to a 19.18% increase in the cascade size, an 8.33% increase in the cascade lifetime, and a 1.69% increase in structural virality.

Secondary dyadic emotional interactions: Rumor cascades with higher values of HOPE vs. UNBELIEF generate more reshares, spread over a longer time horizon, and become more viral. We further find that rumor cascades embedding GUILT, and DESPAIR are negatively associated with the size, lifetime, and structural virality of a cascade. The CURIOSITY–CYNICISM pair is not statistically significant at common statistical significance levels.

Tertiary dyadic emotional interactions: Rumor cascades with higher values of ANXIETY are larger in size, longer-lived, and more viral. We also find a larger size, lifetime, and virality for rumor cascades embedding high levels of DOMINANCE, PESSIMISM, and ANXIETY. We find no statistically significant effect for the SENTIMENTALITY–MORBIDNESS pair.

The control variables tend in a similar direction as in the analysis of the basic emotions. Again, the difference in AIC (comparing the model with and without emotions) is above the common threshold of 10 [52]. Therefore, the models that include emotions are to be preferred.

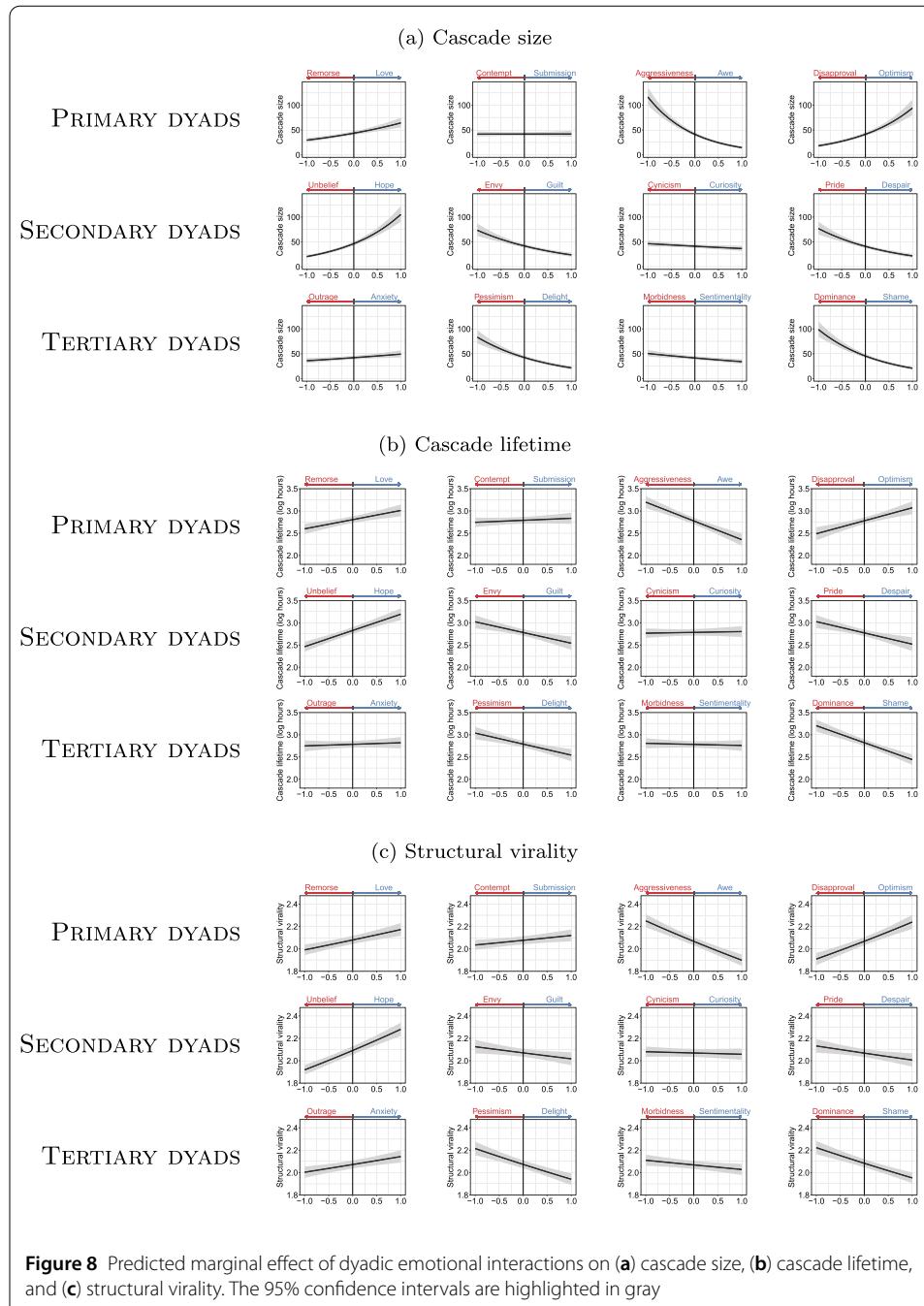
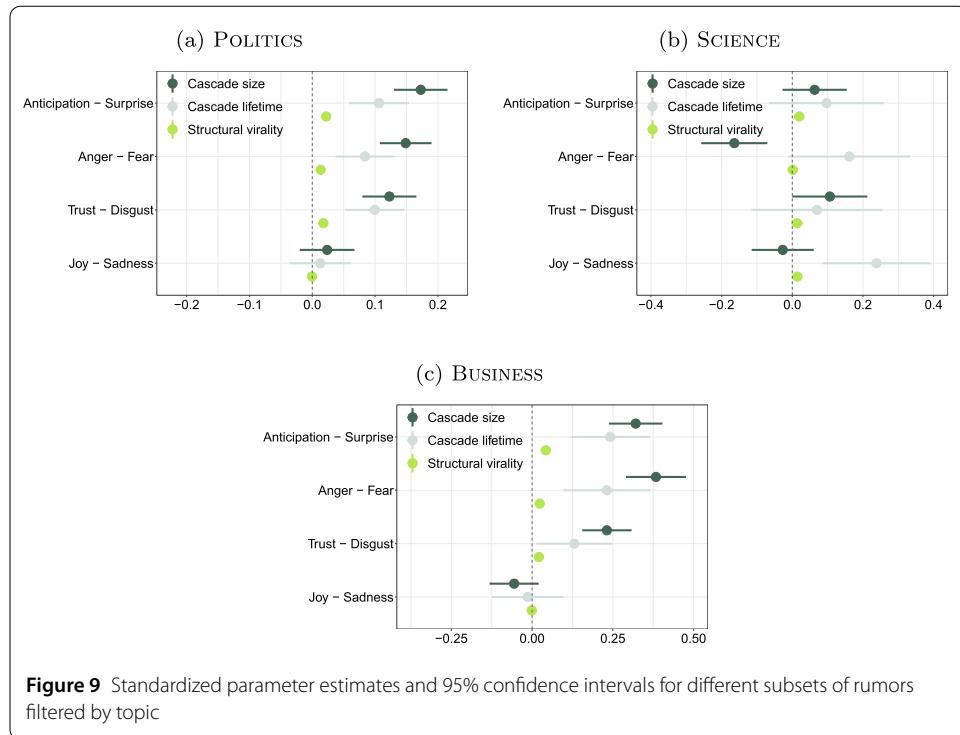


Figure 8 Predicted marginal effect of dyadic emotional interactions on (a) cascade size, (b) cascade lifetime, and (c) structural virality. The 95% confidence intervals are highlighted in gray

3.4 Sensitivity across rumor topics

Our empirical analysis is based on a large-scale dataset with Twitter rumors across varying topics. We now study topic-specific variations. For this purpose, we employ the topic categorization from [11], which classifies Twitter rumors into topics. Here, we focus on the topics POLITICS, BUSINESS, and SCIENCE given their high relevance for society. Note that the topic SCIENCE is broadly defined and also comprises related topics such as health-related rumors. For each of the three topics, we generate a subset of the data and re-estimate our models. The results are visualized in Fig. 9. We find that emotions explain



differences in cascade size, cascade lifetime, and structural virality at a statistically significant level for the topics POLITICS and BUSINESS. In contrast, we find mixed results for SCIENCE. These results are in line with existing literature. For example, [31] find a pronounced role of political content in social media sharing. The authors argue that political topics are more controversial and thus attract more attention, which itself influences sharing behavior.

3.5 Robustness checks

3.5.1 Model checks

We conducted a series of additional model checks that contribute to the robustness of our findings. First, we followed common practice in regression analysis and checked that variance inflation factors as an indicator of multicollinearity were below five [53]. This check led to the desired outcome. Second, we controlled for year-level time effects (i.e., via clustered standard errors and different study horizons) in addition to rumor-level random effects that are already included in our regression model. We obtained conclusive findings. Third, we controlled for non-linear relationships via quadratic terms. In all cases, our findings were supported.

3.5.2 Validation of emotion scores

Our results rely on the validity of dictionaries to extract emotions from online rumors. To check how perceived emotions in rumors align with the dictionary-based emotions, we conducted a survey using the online survey platform Prolific (<https://www.prolific.co/>). We asked $n = 7$ participants (English native speakers) to rate the presence of the eight basic emotions on a Likert scale from -3 to 3 (here: -3 indicates no emotion present while 3 refers to a high degree of emotion present) for a set of 100 randomly sampled rumors. As

Table 1 Kendall's W coefficient for the interrater agreement between survey participants

ANGER	ANTICIPATION	DISGUST	FEAR	JOY	SADNESS	SURPRISE	TRUST
0.474***	0.198***	0.427***	0.406***	0.364***	0.408***	0.227***	0.230***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

shown in Table 1, the participants exhibited a statistically significant interrater agreement according to Kendall's W for each of the 8 basic emotions ($p < 0.01$).

Overall, when aggregating across all 8 basic emotions, the correlation between the dictionary-based emotion scores and human annotations is $\rho = 0.17$ ($p < 0.01$) and thus statistically significant at common significance thresholds. This demonstrates that dictionaries are able to capture emotions in online rumors.

3.5.3 Negation handling

We performed negation scope detection [54, 55] to analyze the robustness to how negations (e.g., “not,” “no”) are handled by the dictionary approach. For example, phrases like “I am surprised” and “I am not surprised” contain the same number of emotional words but convey different emotions to the reader. We analyzed emotional words that are negated by surrounding negation words as follows: (i) We searched for negations using a predefined list of negation words. Here, we used the list of negations from the R package `senti-mentr`. (ii) We recalculated the emotion scores by counting all emotional words in the neighborhood of the negation word as belonging to the opposite emotional dimension (e.g., $Joy = Joy + Sadness_{negated}$). The neighborhood is set to 5 words before and 2 words after the negation. We then compared the emotion scores with negation handling to the values obtained without negation handling. As a result, we found that merely 5.58% of the emotional words in rumors are affected by negations (i.e., lie within negation scopes). Furthermore, the emotion scores with negation handling are highly correlated with the emotion scores without negation handling ($\rho > 0.9$). Altogether, this implies that our analysis and findings are robust to negations.

4 Discussion

In this work, we provided a large-scale study of emotions in online rumor diffusion. For this purpose, 2189 rumors from Twitter with approx. 3.7 million reshares were analyzed with regard to the embedded emotions. Overall, we found that negative emotions are frequently embedded in rumors. Especially frequent are DISGUST (relative share of 23.58%) and SURPRISE (16.44%). (2) The relationship between emotions and the structure of cascade is statistically significant at common significance levels for almost all emotions under study. (3) Rumors embedding ANTICIPATION, ANGER, and, TRUST are estimated to reach a significantly larger number of individuals and diffuse significantly longer and more virally. Interestingly, while negative emotions are more often embedded in rumors, positive emotions are particularly relevant for explaining the diffusion dynamics. (4) A particularly large effect of emotions on the diffusion characteristics is found for AGGRESSIONESS (which is a derived emotion composed of ANTICIPATION and ANGER). A one standard deviation higher level of AGGRESSIONESS is predicted to generate 19.18% more reshares, to be active for 8.33% longer, and to spread 1.69% more virally. Overall, our study establishes emotions as important determinants that describe the spread of online rumor.

Our results contribute to the understanding of online rumor diffusion. As shown by our analysis, emotions are important determinants in explaining the structure of rumor cascades, specifically how many users are involved, the active lifespan and, to a lesser extent, structural virality. The findings are consistent across basic emotions and also dyadic emotion interaction (primary, secondary, tertiary). In addition, our results suggest considerable heterogeneity in the role of emotions. Strong effects are found for most basic emotions (ANTICIPATION, SURPRISE, ANGER, FEAR, TRUST, DISGUST), albeit with the exception of JOY and SADNESS. Similar patterns are observed when studying more complex (derived) emotions. Here, the largest estimated effect size is associated with AGGRESSIONESS. A one standard deviation higher level of AGGRESSIONESS is predicted to generate 19.18% more reshares, cascade that are 8.33% longer, and a 1.69% increase in structural virality. Thereby, we reveal AGGRESSIONESS as a dominant driver of rumor diffusion.

Our work also expands upon rumor theory from offline settings. Offline rumors have a higher chance of dissemination when conveying anxiety [56] and, in particular, negative emotions [42, 43]. However, the underlying evidence stems from offline rumors rather than online rumors. Our work adds in two ways: First, we study the role of emotions in the diffusion of *online* rumors. While rumor diffusion in offline settings is more pronounced for negative emotions, we observe the opposite for online rumors, for which positive emotions appear more influential. Second, we not only compare positive vs. negative emotions but perform a granular study across primary, secondary, and tertiary emotional dyadic interactions. This provides rich findings on the *heterogeneity* of emotion effects. As such, we confirm that ANXIETY is an important driver for rumor diffusion not only in offline but also in online settings. However, further emotions are also relevant: a particularly pronounced role is found with regard to AGGRESSIONESS. To the best of our knowledge, the importance of AGGRESSIONESS in rumor diffusion was previously overlooked.

In our study, inferences were made based on data from Twitter. Twitter has a wide popularity with more than 300 million active users. In addition, it plays an important part in rumor diffusion due to its influential role in the political discourse [10]. This makes our findings directly relevant to both social media platforms and, in particular, public stakeholders. For the same reason, established procedures were followed when compiling the data [11], as this ensures that findings are drawn from a realistic, large-scale dataset of Twitter rumors. To the best of our knowledge, our work is the first statistical analysis linking emotions to online rumor diffusion.

As with other studies, ours is subject to limitations that provide opportunities for future research. First, this study is based on observational inferences, while we leave the extension to (quasi-)experimental settings, and thus causal inferences, to future work. Nevertheless, our study design ensures that many potential confounding factors can be ruled out. This is because of the temporal order (i.e., the emotion-laden wording *precedes* the actual cascade) and the fact that further sources of variability among rumors are captured through rumor-level random effects. Second, our study employs statistical inferences that provide explanatory insights. This allows us to quantify the *marginal* contribution of emotions to online rumor diffusion. A different objective is to use emotions for predictive modeling, which is discussed elsewhere [57–60].

Our work entails several implications. It emphasizes the necessity of considering emotions when studying rumor diffusion. Emotions are also relevant in practice, particularly for social media platforms. To counter the proliferation of online rumors, social media

platforms should seek solutions, based on which emotions can be actively managed. Our study also encourages a granular investigation of emotions for related research questions, whereby not only basic emotions but also derived emotions are considered. Such granular analyses are comparatively more challenging in lab experiments; however, a remedy is offered by computational social science based on which large-scale datasets from online behavior can be mined.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1140/epjds/s13688-021-00307-5>.

Additional file 1. S1—Background Literature (PDF 90 kB)

Funding

Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

All data needed to evaluate the conclusions in the paper are publicly available (and the source reported in the paper). Replication code for this study is available via https://github.com/DominikBaer95/Emotions_Rumor_Diffusion.

Declarations

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

NP and SF designed the study. NP and DB analyzed the data. NP, DB, and SF wrote and revised the manuscript. All authors reviewed the manuscript. All authors read and approved the final manuscript.

Author details

¹University of Giessen, Licher Str. 62, 35394 Giessen, Germany. ²LMU Munich, Geschwister-Scholl-Platz 1, 80539 Munich, Germany. ³ETH Zurich, Weinbergstr. 56/58, 8092 Zurich, Switzerland.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 18 May 2021 Accepted: 20 September 2021 Published online: 18 October 2021

References

- Berger J (2011) Arousal increases social transmission of information. *Psychol Sci* 22(7):891–893. <https://doi.org/10.1177/0956797611413294>
- Berger J, Milkman KL (2012) What makes online content viral? *J Mark Res* 49(2):192–205. <https://doi.org/10.1509/jmr.10.0353>
- Leskovec J, Adamic LA, Huberman BA (2007) The dynamics of viral marketing. *ACM Trans Web* 1(1):5. <https://doi.org/10.1145/1232722.1232727>
- Godes D, Mayzlin D (2004) Using online conversations to study word-of-mouth communication. *Mark Sci* 23(4):545–560. <https://doi.org/10.1287/mksc.1040.0071>
- de Domenico M, Lima A, Mougel P, Musolesi M (2013) The anatomy of a scientific rumor. *Sci Rep* 3:2980. <https://doi.org/10.1038/srep02980>
- Starbird K, Maddock J, Orand M, Achterman P, Mason RM (2014) Rumors, false flags, and digital vigilantes: misinformation on Twitter after the 2013 Boston marathon bombing. In: iConference
- Starbird K (2017) Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on Twitter. In: International AAAI conference on web and social media (ICWSM)
- Aral S, Eckles D (2019) Protecting elections from social media manipulation. *Science* 365(6456):858–861. <https://doi.org/10.1126/science.aaw8243>
- Bakshy E, Messing S, Adamic LA (2015) Exposure to ideologically diverse news and opinion on Facebook. *Science* 348(6239):1130–1132. <https://doi.org/10.1126/science.aaa1160>
- Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D (2019) Fake news on Twitter during the 2016 U.S. presidential election. *Science* 363(6425):374–378. <https://doi.org/10.1126/science.aau2706>
- Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359(6380):1146–1151. <https://doi.org/10.1126/science.aap9559>
- Zaman T, Fox EB, Bradlow ET (2014) A Bayesian approach for predicting the popularity of tweets. *Ann Appl Stat* 8(3):1583–1611. <https://doi.org/10.1214/14-AOAS741>

13. Cha M, Mislove A, Gummadi KP (2009) A measurement-driven analysis of information propagation in the Flickr social network. In: International world wide web conference (WWW)
14. Kwak H, Lee C, Park H, Moon S (2010) What is Twitter, a social network or a news media? In: International world wide web conference (WWW). <https://doi.org/10.1145/1772690.1772751>.
15. Lerman K, Ghosh R (2010) Information contagion: an empirical study of spread of news on Digg and Twitter social networks. In: International AAAI conference on web and social media (ICWSM)
16. Arif A, Shanahan K, Chou F-J, Dosouto Y, Starbird K, Spiro ES (2016) How information snowballs: exploring the role of exposure in online rumor propagation. In: ACM conference on computer-supported cooperative work & social computing (CSCW). <https://doi.org/10.1145/2818048.2819964>
17. Spiro ES, Fitzhugh S, Sutton J, Pierski N, Greczek M, Butts CT (2012) Rumoring during extreme events: a case study of deepwater horizon 2010. In: ACM web science conference (WebSci). <https://doi.org/10.1145/2380718.2380754>
18. Kryvasheyev Y, Chen H, Obradovich N, Moro E, van Hentenryck P, Fowler J, Cebrian M (2016) Rapid assessment of disaster damage using social media activity. *Sci Adv* 2(3):1500779
19. Zeng L, Starbird K, Spiro ES (2016) Rumors at the speed of light? Modeling the rate of rumor transmission during crisis. In: Hawaii international conference on system sciences (HICSS). <https://doi.org/10.1109/HICSS.2016.248>
20. Sauter DA, Eisner F, Ekman P, Scott SK (2010) Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc Natl Acad Sci USA* 107(6):2408–2412. <https://doi.org/10.1073/pnas.0908239106>
21. Ekman P (1992) An argument for basic emotions. *Cogn Emot* 6(3–4):169–200. <https://doi.org/10.1080/02699939208411068>
22. 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. *Am Sci* 89(4):344–350
23. Zhang P (2013) The affective response model: a theoretical framework of affective concepts and their relationships in the ICT context. *MIS Q* 37(1):247–274
24. Barsade SG (2002) The ripple effect: emotional contagion and its influence on group behavior. *Adm Sci Q* 47(4):644. <https://doi.org/10.2307/3094912>
25. Kramer ADI, Guillory JE, Hancock JT (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proc Natl Acad Sci USA* 111(24):8788–8790. <https://doi.org/10.1073/pnas.1320040111>
26. Goldenberg A, Gross JJ (2020) Digital emotion contagion. *Trends Cogn Sci* 24(4):316–328. <https://doi.org/10.1016/j.tics.2020.01.009>
27. Fan R, Varol O, Varamesh A, Barron A, van de Leemput IA, Scheffer M, Bollen J (2019) The minute-scale dynamics of online emotions reveal the effects of affect labeling. *Nat Hum Behav* 3(1):92–100. <https://doi.org/10.1038/s41562-018-0490-5>
28. Ferrara E, Yang Z (2015) Measuring emotional contagion in social media. *PLoS ONE* 10(11):e0142390. <https://doi.org/10.1371/journal.pone.0142390>
29. Alvarez R, Garcia D, Moreno Y, Schweitzerer F (2015) Sentiment cascades in the 15M movement. *EPJ Data Sci* 4(1):407. <https://doi.org/10.1140/epjds/s13688-015-0042-4>
30. Zollo F, Novak PK, Del Vicario M, Bessi A, Mozzetti I, Scala A, Caldarelli G, Quattrociocchi W (2015) Emotional dynamics in the age of misinformation. *PLoS ONE* 10(9):e0138740. <https://doi.org/10.1371/journal.pone.0138740>
31. Stieglitz S, Dang-Xuan L (2013) Emotions and information diffusion in social media: sentiment of microblogs and sharing behavior. *J Manag Inf Syst* 29(4):217–248. <https://doi.org/10.2753/MIS0742-1222290408>
32. Naveed N, Gottron T, Kunegis J, Alhadi AC (2011) Bad news travel fast: a content-based analysis of interestingness on Twitter. In: International web science conference (WebSci). <https://doi.org/10.1145/2527031.2527052>
33. Kim J, Yoo J (2012) Role of sentiment in message propagation: reply vs. retweet behavior in political communication. In: International conference on social informatics. <https://doi.org/10.1109/SocialInformatics.2012.33>
34. Kissler J, Herbert C, Peyk P, Junghofer M (2007) Buzzwords: early cortical responses to emotional words during reading. *Psychol Sci* 18(6):475–480. <https://doi.org/10.1111/j.1467-9280.2007.01924.x>
35. Luminet O, Bouts P, Delie F, Manstead ASR, Rimé B (2000) Social sharing of emotion following exposure to a negatively valenced situation. *Cogn Emot* 14(5):661–688. <https://doi.org/10.1080/02699930050117666>
36. Rimé B (2009) Emotion elicits the social sharing of emotion: theory and empirical review. *Emot Rev* 1(1):60–85. <https://doi.org/10.1177/1754073908097189>
37. Peters K, Kashima Y, Clark A (2009) Talking about others: emotionality and the dissemination of social information. *Eur J Soc Psychol* 39(2):207–222. <https://doi.org/10.1002/ejsp.523>
38. Chuai Y, Zhao J (2020) Anger makes fake news viral online. [arXiv:2004.10399](https://arxiv.org/abs/2004.10399)
39. Wu S, Tan C, Kleinberg J, Macy M (2011) Does bad news go away faster? In: International AAAI conference on web and social media (ICWSM)
40. Bakshy E, Hofman JM, Mason WA, Watts DJ (2011) Everyone's an influencer. In: International conference on web search and data mining (WSDM). <https://doi.org/10.1145/1935826.1935845>
41. Brady WJ, Wills JA, Jost JT, Tucker JA, van Bavel JJ (2017) Emotion shapes the diffusion of moralized content in social networks. *Proc Natl Acad Sci USA* 114(28):7313–7318. <https://doi.org/10.1073/pnas.1618923114>
42. Anthony S (1973) Anxiety and rumor. *J Soc Psychol* 89(1):91–98. <https://doi.org/10.1080/00224545.1973.9922572>
43. Knapp RH (1944) A psychology of rumor. *Public Opin Q* 8(1):22–37
44. Martel C, Pennycook G, Rand DG (2020) Reliance on emotion promotes belief in fake news. *Cogn Res Princ Implic* 5(1):47. <https://doi.org/10.1186/s41235-020-00252-3>
45. Weeks BE (2015) Emotions, partisanship, and misperceptions: how anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *J Commun* 65(4):699–719. <https://doi.org/10.1111/jcom.12164>
46. Acerbi A (2019) Cognitive attraction and online misinformation. *Palgrave Commun* 5(1):15. <https://doi.org/10.1057/s41599-019-0224-y>
47. Goel S, Anderson A, Hofman J, Watts DJ (2016) The structural virality of online diffusion. *Manag Sci* 62(1):180–196. <https://doi.org/10.1287/mnsc.2015.2158>
48. Allport GW, Postman L (1947) The psychology of rumor. Holt, New York
49. Kratzwald B, Ilić S, Kraus M, Feuerriegel S, Prendinger H (2018) Deep learning for affective computing: text-based emotion recognition in decision support. *Decis Support Syst* 115:24–35. <https://doi.org/10.1016/j.dss.2018.09.002>

50. Mohammad SM, Turney PD (2013) Crowdsourcing a word-emotion association lexicon. *Comput Intell* 29(3):436–465. <https://doi.org/10.1111/j.1467-8640.2012.00460.x>
51. Breiman L (2001) Statistical modeling: the two cultures. *Stat Sci* 16(3):199–231
52. Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociol Methods Res* 33(2):261–304
53. Akinwande MO, Dikko HG, Samson A et al (2015) Variance inflation factor: as a condition for the inclusion of suppressor variable (s) in regression analysis. *Open J Stat* 5(7):754–767
54. Pröllochs N, Feuerriegel S, Neumann D (2019) Learning interpretable negation rules via weak supervision at document level: a reinforcement learning approach. In: Conference of the North American chapter of the association for computational linguistics: human language technologies (NAACL-HLT)
55. Pröllochs N, Feuerriegel S, Lutz B, Neumann D (2020) Negation scope detection for sentiment analysis: a reinforcement learning framework for replicating human interpretations. *Inf Sci* 536:205–221. <https://doi.org/10.1016/j.ins.2020.05.022>
56. Rosnow RL (1991) Inside rumor: a personal journey. *Am Psychol* 46(5):484–496
57. Castillo C, Mendoza M, Poblete B (2011) Information credibility on Twitter. In: International world wide web conference (WWW). <https://doi.org/10.1145/1963405.1963500>
58. Kwon S, Cha M, Jung K, Chen W, Wang Y (2013) Prominent features of rumor propagation in online social media. In: International conference on data mining (ICDM). <https://doi.org/10.1109/ICDM.2013.61>
59. Kwon S, Cha M, Jung K (2017) Rumor detection over varying time windows. *PLoS ONE* 12(1):e0168344. <https://doi.org/10.1371/journal.pone.0168344>
60. Ducci F, Kraus M, Feuerriegel S (2020) Cascade-LSTM: a tree-structured neural classifier for detecting misinformation cascades. In: ACM SIGKDD conference on knowledge discovery and data mining (KDD)

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com