

RESEARCH ARTICLE

Measuring Emotional Contagion in Social Media

Emilio Ferrara^{1,2}*, Zeyao Yang¹

1 School of Informatics and Computing, Indiana University, Bloomington, IN, United States of America, **2** Information Sciences Institute, University of Southern California, Marina Del Rey, CA, United States of America

* These authors contributed equally to this work.

* ferrarae@isi.edu



OPEN ACCESS

Citation: Ferrara E, Yang Z (2015) Measuring Emotional Contagion in Social Media. PLoS ONE 10 (11): e0142390. doi:10.1371/journal.pone.0142390

Editor: Chris T. Bauch, University of Waterloo, CANADA

Received: June 17, 2015

Accepted: October 21, 2015

Published: November 6, 2015

Copyright: © 2015 Ferrara, Yang. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data have been collected through the public Twitter API (<https://dev.twitter.com/overview/api>). To comply with Twitter terms of service, data cannot be publicly shared. Interested future researchers may reproduce the experiments by following the procedure described in the paper. Anonymized data may be available upon request from Dr. Emilio Ferrara (ferrarae@isi.edu).

Funding: ZY was partly supported by NSF grant no. IIS-0811994. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Social media are used as main discussion channels by millions of individuals every day. The content individuals produce in daily social-media-based micro-communications, and the emotions therein expressed, may impact the emotional states of others. A recent experiment performed on Facebook hypothesized that emotions spread online, even in absence of non-verbal cues typical of in-person interactions, and that individuals are more likely to adopt positive or negative emotions if these are over-expressed in their social network. Experiments of this type, however, raise ethical concerns, as they require massive-scale content manipulation with unknown consequences for the individuals therein involved. Here, we study the dynamics of emotional contagion using a random sample of Twitter users, whose activity (and the stimuli they were exposed to) was observed during a week of September 2014. Rather than manipulating content, we devise a null model that discounts some confounding factors (including the effect of emotional contagion). We measure the emotional valence of content the users are exposed to before posting their own tweets. We determine that on average a negative post follows an over-exposure to 4.34% more negative content than baseline, while positive posts occur after an average over-exposure to 4.50% more positive contents. We highlight the presence of a linear relationship between the average emotional valence of the stimuli users are exposed to, and that of the responses they produce. We also identify two different classes of individuals: highly and scarcely susceptible to emotional contagion. Highly susceptible users are significantly less inclined to adopt negative emotions than the scarcely susceptible ones, but equally likely to adopt positive emotions. In general, the likelihood of adopting positive emotions is much greater than that of negative emotions.

Introduction

The study of socio-technical systems, and their effects on our increasingly interconnected society, is playing a significant role in the emerging field of *computational social science* [1–7]. Online social platforms like Facebook and Twitter provide millions of individuals with near-

unlimited access to information and connectivity [8–10]. The content produced on such platforms has proved to impact society at large: from social and political discussions [11–16], to emergency and disaster response [17–19], social media conversation affects the offline, physical world in tangible ways.

The central issue that inspires this work is how the content produced and consumed on social media affects individuals emotional states and behaviors. We are concerned in particular with the theory of emotional contagion [20]. Data from a 20-years longitudinal study suggest that emotions can be passed via social networks, and have long-term effects [21]. Various recent contributions advanced the hypothesis that emotions may be passed also via online interactions [22–28]. A recent study performed by Facebook suggests that emotional contagion occurs online even in absence of non-verbal cues typical of in-person interactions [29]. The authors of such study performed a controlled experiment selecting a sample of users and, by manipulating the content on their time-lines, exposed some to increased levels of positive or negative emotions, as conveyed by the posts produced by their contacts. This experiment revealed a small but significant correlation between the number of emotionally positive/negative words in users' posts and that of the stream they have been exposed to.

The possibility to manipulate the information that users see is clearly well suited to address questions about the existence and magnitude of emotional contagion, but raises ethical concerns [30–32]: the consequences of massive-scale content manipulations are unknown, and might include long-term effects on the mental and physical well-being of individuals.

In this study, we use Twitter as case study, and explore the hypothesis of emotional contagion via the social stream. A reasonable expectation is that Twitter connections carry a smaller emotional contagion power than Facebook ones: users generally adopt these platforms for different purposes —Twitter for information sharing [8], and Facebook to keep in touch with family and friends (or other social internetworking activities) [33–35]. Yet, a recent neuroscience study found that “*reading a Twitter timeline generates 64 percent more activity in the parts of the brain known to be active in emotion than normal Web use; tweeting and retweeting boosts that to 75 percent more than a run-the-mill website*” (This is your brain on Twitter: <https://medium.com/backchannel/this-is-your-brain-on-twitter-cac0725cea2b>). In our approach we observe the Twitter stream without performing content manipulation or re-engineering of any type (no information filtering, prioritization, ranking, etc.). We rather devise a clever null model that tries to discount for emotional contagion and other correlational biases, and a method to reconstruct the stimuli (in terms of contents and their emotions) users were exposed to before posting their tweets. This allows us to delve into the theory of emotional contagion studying single individuals and their responses to different emotions: our analysis suggests a significant presence of emotional contagion. We show that negative posts on average follow a 4.34% over-exposure to negative contents prior to their production, while positive tweets occur on average after a 4.50% over-exposure to positive contents. We infer a linear relationship between the emotional valence of the stimuli and the response for a sample of users whose activity, and the activity of all their followees, has been monitored for an entire week during September 2014. Our experiments highlight that different extents of emotional contagion may occur: in particular, we identify two classes of individuals, namely those highly or scarcely susceptible to emotional contagion. These two classes respond differently to different stimuli: highly susceptible individuals are less inclined to adopt negative emotions but equally likely to adopt positive emotions than the scarcely susceptible ones. Also, the adoption rate of positive emotions is in general greater than that of negative emotions.

It is worth noting upfront that the observational nature of the experiments, and the technical limits posed by sentiment analysis algorithms, make emotional contagion a plausible yet not exclusive explanation: (i) the presence of confounding factors, including network effects

like *homophily* and *latent homophily*, may effect the size of the observed effects; (ii) even the state of the art among sentiment analysis algorithms, like SentiStrength here employed, is not able to capture complex language nuances such as sarcasm or irony; (iii) finally, emotional contagion may be mixed to other emotional alignment effects, such as empathy or sympathy. In the Discussion section we detail these issues.

Our work furthers the understanding of human emotions expressed via online interactions while avoiding the inconveniences and ethically-problematic consequences of previous experimental work carried out on other social platforms.

Materials and Methods

Sentiment Analysis

The analysis of the emotional valence of content can be leveraged to produce reliable forecasts in a variety of different circumstances. [36–40]. There exists a variety of sentiment analysis algorithms able to capture positive and negative sentiment, some specifically designed for short, informal texts [41–43]. In this work, we use SentiStrength [44–46] to annotate the tweets with positive and negative sentiment scores. Compared with other tools, SentiStrength provides several advantages: it is designed for short informal texts with abbreviations and slang (features commonly observed in Twitter), and it employs linguistic rules for negations, amplifications, booster words, emoticons, spelling corrections, particularly well suited to process social media data. SentiStrength was proven able to capture positive emotions with 60.6% accuracy and negative emotions with 72.8% accuracy on MySpace [44–46].

SentiStrength assigns to each tweet t a positive $S^+(t)$ and negative $S^-(t)$ sentiment score. Both scores are on a scale ranging between 1 (neutral) and 5 (strongly positive or negative). To capture in one single measure the sentiment expressed by each tweet, we define the *polarity score* $S(t)$ as the difference between positive and negative sentiment scores assigned to tweet t :

$$S(t) = S^+(t) - S^-(t). \quad (1)$$

The polarity score S ranges between -4 (extremely negative: $S^+(t) = 1$ and $S^-(t) = 5$) to +4 (extremely positive: $S^+(t) = 5$ and $S^-(t) = 1$). When positive and negative sentiment scores for tweet t are the same ($S^+(t) = S^-(t)$), we say that the polarity of tweet t is neutral ($S(t) = 0$). The choice of focusing on the polarity score rather than on both positive and negative sentiment scores is justified by previous studies that showed how is preferable to measure the overall sentiment rather than the intensity of sentiment when dealing with short pieces of texts like tweets [40, 44–46] —this is intuitively due to the paucity of information conveyed in 140 characters, and the somewhat simplicity of the sentiment analysis tools adopted as opposed to the intrinsic difficulty of the task. The distribution of polarity scores is peaked around neutral tweets, and overall slightly skewed toward positiveness (see Fig 1 and the related discussion). We also observed that extreme values of positive and negative tweets are comparably represented, assessing that the algorithm is not producing systematically biased results. In such cases, we observed that among the most recurring keywords in the tweets with negative polarities exhibit have feelings like anger (*hate, blame, bored, tired, annoyed, etc.*), fear (*scared, lonely, sadness, etc.*), and contain cussing (*wtf, omfg, fuck, etc.*), negative superlative adjectives (*worst, weirdest, nastiest, grossest, etc.*); on the other hand, tweets annotated as positive intuitively exhibit feelings like joy, excitement, happiness, love, etc.

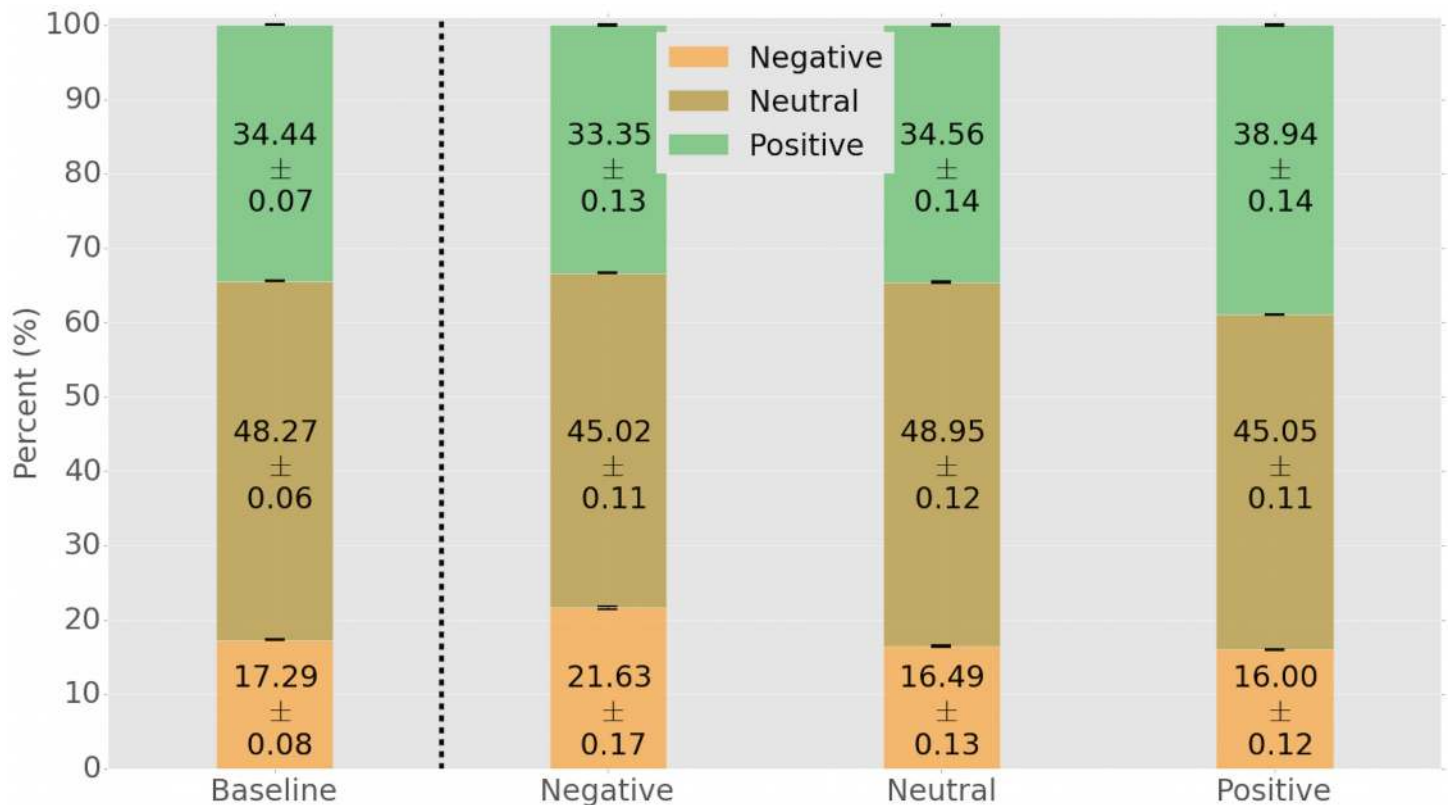


Fig 1. Average proportions of positive, neutral, and negative emotions prior to each observed tweet. The Baseline model (left) discounts for the effect of emotional contagion by means of a reshuffling strategy. The three bars (Negative, Neutral, and Positive) respectively show the average proportions of emotions prior to posting a negative, neutral, or positive tweet. For each negative tweet posted, on average its author was previously exposed to about 4.34% more negative tweets than expected by the Baseline model. For each positive tweet posted, on average its author was previously exposed to about 4.50% more positive content. Note how the distribution of emotions before posting a neutral tweet almost perfectly matches that of the Baseline model. The numbers inside the columns represent the exact proportions \pm the standard errors. Error bars represent standard errors.

doi:10.1371/journal.pone.0142390.g001

Data

Our goal is to establish a relation between the sentiment of a tweet and that of the tweets that its author may have seen in a short time period preceding its posting. To achieve that, we first collected a set U consisting of a random sample of 3800 users who posted at least one tweet in English (among those provided by the Twitter *gardenhose*) in the last week of September 2014. Via the appropriate Twitter API we also collected the set F of followees of all users in U .

For each tweet t produced by an user u in U in said last week of September 2014, we constructed h_t , the set of tweets produced by any of u 's followees in a time span of one hour preceding the posting of t .

For the purpose of our analysis we considered only tweets t such that $|h_t| \geq 20$. Also, we considered only tweets (i) in English, and (ii) that do not contain URLs or media content (photos, videos, etc.). Finally, each tweet, both from the target set of users and their followees were annotated by their sentiment score as discussed above.

It is worth to briefly justify some of the choice we made. The filter English + non-media was applied to be able to unambiguously attribute a sentiment score to the tweets. The choice of limiting ourselves to sampling from the last week of September was dictated by the technical

limitations of the Twitter API to recover the 100% of tweets posted by any given user only to one week prior to the query time. This precaution allows us to discount for possible sampling issues, so to reconstruct the full exposures to contents prior to any posting from this established set of users. Dealing with the 100% of the content excludes possible sampling biases common to many social media studies [47]. The choice to focus on tweets for which the user was exposed to at least 20 tweets within 1 hour from their posting allows us to obtain a significant description of the stimuli the users were exposed to.

We finally separated all tweets in three classes of emotions: negative (polarity score $S \leq -1$), neutral ($S = 0$), and positive ($S \geq 1$). Focusing on the classes of emotions rather than the intensity of emotions will facilitate our analysis and also discount for possible inaccuracies of the sentiment analysis procedure: several previous studies showed that it is much easier to capture the overall emotion of a short piece of text, rather than emotion intensities [41–46]. We experimented with other thresholds and the results presented later don't vary, exhibiting the same effects: the only differences are the proportions of tweets assigned to the different classes.

Results

Effect of emotional contagion

We here want to test the hypothesis that emotional contagion occurs among social media users, as suggested by recent works on various social platforms [21, 29]. The idea is that emotions can be passed via online interactions even in absence of non-verbal cues typical of in-person interactions, which are deemed by traditional psychology to be an essential ingredient for emotional contagion [20]. To test this hypothesis, we need to reconstruct the emotions conveyed by the tweets each user was exposed to before posting their own tweets: this will allow us to determine whether the stimuli are correlated with the responses, namely the emotions subsequently expressed by the user.

Our study is purely observational, as we don't perform any type of controlled experiment differently from other works [29]. We aim to show that the average sentiment of tweets preceding a positive, negative or neutral tweet are significantly different, and determine the effect size which, even if small, at scale would have important implications.

To do so, we adopt the following reshuffling strategy aimed at determining the baseline distributions of positive, neutral, and negative contents independently of emotional contagion: for each user u in the set of 3,800 users, and for each tweet t_u produced by u , we have the history $\ell(t_u)$ of all tweets preceding t_u in the 1 hour period prior to t_u 's publication, and we record how many such tweets $s_{\ell(t_u)}$ user u was exposed to. We then put all these tweets $\ell(t_u)$, that represent the stimuli prior to the users' activities, for all tweets, for all users, in one single bucket.

To create our reshuffled null model that discounts for the effect of emotional contagion, we therefore sample with replacement from bucket B , for each tweet t_u of each user u , a number of tweets equal to the size $s_{\ell(t_u)}$. The results for sampling without replacement are substantially identical. At the end of the procedure, we obtain a baseline distribution of positive, neutral, and negative sentiment prior to the publication of any tweet, which discounts for the effect of exposure and the possibility of emotional contagion. The baseline distribution of sentiment in the null model is displayed in Fig 1: the proportion of positive, neutral, and negative sentiment after the exposure reshuffling is equal to, respectively, 34.44% (± 0.07), 48.27% (± 0.06), and 17.29% (± 0.08). These proportions reflect the three classes of emotions defined as follows: negative ($S \leq -1$), neutral ($S = 0$), and positive ($S \geq 1$).

To verify the hypothesis of emotional contagion, we divide all tweets t_u posted by each user u , in three categories (positive, neutral, and negative) according to their sentiment. For each

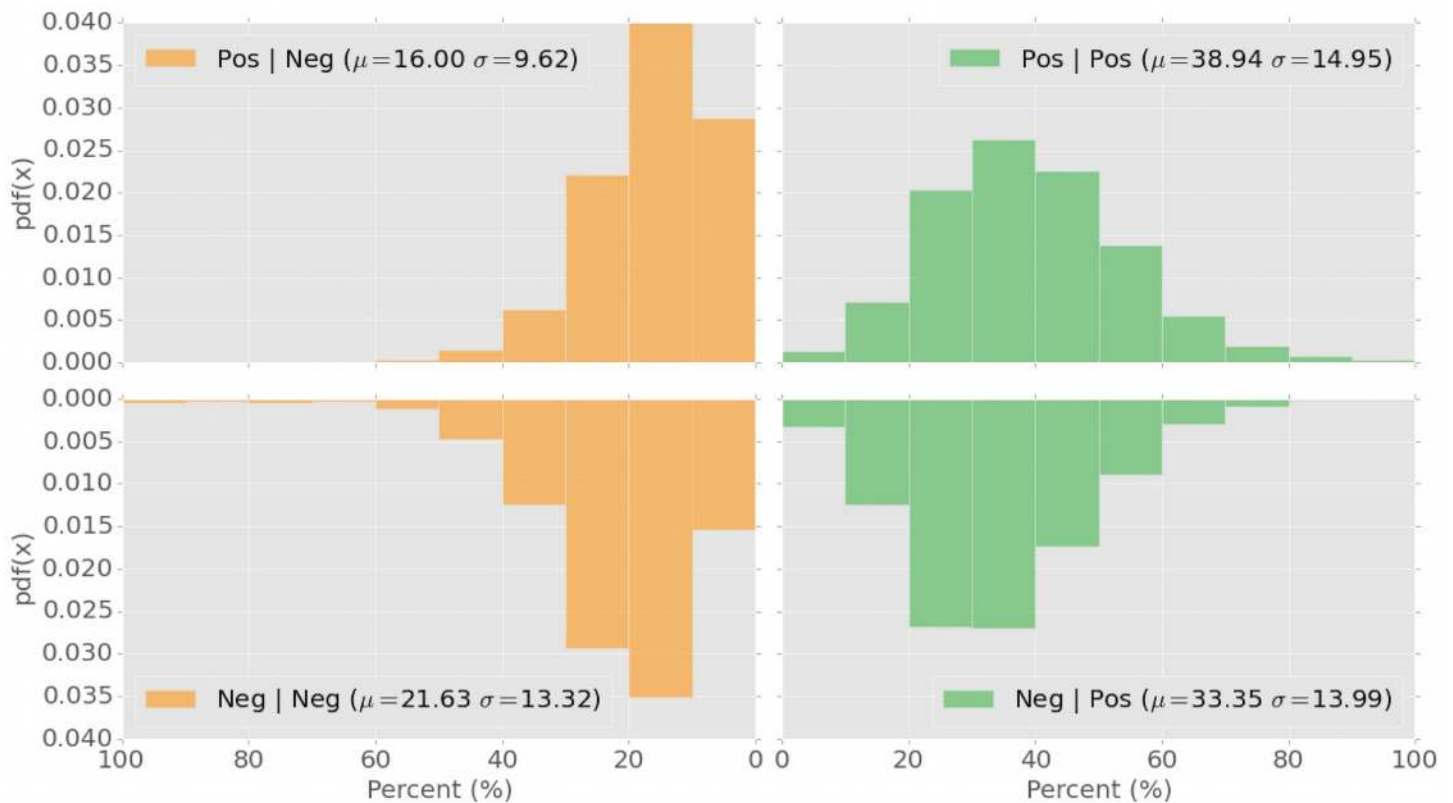


Fig 2. Distributions of positive and negative stimuli before positive and negative responses. The four quadrants show the probability distributions of a negative response prior to a negative (bottom left) or positive (bottom right) stimulus, or a positive response prior to a negative (top left) or positive (top right) stimulus.

doi:10.1371/journal.pone.0142390.g002

category, then, we generate the distribution of fraction of positive, neutral, and negative sentiments observed in the stimuli, the tweets produced by u 's followees prior to the posting of each t_u . The results, displayed in Fig 1, are interpreted as follows: the three stacked-columns identify the distributions of sentiment prior to posting (from left to right) a negative, neutral, or positive tweet. For example, a user in our set prior to posting a negative tweet is exposed, on average, to 21.63% (± 0.17) negative tweets, 45.02% (± 0.11) neutral, and 33.35% (± 0.13) positive ones. This signifies an over-exposure to 4.34% more negative tweets, at the expenses of 1.09% less positive ones, if compared with our null model in Fig 1. Similarly, prior to posting a positive tweet, a user in our dataset is exposed, on average, to 16.00% (± 0.12) negative tweets, 45.05% (± 0.11) neutral, and 38.94% (± 0.14) positive ones. This amounts for an over-exposure of 4.50% more positive tweets, at the expenses of 1.29% less negative ones, if compared with the null model. Notably, the distribution of the sentiment of tweets before the posting of a neutral one matches almost perfectly the distribution of the null model in Fig 1, suggesting that no emotional contagion occurs in the case of neutral tweets. To prove the statistical significance of these differences, we run a Mann-Whitney U test between the observed distributions in presence of emotional contagion, and the expected baseline of the null model. Both p values for negative and positive emotional contagion tests are $p < 10^{-6}$ while no significant difference occurs for the neutral case; the strength of the statistical significance is further illustrated by the narrow error bars in Fig 1. The distributions of the positive and negative stimuli, respectively, before positive and negative responses, are also reported in Fig 2.

These results suggest the presence of emotional contagion for both negative and positive sentiment, and seem to show that no emotional contagion occurs prior to posting neutral contents. To verify that these findings were not strongly dependent on some particular conditions, we performed additional experiments and observed consistent results across different comparable datasets (not discussed here to avoid confusion) and sampling methods.

To further validate this hypothesis, and in particular to focus only on positive and negative contagion, we here propose another measure, that we call *valence*, that can be computed on any set (bucket) of tweets for which the sentiment is computed. Given a bucket of tweets b , its valence $V(b)$ is given by the following formula:

$$V(b) = 2 \cdot \frac{p_b}{n_b + p_b} - 1 \quad (2)$$

where p_b and n_b represent, respectively, the fraction of positive and negative tweets in bucket b . This measure ranges between -1 and +1: the lower the score, the larger the disproportion toward negative emotion, and vice-versa.

Since for each tweet t_u produced by each user u we already obtained the history $\ell(t_u)$ of all tweets preceding t_u in the 1 hour period prior to t_u 's publication, we can compute the valence scores $V(\ell(t_u))$ for all histories. This allows us to represent the difference in intensity between positive and negative stimuli each user u was exposed to prior to posting each tweet t_u . Therefore, we calculate the valence scores $V(\ell(t_u))$ for all tweets t_u in our dataset. This generates a distribution of values between -1 and +1, each value representing the valence of the stimulus of the associated tweet. We then bin these *stimuli valence* values, in 20 bins of length 0.05 (see the x -axis of Fig 3). Each bin x_b contains, again, a set of tweets (the responses) for which we already calculated the sentiment (positive, negative, or neutral). We can calculate also the valence of each x_b . Such values will represent the *response valence* for a given value (bin) of stimulus valence. The results, illustrated in Fig 3, show a very strong linear relationship ($R^2 = 0.975$) between the valence of the stimulus and the valence of the response. For example, a very strong negative stimulus with valence -1 generates a response valence of about -0.8. Similarly, a very strong positive stimulus of valence +1 will trigger a response of valence around +0.6. Other regression models have also been tried, on this and other similar datasets: the linear model seems to best capture the stimulus-response dynamics without over-fitting the data. These results suggest a common mechanism of contagion in both negative and positive contents: in general, a strongly negative stimulus is followed by negative responses, while a strongly positive stimulus generates positive responses. Neutral stimuli also trigger neutral responses.

Extent of emotional contagion and individuals' susceptibility

Using data collected for the previous experiment, we can also explore if different users have different susceptibility to emotional contagion, for example by measuring how many of their tweets reflect the over-represented stimulus prior to the postings. We now focus our attention on the tweets posted by each of the 3,800 users in our dataset and all tweets produced by their followees.

To determine whether user u was susceptible to emotional contagion prior to posting any of her/his tweets, for each tweet t_u posted by u we calculate the proportions of positive p^+ , neutral p^0 , and negative p^- polarities computed from the distribution of all tweets produced by u 's followees in the 1 hour prior to t_u 's posting time. This triplet $O = \{p^+, p^0, p^-\}$ has three entries that indicate the proportion of each of the three sentiment states $\{+, ^0, -\}$. These tweets are considered as the stimulus to which user u was exposed prior to posting tweet t_u .

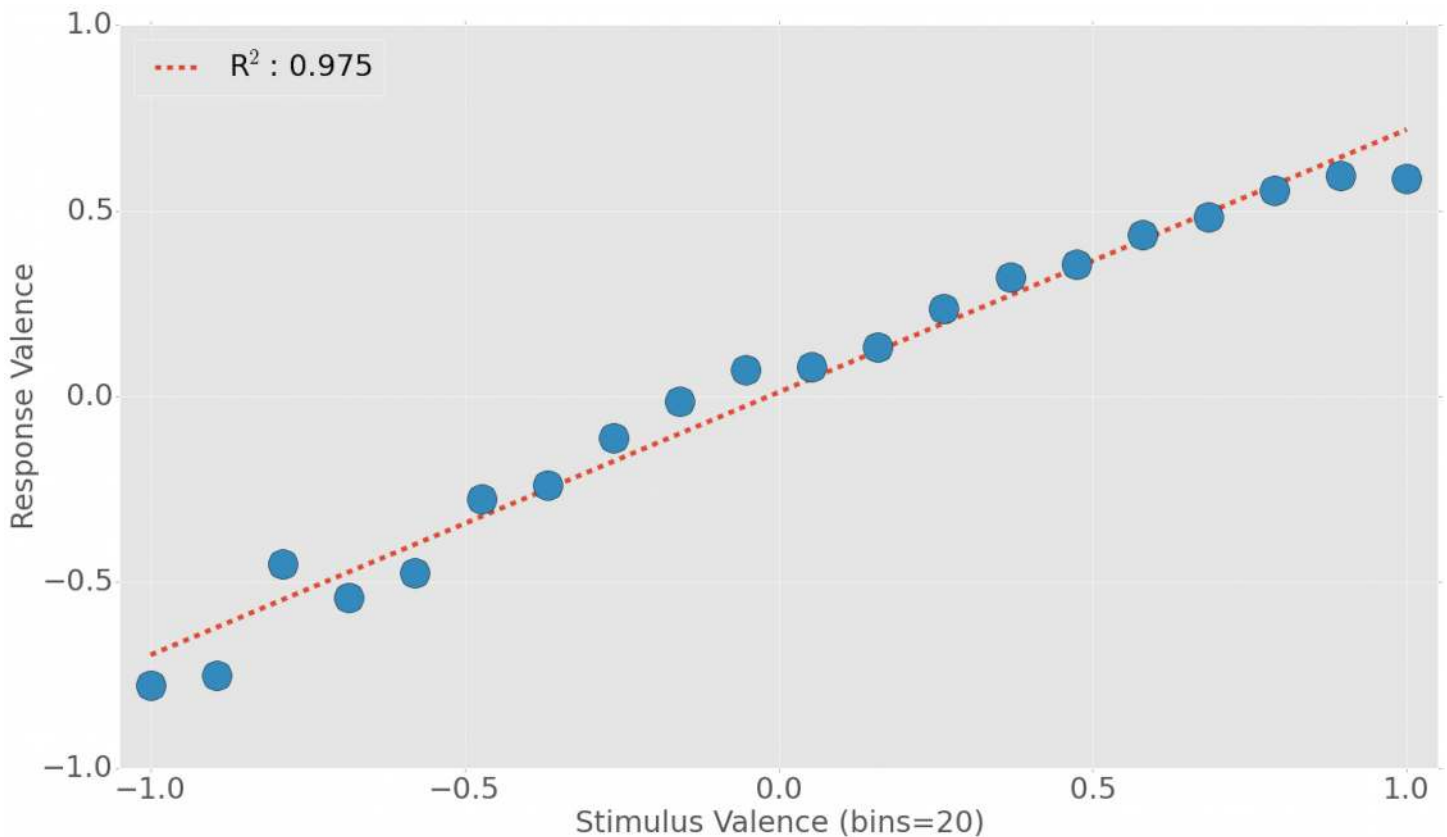


Fig 3. Relationship between stimulus and response valence in Twitter. The emerging linear relationship ($R^2 = 0.975$) suggests that there is a strong correlation between stimuli and responses in terms of valence (difference between positive and negative sentiments in the set of tweets).

doi:10.1371/journal.pone.0142390.g003

The following baseline proportions are derived by the previous experiment (Fig 3):

$$\begin{aligned}
 B^- &= \{21.63, 45.02, 33.35\} \\
 B^\circ &= \{16.49, 48.95, 34.56\} \\
 B^+ &= \{16.00, 45.05, 38.94\}.
 \end{aligned}$$

We therefore determine the smallest Euclidean distance $\bar{L}_2(O, B^s)$ among the distances between the observed distribution O , and any of the three baseline sentiment proportions B^-, B° , and B^+ . This to determine the nature of the stimulus to which u is exposed prior to posting (over-exposure to negative, neutral, or positive):

$$\bar{L}_2(O, B^s) = \min_{s \in \{+, \circ, -\}} \left\{ \sqrt{\sum_{i \in \{+, \circ, -\}} (O_i - B_i^s)^2} \right\} \quad (3)$$

If the smallest distance is, say, $\bar{L}_2(O, B^-)$ it means that, in presence of emotional contagion, u would be expected to post a negative tweet given the over-exposure to negative content. Similarly, if the smallest distance is $\bar{L}_2(O, B^+)$, then u is expected to post a positive tweet, in case $s/$

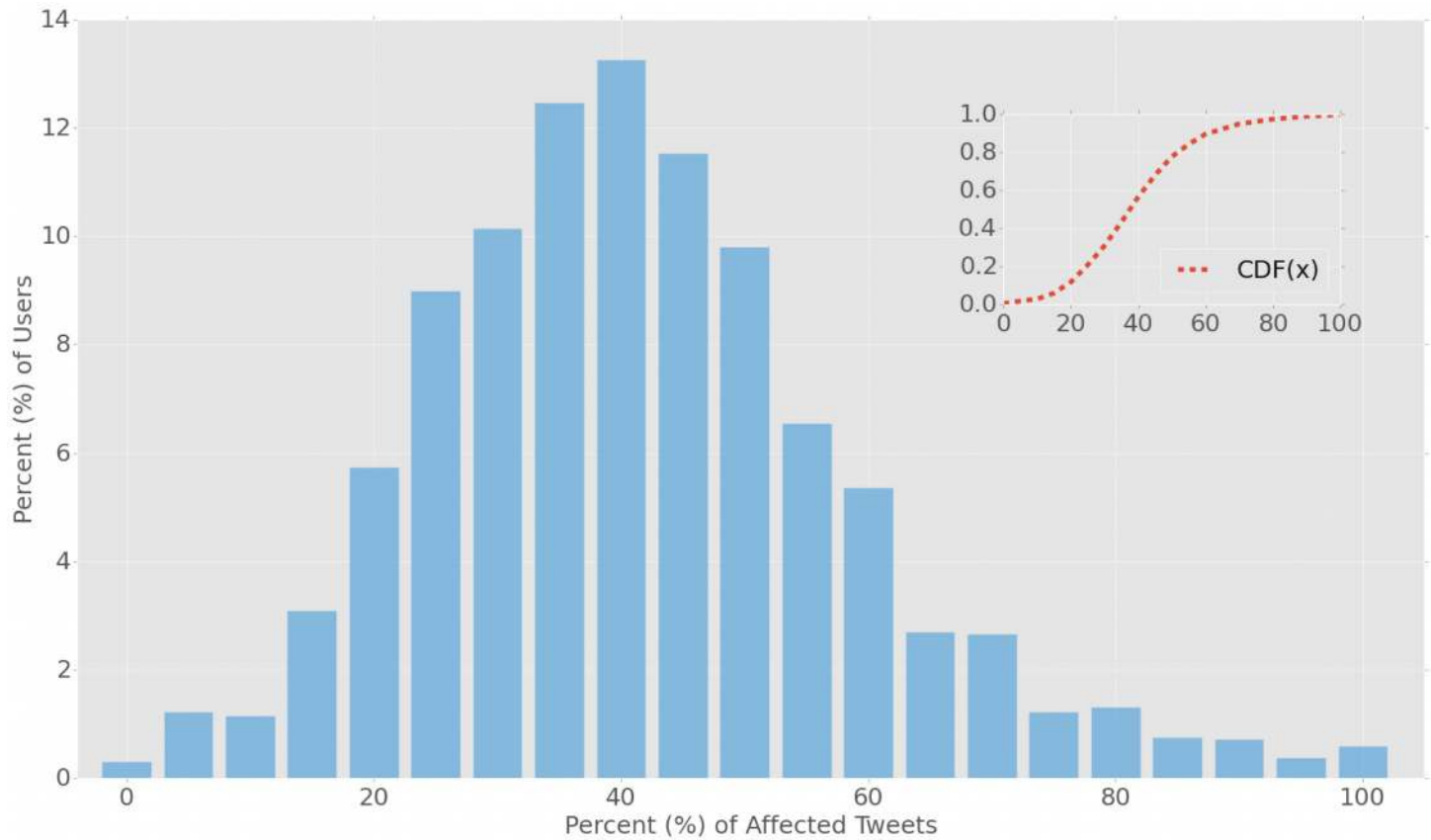


Fig 4. Measurement of emotional contagion on users' content posted on Twitter. The main plot shows the number of users as function of the fraction of their tweets affected by emotional contagion. The inset shows the cumulative distribution. About 80% of the users have up to 50% of their tweets affected by emotional contagion, while the remainder 20% of users exhibits effects of emotional contagion on more than 50% of the posts they produce.

doi:10.1371/journal.pone.0142390.g004

he being affected by emotional contagion. If u were to tweet according to the stimuli (s) he is exposed to, then we consider t_u to be outcome of susceptibility to emotional contagion; vice-versa, t_u is counted as instance of u being insusceptible to emotional contagion given the stimuli.

We perform this analysis for all tweets of all users, and characterize each user u with a fraction summarizing the proportion of tweets affected by emotional contagion. Fig 4 shows the distribution of this measure for all users (the inset of Fig 4 illustrates the cumulative distribution): it is evident that about 80% of the users have up to 50% of their tweets affected by emotional contagion, while the remainder 20% exhibits very high susceptibility and demonstrate that more than 50% of the content they post suggests the presence of emotional contagion.

We further divide the users in two categories, highly and scarcely susceptible to emotional contagion, by selecting the top and bottom 15% of the distribution, respectively. For these two classes independently we compute the fraction of susceptible tweets that are positively or negatively affected by emotional contagion, we average these fractions across users, and we plot the results in Fig 5. We can note that two very different emotional contagion dynamics exist: the group of users who are more susceptible to emotional contagion, are significantly more inclined to adopt positive emotions rather than negative. The vice-versa happens for users scarcely susceptible to emotional contagion: they adopt much more frequently negative

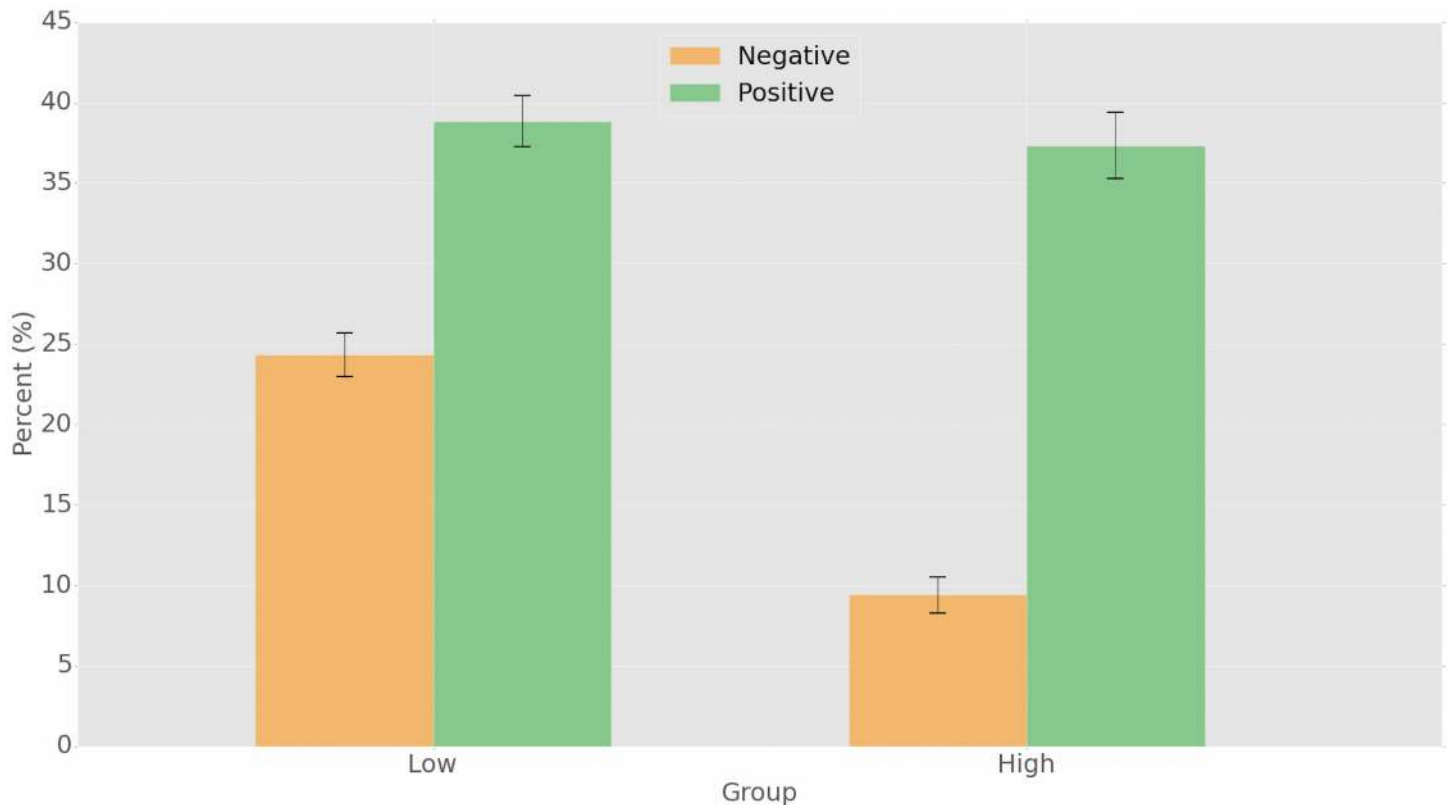


Fig 5. Different extent of emotional contagion on the two groups of scarcely and highly susceptible users. Highly susceptible users are significantly less inclined to adopt negative emotions than the scarcely susceptible ones, but equally likely to adopt positive emotions. In general, the likelihood of adopting positive emotions is much greater than that of negative emotions.

doi:10.1371/journal.pone.0142390.g005

emotions in the uncommon occurrences when they are susceptible to emotional contagion. However, the probability of a contagion of positive emotions is much greater than the negative case in both susceptibility classes: the low- and high-susceptibility groups are, respectively, 1.6 times and 3.96 times more likely to adopt positive emotions, with respect to negative ones.

Discussion

In this study we performed an extensive observational analysis of the patterns of emotional contagion on a sample of Twitter users. Differently from a study carried out on Facebook [29], where controlled experiments were performed to manipulate the exposure to arbitrary emotions, in this study we observe and measure emotional contagion without interacting with the users. The design of a clever null model, which discounts some confounding factors including contagion, allows us to highlight the effect of emotional contagion on a sample of 3,800 users whose activity comprehending the entire history of stimuli they were exposed to and responses they produced has been observed throughout a week during the end of September 2014. Our results suggest a number of insights: we can hypothesize the presence of emotional contagion even without the hassle (and ethical concerns) of manipulating users time-lines. We observed that, on average, on our sample of Twitter users a negative tweet follows an over-exposure to

4.34% more negative stimuli, whereas a positive one follows an over-exposure to 4.50% more positive tweets. A strong linear relation emerges between the valence of the stimuli and that of the responses, suggesting that a common mechanism of contagion exists regulating both negative and positive emotions. Finally, by dividing the users in two categories (highly and scarcely susceptible), we observed how, in general, positive emotions are more prone to contagion, and that highly-susceptible users are significantly more inclined to adopt positive emotions.

Due to the observational nature of our experiment, our study is certainly not immune of possible shortcomings: emotional contagion may not be the only observed phenomenon, but it might co-occur with other network effects. For example, theoretical work by Shalizi and Thomas [48] suggests that in observational studies like ours it is not possible to separate contagion from homophily. In a world entirely dominated by homophily, our observation would not imply an effect: users prone to produce negative contents would link only to others with same emotional alignment (and vice-versa for positive-inclined ones).

However, in the real world it makes sense to assume a mixture of contagion and homophily dynamics, as illustrated by some recent studies [49–52]. In such a scenario, our observations suggest the presence and the extent of emotional contagion, while further work will be needed to understand the effect of homophily and how that intertwines with emotional contagion. To this end, we suggest the possibility of designing an in-silico experiment in which homophily is arbitrarily tuned by artificially affecting the social network structure, for example by introducing an ad-hoc community structure [53–55], and contagion is thus analyzed in light of the controlled homophily mechanism.

Another interesting aspect worth mentioning is that, in absence of non-verbal cues, disentangling some nuances of the dynamics behind emotions and contagion may be challenging: for example, it is quite hard to tell apart contagion from empathy; in some instances, users might also pretend to sympathize (without necessarily doing so) by aligning their expressions with others' contents; yet, parts of emotional contents may be picked up within conversations as responses but without actually being indications of one's emotions. Interestingly, these limitations are not due to the observational nature of the study—the above challenges would hold true also for the most carefully designed controlled experiments—and we point the interested readers to social and cognitive psychology literature for further investigations of such phenomena [56–59].

Other fundamental limits arise from the current state of the art in sentiment analysis algorithms: modern approaches, like SentiStrength here employed, although more robust and precise than ever before, still produce crude heuristics and hardly capture the many nuanced expressions that human language is able to convey. The inability to capture complex contexts triggering expressions like sarcasm or irony, the attribution of equal weights to all emotions, the suppression of multiple emotions, the presence of ambiguity (i.e., tweets that include at the same time positive and negative emotions), etc., are only few examples of the sources of potential noisy outputs of such methods. It is however worth noting that such limits apply to all studies that make usage of sentiment analysis tools, and don't inherently affect the validity of the presented findings.

Concluding, our study relies on an idealized world in which each user reads all contents (stimuli tweets) he/she is exposed to during one hour prior to the production of one of his/her own tweets. Certainly, this oftentimes might not correspond to the reality: recent studies explored the effects of limited cognitive capacity on social media users, unveiling that memory and limited attention play a crucial role in the dynamics of information production and consumption [60–62]. In the future, it would be interesting to perform controlled experiments in which these dynamics are intermingled with contagion effects.

Acknowledgments

EF is grateful to Filippo Menczer, Y.Y. Ahn, Sune Lehmann, and Johan Bollen for interesting discussions, and to Alessandro Flammini and Lorenzo Coviello for their precious feedback on the project and extensive comments on the manuscript.

Author Contributions

Conceived and designed the experiments: EF. Performed the experiments: EF ZY. Analyzed the data: EF ZY. Contributed reagents/materials/analysis tools: EF ZY. Wrote the paper: EF.

References

1. Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, et al. Life in the network: the coming age of computational social science. *Science*. 2009; 323(5915):721.
2. Vespignani A. Predicting the behavior of techno-social systems. *Science*. 2009; 325(5939):425. doi: [10.1126/science.1171990](https://doi.org/10.1126/science.1171990) PMID: [19628859](https://pubmed.ncbi.nlm.nih.gov/19628859/)
3. Gilbert E, Karahalios K. Predicting tie strength with social media. In: 27th SIGCHI Conference on Human Factors in Computing Systems. ACM; 2009. p. 211–220.
4. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*. 2010; 53(1):59–68. doi: [10.1016/j.bushor.2009.09.003](https://doi.org/10.1016/j.bushor.2009.09.003)
5. Asur S, Huberman BA. Predicting the future with social media. In: 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. IEEE; 2010. p. 492–499.
6. Tang J, Lou T, Kleinberg J. Inferring social ties across heterogenous networks. In: Proceedings of the fifth ACM international conference on Web search and data mining. ACM; 2012. p. 743–752.
7. Cheng J, Adamic L, Dow PA, Kleinberg JM, Leskovec J. Can cascades be predicted? In: Proceedings of the 23rd international conference on World wide web. ACM; 2014. p. 925–936.
8. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In: Proceedings of the 19th international conference on World wide web. ACM; 2010. p. 591–600.
9. Gomez Rodriguez M, Leskovec J, Schölkopf B. Structure and dynamics of information pathways in online media. In: Proceedings of the sixth ACM international conference on Web search and data mining. ACM; 2013. p. 23–32.
10. Ferrara E, Varol O, Menczer F, Flammini A. Traveling trends: social butterflies or frequent fliers? In: First ACM conference on Online social networks. ACM; 2013. p. 213–222.
11. Ratkiewicz J, Conover M, Meiss M, Gonçalves B, Flammini A, Menczer F. Detecting and Tracking Political Abuse in Social Media. In: 5th International AAAI Conference on Weblogs and Social Media; 2011. p. 297–304.
12. Metaxas PT, Mustafaraj E. Social media and the elections. *Science*. 2012; 338(6106):472–473. doi: [10.1126/science.1230456](https://doi.org/10.1126/science.1230456) PMID: [23112315](https://pubmed.ncbi.nlm.nih.gov/23112315/)
13. Bond RM, Fariss CJ, Jones JJ, Kramer AD, Marlow C, Settle JE, et al. A 61-million-person experiment in social influence and political mobilization. *Nature*. 2012; 489(7415):295–298. doi: [10.1038/nature11421](https://doi.org/10.1038/nature11421) PMID: [22972300](https://pubmed.ncbi.nlm.nih.gov/22972300/)
14. Conover MD, Ferrara E, Menczer F, Flammini A. The digital evolution of Occupy Wall Street. *PloS ONE*. 2013; 8(5):e64679. doi: [10.1371/journal.pone.0064679](https://doi.org/10.1371/journal.pone.0064679) PMID: [23734215](https://pubmed.ncbi.nlm.nih.gov/23734215/)
15. Conover MD, Davis C, Ferrara E, McKelvey K, Menczer F, Flammini A. The geospatial characteristics of a social movement communication network. *PloS ONE*. 2013; 8(3):e55957. doi: [10.1371/journal.pone.0055957](https://doi.org/10.1371/journal.pone.0055957) PMID: [23483885](https://pubmed.ncbi.nlm.nih.gov/23483885/)
16. Varol O, Ferrara E, Ogan CL, Menczer F, Flammini A. Evolution of online user behavior during a social upheaval. In: 2014 ACM conference on Web Science. ACM; 2014. p. 81–90.
17. Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users: real-time event detection by social sensors. In: 19th International Conference on World Wide Web. ACM; 2010. p. 851–860.
18. Merchant RM, Elmer S, Lurie N. Integrating social media into emergency-preparedness efforts. *New England Journal of Medicine*. 2011; 365(4):289–291. doi: [10.1056/NEJMp1103591](https://doi.org/10.1056/NEJMp1103591) PMID: [21793742](https://pubmed.ncbi.nlm.nih.gov/21793742/)
19. Lazer D, Kennedy R, King G, Vespignani A. The Parable of Google Flu: Traps in Big Data Analysis. *Science*. 2014; 343(6176):1203–1205. doi: [10.1126/science.1248506](https://doi.org/10.1126/science.1248506) PMID: [24626916](https://pubmed.ncbi.nlm.nih.gov/24626916/)
20. Hatfield E, Cacioppo JT. Emotional contagion. Cambridge Univ. Press; 1994.

21. Fowler JH, Christakis NA, et al. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *BMJ*. 2008; 337:a2338. doi: [10.1136/bmj.a2338](https://doi.org/10.1136/bmj.a2338) PMID: [19056788](https://pubmed.ncbi.nlm.nih.gov/19056788/)
22. Harris RB, Paradice D. An investigation of the computer-mediated communication of emotions. *Journal of Applied Sciences Research*. 2007; 3(12):2081–2090.
23. Mei Q, Ling X, Wondra M, Su H, Zhai C. Topic sentiment mixture: modeling facets and opinions in weblogs. In: *Proceedings of the 16th international conference on World Wide Web*. ACM; 2007. p. 171–180.
24. Golder SA, Macy MW. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*. 2011; 333(6051):1878–1881. doi: [10.1126/science.1202775](https://doi.org/10.1126/science.1202775) PMID: [21960633](https://pubmed.ncbi.nlm.nih.gov/21960633/)
25. Choudhury MD, Counts S, Gamon M. Not All Moods Are Created Equal! Exploring Human Emotional States in Social Media. In: *International AAAI Conference on Weblogs and Social Media*; 2012. p. 66–73.
26. Garcia D, Garas A, Schweitzer F. Positive words carry less information than negative words. *EPJ Data Science*. 2012; 1(1):3. doi: [10.1140/epjds3](https://doi.org/10.1140/epjds3)
27. Coviello L, Sohn Y, Kramer AD, Marlow C, Franceschetti M, Christakis NA, et al. Detecting emotional contagion in massive social networks. *PloS one*. 2014; 9(3):e90315. doi: [10.1371/journal.pone.0090315](https://doi.org/10.1371/journal.pone.0090315) PMID: [24621792](https://pubmed.ncbi.nlm.nih.gov/24621792/)
28. Coviello L, Franceschetti M. Words on the Web: Noninvasive Detection of Emotional Contagion in Online Social Networks. *Proceedings of the IEEE*. 2014; 102(12). doi: [10.1109/JPROC.2014.2366052](https://doi.org/10.1109/JPROC.2014.2366052)
29. Kramer AD, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014;p. 201320040.
30. Tufekci Z. Engineering the public: Big data, surveillance and computational politics. *First Monday*. 2014; 19(7). doi: [10.5210/fm.v19i7.4901](https://doi.org/10.5210/fm.v19i7.4901)
31. Fiske ST, Hauser RM. Protecting human research participants in the age of big data. *Proceedings of the National Academy of Sciences*. 2014; 111(38):13675–13676. doi: [10.1073/pnas.1414626111](https://doi.org/10.1073/pnas.1414626111)
32. Acquisti A, Taylor C, Wagman L. The economics of privacy. *Journal of Economic Literature*. 2014;.
33. Ferrara E. A large-scale community structure analysis in Facebook. *EPJ Data Science*. 2012; 1(1):1–30. doi: [10.1140/epjds9](https://doi.org/10.1140/epjds9)
34. De Meo P, Ferrara E, Fiumara G, Provetti A. On Facebook, most ties are weak. *Communications of the ACM*. 2014; 57(11):78–84. doi: [10.1145/2629438](https://doi.org/10.1145/2629438)
35. Backstrom L, Kleinberg J. Romantic partnerships and the dispersion of social ties: a network analysis of relationship status on facebook. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM; 2014. p. 831–841.
36. Pang B, Lee L. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*. 2008; 2(1–2):1–135. doi: [10.1561/1500000011](https://doi.org/10.1561/1500000011)
37. Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. *Journal of Computational Science*. 2011; 2(1):1–8. doi: [10.1016/j.jocs.2010.12.007](https://doi.org/10.1016/j.jocs.2010.12.007)
38. Bollen J, Mao H, Pepe A. Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena. In: *International AAAI Conference on Weblogs and Social Media*. AAAI; 2011. p. 450–453.
39. Le L, Ferrara E, Flammini A. On predictability of rare events leveraging social media: a machine learning perspective. In: *COSN'15: 2015 ACM SGB Conference on Online Social Networks*. ACM; 2015.
40. Ferrara E, Yang Z. Quantifying the Effect of Sentiment on Information Diffusion in Social Media. *PeerJ Computer Science*. 2015;p. 1e26.
41. Akkaya C, Wiebe J, Mihalcea R. Subjectivity word sense disambiguation. In: *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*. ACL; 2009. p. 190–199.
42. Paltoglou G, Thelwall M. A study of information retrieval weighting schemes for sentiment analysis. In: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. ACL; 2010. p. 1386–1395.
43. Hutto C, Gilbert E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In: *International AAAI Conference on Weblogs and Social Media*. AAAI; 2014. p. 216–225.
44. Thelwall M, Buckley K, Paltoglou G, Cai D, Kappas A. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*. 2010; 61(12):2544–2558. doi: [10.1002/asi.21416](https://doi.org/10.1002/asi.21416)
45. Thelwall M, Buckley K, Paltoglou G. Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*. 2011; 62(2):406–418. doi: [10.1002/asi.21462](https://doi.org/10.1002/asi.21462)

46. Stieglitz S, Dang-Xuan L. Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*. 2013; 29(4):217–248. doi: [10.2753/MIS0742-1222290408](https://doi.org/10.2753/MIS0742-1222290408)
47. Morstatter F, Pfeffer J, Liu H, Carley KM. Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose. In: *Seventh International AAAI Conference on Weblogs and Social Media*; 2013.
48. Shalizi CR, Thomas AC. Homophily and contagion are generically confounded in observational social network studies. *Sociological methods & research*. 2011; 40(2):211–239. doi: [10.1177/0049124111404820](https://doi.org/10.1177/0049124111404820)
49. Aral S, Muchnik L, Sundararajan A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*. 2009; 106(51):21544–21549. doi: [10.1073/pnas.0908800106](https://doi.org/10.1073/pnas.0908800106)
50. VanderWeele TJ. Sensitivity analysis for contagion effects in social networks. *Sociological Methods & Research*. 2011; 40(2):240–255. doi: [10.1177/0049124111404821](https://doi.org/10.1177/0049124111404821)
51. Bakshy E, Rosenn I, Marlow C, Adamic L. The role of social networks in information diffusion. In: *Proceedings of the 21st international conference on World Wide Web*. ACM; 2012. p. 519–528.
52. Lewis K, Gonzalez M, Kaufman J. Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*. 2012; 109(1):68–72. doi: [10.1073/pnas.1109739109](https://doi.org/10.1073/pnas.1109739109)
53. Centola D. The spread of behavior in an online social network experiment. *Science*. 2010; 329(5996):1194–1197. doi: [10.1126/science.1185231](https://doi.org/10.1126/science.1185231) PMID: [20813952](https://pubmed.ncbi.nlm.nih.gov/20813952/)
54. Nematzadeh A, Ferrara E, Flammini A, Ahn YY. Optimal network modularity for information diffusion. *Physical review letters*. 2014; 113(8):088701. doi: [10.1103/PhysRevLett.113.088701](https://doi.org/10.1103/PhysRevLett.113.088701) PMID: [25192129](https://pubmed.ncbi.nlm.nih.gov/25192129/)
55. Centola D, Baronchelli A. The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*. 2015; 112(7):1989–1994. doi: [10.1073/pnas.1418838112](https://doi.org/10.1073/pnas.1418838112)
56. Hatfield E, Rapson RL, Le YCL. Emotional Contagion and Empathy. In: *The social neuroscience of empathy*. MIT Press; 2011. p. 19–30.
57. Dimberg U, Thunberg M. Empathy, emotional contagion, and rapid facial reactions to angry and happy facial expressions. *Psych Journal*. 2012; 1(2):118–127. doi: [10.1002/pchj.4](https://doi.org/10.1002/pchj.4) PMID: [26272762](https://pubmed.ncbi.nlm.nih.gov/26272762/)
58. Tsai J, Bowring E, Marsella S, Wood W, Tambe M. A Study of Emotional Contagion with Virtual Characters. In: *Intelligent Virtual Agents*. vol. 7502 of *Lecture Notes in Computer Science*. Springer; 2012. p. 81–88.
59. Mackie DM, Maitner AT, Smith ER. Emotion and Intergroup Relations. In: *Emerging Trends in the Social and Behavioral Sciences*. John Wiley & Sons, Inc.; 2015.
60. Hodas NO, Lerman K. How visibility and divided attention constrain social contagion. In: *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Conference on Social Computing (SocialCom)*. IEEE; 2012. p. 249–257.
61. Kang JH, Lerman K. Structural and cognitive bottlenecks to information access in social networks. In: *Proceedings of the 24th ACM Conference on Hypertext and Social Media*. ACM; 2013. p. 51–59.
62. Hodas NO, Lerman K. The simple rules of social contagion. *Scientific reports*. 2014; 4. doi: [10.1038/srep04343](https://doi.org/10.1038/srep04343) PMID: [24614301](https://pubmed.ncbi.nlm.nih.gov/24614301/)