

RESEARCH ARTICLE

Emoticon-Based Ambivalent Expression: A Hidden Indicator for Unusual Behaviors in Weibo

Yue Hu, Jichang Zhao, Junjie Wu*

School of Economics and Management, Beihang University, Beijing, China

* wujj@buaa.edu.cn



OPEN ACCESS

Citation: Hu Y, Zhao J, Wu J (2016) Emoticon-Based Ambivalent Expression: A Hidden Indicator for Unusual Behaviors in Weibo. PLoS ONE 11(1): e0147079. doi:10.1371/journal.pone.0147079

Editor: Zi-Ke Zhang, Hangzhou Normal University, CHINA

Received: July 22, 2015

Accepted: December 27, 2015

Published: January 22, 2016

Copyright: © 2016 Hu et al. 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: All relevant data are uploaded to figshare under DOI <https://dx.doi.org/10.6084/m9.figshare.2060046>.

Funding: This work was partially supported by National Natural Science Foundation of China (71322104, 71171007, 71471009 and 71531001), National Center for International Joint Research on E-Business Information Processing (2013B01035), National High Technology Research and Development Program of China (863 Program) (SS2014AA012303), Foundation for the Author of National Excellent Doctoral Dissertation of PR China (201189), and Program for New Century Excellent Talents in University (NCET-11-0778). JZ thanks NSF

Abstract

Recent decades have witnessed online social media being a big-data window for testifying conventional social theories quantitatively and exploring much detailed human behavioral patterns. In this paper, by tracing the emoticon use in Weibo, a group of hidden “ambivalent users” are disclosed for frequently posting ambivalent tweets containing both positive and negative emotions. Further investigation reveals that this ambivalent expression could be a novel indicator of many unusual social behaviors. For instance, ambivalent users with the female as the majority like to make a sound in midnights and at weekends. They mention their close friends frequently in ambivalent tweets, which attract more replies and serve as a more private communication way. Ambivalent users also respond differently to public affairs from others and demonstrate more interests in entertainment and sports events. Moreover, the sentiment shift in ambivalent tweets is more evident than usual and exhibits a clear “negative to positive” pattern. The above observations, though being promiscuous seemingly, actually point to the self-regulation of negative mood in Weibo, which could find its basis from the traditional emotion management theories in sociology but makes an important extension to the online environment in this study. Finally, as an interesting corollary, ambivalent users are found connected with compulsive buyers and turn out to be perfect targets for online marketing.

Introduction

Emotional expression is a psychological behavior with the purpose of communicating affective states between different individuals. This behavior could be either verbal or nonverbal, including texts, voice, face expressions and body languages [1]. In the epoch of the Internet, tremendous developments of online social media provide abundant innovative and powerful means of information exchange, bringing unprecedented richness and diversity to the forms of emotional expression.

Among these new forms, the *emoticon* is becoming more and more popular with the rapid growth of Weibo, a Twitter-like service in China. In general, an emoticon is a small image used to convey extra meaning in a statement outside of written texts and derives its name from a

of China (71501005). 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.

merging of the words “emotion” and “icon”. This differs from a *smiley* which is formed by combining punctuation marks as used in Twitter. Indeed, users seem to show increasing interests in expressing sentiments through approximately 2,000 emoticons designed by Weibo, ranging from facial expressions like laughing or crying, body languages like applause or hug, to some icons like the sun, a Christmas tree or birthday cake. Utilization of emoticons in posts is very common since posts to Weibo are very short and lack context due to the 140-character limitation. Emoticons allow posts to convey emotions and context in a vivid way without wasting precious character slots.

It has been found that both smileys and emoticons are strongly related with typical sentiment words, and could serve as convincing indicators of different emotions [2]. Tossell et al. confirm that emoticon usage is contextual, and people generally use more negative emoticons than positive ones [3]. Researchers also realize that emoticons and smileys could be effective features for texts to improve the precision of sentiment analysis [4–9], or be treated as sentiment labels to avoid intensive labor costs for preparing training data [5].

As a special social behavior, emoticon usage is shaped by cultural and social factors. The pattern of emoticon usage in short-message texting is investigated and the discrepancies between male and female users are revealed [10]. For instance, female users send more messages containing emoticons, while male users use a wider range of emoticons. Schnoebelen et al. distinguish users by whether they use nose smiley “:-)” or non-nose smiley “:)”, and demonstrate that the variants correspond to different types of users, tweeting with different vocabularies and writing styles [11]. Park et al. point out that emoticons are socio-cultural norms and their meanings could be affected by the identity of the speaker [12].

The above research, though being very interesting, is mainly focused on emoticons rather than the emoticon users. Who they are, what they are talking about, how they behave online and why they adopt emoticons—these human-centric problems are actually very interesting to sociologists and perhaps marketers, who are always seeking for potential consumers. Meanwhile, different from expensive but spatially or temporally limited surveys in tradition [13], digital traces created by interactions with technology indeed offer a new probe to collective human behaviors [14–16], and these new data are fueling the rapid development of computational social science [17, 18]. These indeed motivate our study in this paper, which takes ambivalent emoticons in a tweet as latent clue for tracing the unusual behaviors of a special group called “ambivalent users”.

More specifically, we manually select 79 positive and 36 negative emoticons with unambiguous polarities to label the sentiments of tweets; that is, tweets containing only positive (negative) emoticons are defined as positive (negative) tweets and vice versa. Surprisingly, we find many *ambivalent tweets* containing both positive and negative emoticons, indicating inconsistent emotions [19]. The *ambivalent users* are then defined as the users who have published more than 30 ambivalent tweets in 2012. The *senior ambivalent users* are defined accordingly as a group of special ambivalent users who have published more than 50 ambivalent tweets in 2012. As contrast, the *ordinary users* are the ones who have not published any ambivalent tweet in 2012. After excluding abnormal users like verified celebrities, organizations, inactive users (who post less than 12 tweets in 2012 or have less than 65 followers) and spam accounts, we are left with 1,069 ambivalent users, among whom 357 are senior users, and the ordinary users total 46,245. Note that the threshold used to identify ambivalent users (i.e., the least number of ambivalent tweets posted, 30) will not significantly affect our following findings, but setting it as high as 30 or even larger can help to find more representative ambivalent users who are markedly different to ordinary users.

A first study of demographics shows that ambivalent expression is more popular in Weibo than in Twitter, and that the female makes up a significantly higher proportion of ambivalent

users. This indeed validates the conventional theory in psychology that people from the east show significantly stronger association of positive and negative in emotion expression than the west, especially the female [20]. Further explorations on the behaviors of ambivalent users disclose more interesting patterns. Specifically, ambivalent users like to express ambivalence in midnights or at weekends, focus much more on topics of entertainment and sports, actively mention their close friends in tweets and gain more replies rather than reposts than ordinary users, and frequently use positive terms while expressing negative mood.

The above observations suggest the subtle link between users' mood status and the usage of online social media, which has become a very hot topic in recent years and attracted increasing research interests [21–28]. Specifically, based on the theories of emotion management in sociology [20, 29, 30], we point out that ambivalent users' unusual behaviors can be well explained from the view of self-regulation against negative feelings. Moreover, this self-regulation behavior of ambivalent users seems to be conscious, which makes our findings distinct from existing studies mostly on unconscious behaviors. Finally, as a natural conjecture, we testify the self-correcting effect of shopping on ambivalent users' negative mood using the tweets posted around the Singles' Day in 2012. This indicates that ambivalent users like compulsive users are ideal targets for online marketing.

Results

Ambivalent Expression

We first observe ambivalent expression on Weibo. It is interesting that nearly 1.9% of the emotional tweets on Weibo are ambivalent, but the figure reduces to 1.1% for Twitter, indicating stronger inclination of Chinese Weibo users for ambivalent expression (see Fig 1 for typical emoticons in Weibo and more details in Methods). This may owe to the added clarity and ease-of-use of Weibo's emoticon feature. Linguistic cultural differences between Chinese and English may also contribute to these discrepancies.

We then focus on the ambivalent users on Weibo. It is interesting that ambivalent users generally show more passive mood by publishing significantly more negative tweets than ordinary users (the negative-tweet ratios are 27.1% against 18.8% for ambivalent and ordinary users, respectively, in 2012). Moreover, the female occupies 79.0% of total ambivalent users but only 42.8% of the ordinary users, which indicates that the female dominates the ambivalent users. Among them we find a verified community named *Weibo Lady*, whose members are most active females passionate about sharing their personal lives. In this community, 82.8% of the users ever posted ambivalent tweets, which is consistent with the theory that people from the east show a stronger association between positive and negative affect than the west, especially the female [20].

We further observe the posting behavior of ambivalent users. Fig 2(a) shows the daily pattern. As can be seen, compared with the normal ones, ambivalent tweets seem more noticeable in the midnight from 11PM to 2AM. Similar situation happens for the weekly pattern in Fig 2(b), where significantly more ambivalent tweets are posted at the weekend, especially on Sunday. Since midnights and weekends are usually the leisure time for individuals, we could conjecture that ambivalent users are apt to express mixed feelings when relaxed.

Topic Preference

We further explore the topic preference of ambivalent users. It has been reported that the variation of social behavioral indexes, like "the number of posts", could be employed to detect hot topics or external events in social media [5, 31]. Accordingly, we trace the time series of positive





Positive		Negative	
Emoticon	Description	Emoticon	Description
	Hug		Aggrieved
	Gift		Dizzy
	Applaud		Snort
	Steal happiness		Sick
	Like		Disappointed
	Heart		Rant
	Smirk		Look down upon
	Cool		Cry
	Rose		Bad luck
	Mighty		Tired

Fig 1. Typical positive and negative emoticons in Weibo. The tweets containing both the positive and negative emoticons will be identified as ambivalent ones.

doi:10.1371/journal.pone.0147079.g001

and negative tweets posted by ambivalent users, and highlight the peaks with clear event semantics (see [Methods](#)), as shown in [Fig 3](#).

[Fig 3\(a\)](#) suggests that, in terms of positive emotion, besides the festivals such as Spring Festival, Valentine’s Day, and Moon Cake Festival, ambivalent users enjoy talking more about some niche topics of sports and entertainment, like Day to speak out love (*J*), a singer Xin Liu’s Birthday (*L*), European Cup competitions (*L* and *M*), and Olympic Games (*N*). The variations of negative sentiment in [Fig 3\(b\)](#) further validate this point; that is, besides the natural disasters with universal concerns, ambivalent users pay more attention to topics like the rescue of dogs (*B*), the competitions in European Cup (*D*), the final episode of a TV play (*F*) and the final of Voice of China (*H*), while contrarily these topics might be very trivial (like the insignificant peaks *A*, *B* and *D*) to ordinary users. These results unveil that ambivalent users are indeed more sensitive and emotional than the ordinary ones. The extracted topic words in [Tables 1](#) and [2](#) provide more details (see [Methods](#)).

Social Interaction

It is interesting that ambivalent tweet is an important channel for ambivalent users to socialize online. To illustrate, we take senior ambivalent users as well as ordinary users for comparative

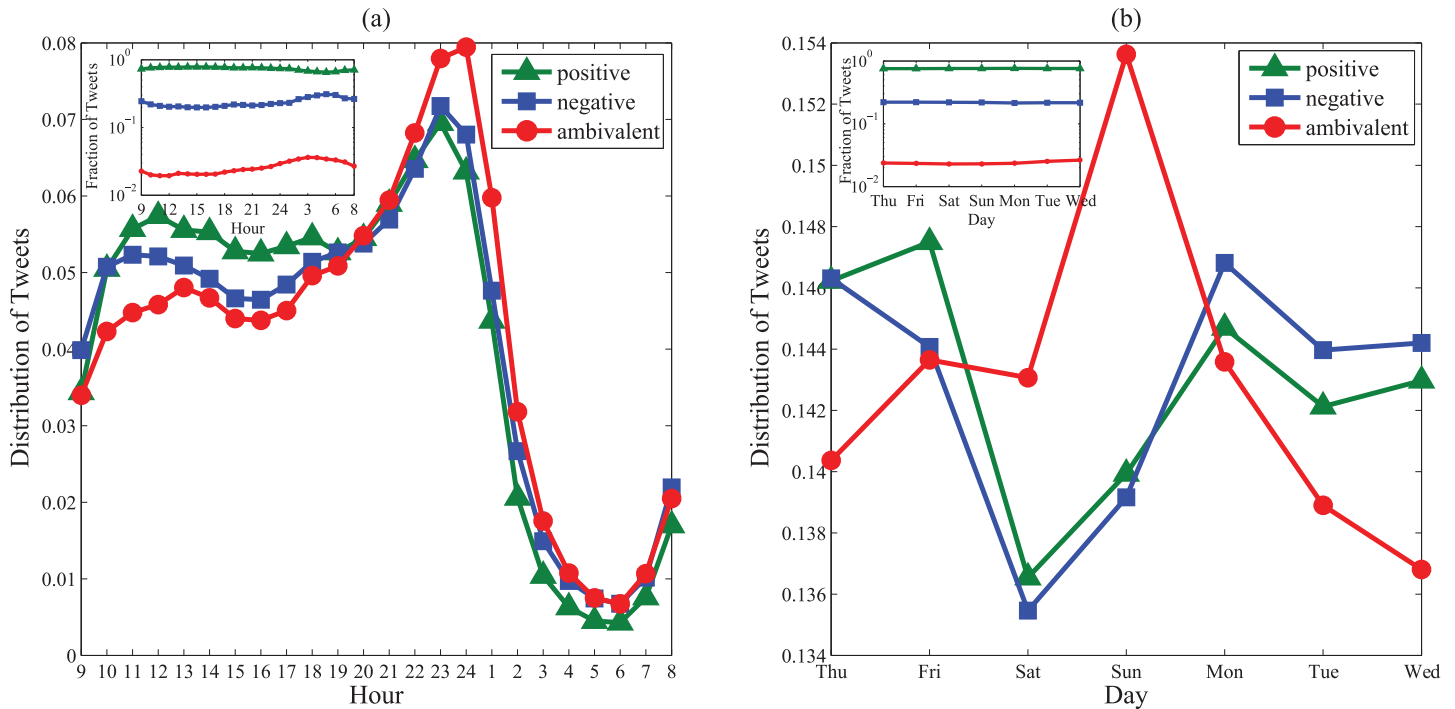


Fig 2. Distribution of tweets posted at different time. (a) The hourly pattern. (b) The weakly pattern. In both (a) and (b), insets show the absolute fraction of each type of tweets at different time, from which it can be seen that the ambivalent tweets only occupy a small fraction in Weibo.

doi:10.1371/journal.pone.0147079.g002

study. Simple statistics show that ambivalent users are a group of people who are socially more active online, e.g., having more fans (2506 versus 735) and bi-friends (270 versus 194) and posting more tweets (1890 versus 336), than the ordinary users. Note that we here define bi-friends as users following each other mutually in Weibo.

More interestingly, ambivalent users have a subtle communication mode. To illustrate this, we trace the “@” behavior in all the tweets they posted in 2012 (a Weibo user can use the “@” symbol in her tweet to inform someone deliberately, which could be regarded as a way for more “private” communication on Weibo). Fig 4 shows that ambivalent users are generally more apt to mention a friend in an ambivalent tweet. Further exploration reveals that the friends mentioned in ambivalent tweets share averagely 20 common followers with the ambivalent users, but the number reduces to 12 for the friends mentioned in positive tweets and 7 in negative tweets. Since to share more friends indicates a stronger social tie between two users [32], the above implies that ambivalent users prefer to interact with their closer friends via ambivalent tweets. In other words, we can track close ties on Weibo by simply checking who is mentioned by whom in ambivalent tweets.

The reactions from the friends of ambivalent users are also very interesting. To illustrate this, we observe the *replies* and *reposts* of the tweets posted by senior ambivalent users in 2012. We find that ambivalent tweets are generally more apt to attract *replies*—a tweet containing dialectic emotions has 36 replies on average, whereas a positive (negative) tweet only has 34 (23) replies. This contradicts with the *reposts*’ case, where a positive tweet elicits more reposts than an ambivalent tweet (63 versus 53). Since a reply is a more private feedback than a repost, the above implies that ambivalent users are easier to gain private feedbacks via ambivalent tweets. In a nutshell, we can conclude that ambivalent users on Weibo depend heavily on ambivalent tweets to interact with their close friends in a relatively private way.

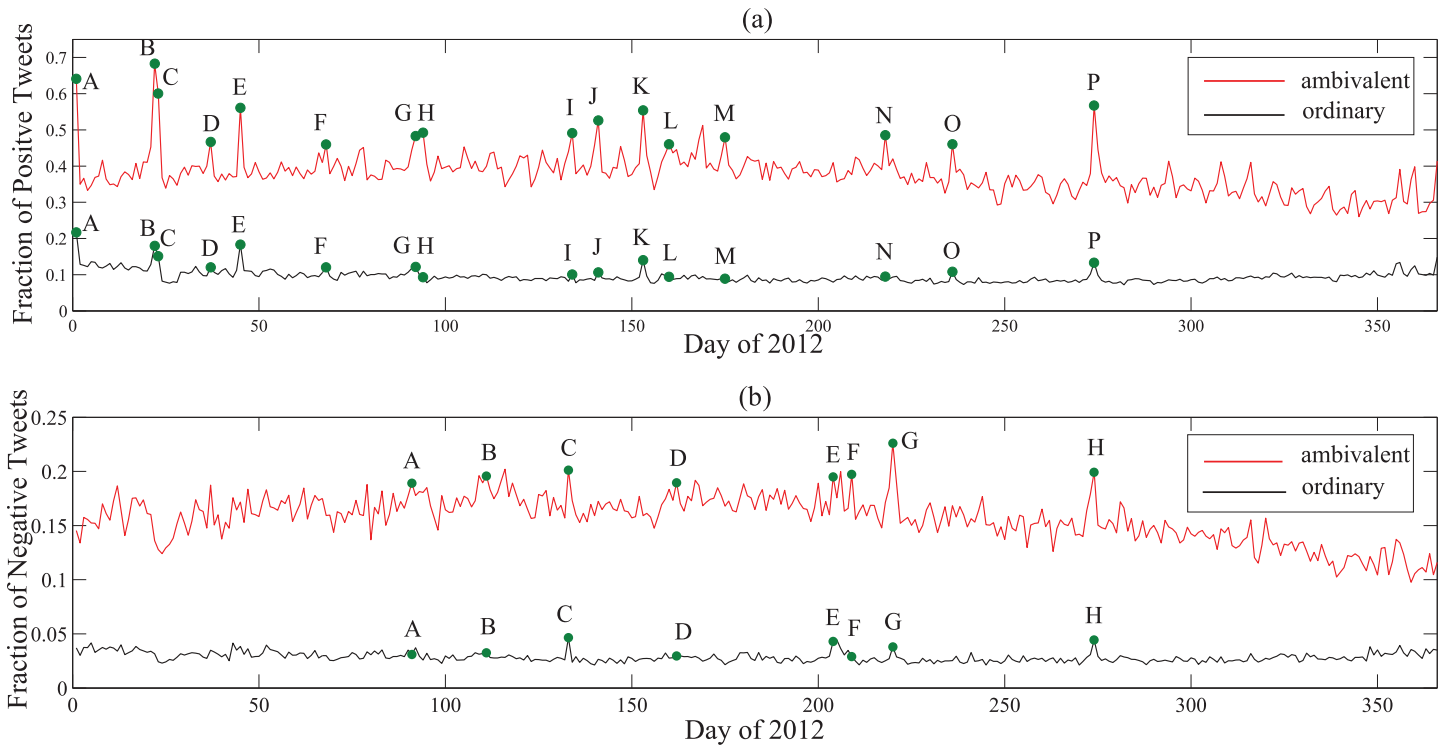


Fig 3. Comparison of topic preferences between ambivalent and ordinary users. (a) Variation of positive emotion for topic detection. (b) Variation of negative emotion for topic detection. Peaks with clear event semantics are selected from the positive or negative emotional line of the ambivalent users and the points with the same time stamp are also highlighted for ordinary users.

doi:10.1371/journal.pone.0147079.g003

Sentiment Shift

We here demonstrate that posting ambivalent tweets can lead to sentiment shift and thus help to ease negative emotion. This is particularly important for ambivalent users, who seem generally more passive than ordinary users. This may also uncover the root of ambivalent expression on social media.

To this end, we collect all the ambivalent tweets posted in 2012 and analyze the sentiment shift indicated by ambivalent emoticons within a tweet. Simple statistics reveal that there are 60,293 tweets shifting from negative emotion to positive one, but only 39,457 in reverse (we abandon tweets with ambiguous sentiment shift). The “negative → positive” pattern of emotional shift implies that Weibo indeed can ease negative mental state of ambivalent users.

We also try “perceiving” subtle emotional shifts from the words adopted in ambivalent tweets. To this end, we collect all the ambivalent tweets posted in 2012 and split them into short clauses, which are then tagged as positive (negative) clauses if they only contain positive (negative) emoticons. The frequent terms are then extracted from the positive (negative) clauses, and are compared with the terms used in purely positive (negative) tweets (see [Methods](#)). It is interesting that the sentiment shift in ambivalent tweets in the term level is surprisingly evident. Specifically, the ambivalent users tend to use some special terms in negative (positive) clauses that are actually more frequently occurred in positive (negative) tweets. This contradiction is particularly evident for the case of negative clauses, implying the more positive inclination of ambivalent tweets. This agrees with the mainstream pattern of emotion shift: negative → positive, and implies that posting ambivalent tweets indeed can help Weibo users realize self-correcting. We will have a more detailed discussion below.

Table 1. Positive Topics Discussed by Ambivalent Users.

A: Jan 1, 2012, New Year
New Year, happy, one year, the first day, hope, ha-ha, 2011, happiness, together
B: Jan 22, 2012, Eve of lunar New Year
New year, happy, Spring Festival Gala, Year of Dragon, ha-ha, one year, happiness, celebrate the spring festival, healthy, friends
C: Jan 23, 2012, Spring Festival
New year, happy, Year of Dragon, ha-ha, healthy, one year, happiness, first
D: Feb 6, 2012, Lantern Festival
Happy, Lantern Festival, rice dumpling, happiness, firework, Valentine's Day of China
E: Feb 14, 2012, Valentine's Day
Valentine's Day, happy, together, gifts, lovers, chocolate, enjoy, husband, dear, bachelor
F: Mar 8, 2012, Women's Day
Happy, festival, Mother, thanks, women, enjoy, March 8th, girls
G: Apr 1, 2012, April Fools' Day
Ha-ha, happy, April Fools Day
H: Apr 3, 2012, Weibo opened comments to the users
Comments, Weibo, finally, the first comment
I: May 13, 2012, Mothers' Day
Mother, happy, Mothers' Day, happiness, healthy, I love you, enjoy
J: May 20, 2012, Day to speak out love
Ha-ha, happy, love, I love you, happiness, thanks, fighting, smile, speak out of love
K: Jun 1, 2012, Children's Day
Children's Day, care, I love you, ha-ha, festival, enjoy, gifts, kids, lovely, cute
L: Jun 8, 2012, Xin Liu's Birthday and the beginning of European Cup
Birthday, happy, Xin Liu, thanks; European Cup, beginning
M: Jun 23, 2012, Dragon Boat Festival and competitions of European Cup
Dragon Boat Festival, zongzi, Germany, fighting, Greece, European Cup
N: Aug 5, 2012, Dan Lin won the gold medal
Fighting, Dan Lin, China, Super Dan, Olympic Games, Chong Wei Lee, Yang Sun, congratulations, Champaign, badminton
O: Aug 23, 2012, Star Festival
Star Festival, happy, ha-ha, Lovers' Day, love, smile, together, gifts
P: Sep 30, 2012, Moon Cake Festival and the finals of the Voice of China
Mookcake Festival, Voice of China, mooncake, birthday of China, go home, together, Mochou Wu, National Day, moon, Bo Liang

doi:10.1371/journal.pone.0147079.t001

Table 2. Negative Topics Discussed by Ambivalent Users.

A: Mar 31, 2012, Sina closed the comments in Weibo
Comments, reposts, mean jokes, Sina, rumor, close
B: Apr 20, 2012, Volunteers rescued dogs in Kunming
Dogs, raining, rainstorm, Guangzhou, animals, dog dealers, volunteers, Kunming
C: May 12, 2012, The 4th Anniversary of Wenchuan Earthquake
Four years, Wenchuan, victims, rest in peace, 5.12, earthquake, survivals, fellows, blessing, commemorate
D: Jun 10, 2012, Group Stages in European Cup
Holland, European Cup, fighting, Italy, Germany, Portugal, Denmark
E: Jul 22, 2012, Torrential Rains in Beijing
Beijing, rest in peace, salute, heroes, rainstorm, sacrifice, death, local police station, policemen
F: Jul 27, 2012, Witness Insecurity
Ending, Witness Insecurity, Mr Xu, TVB, die
G: Aug 7, 2012, Xiang Liu Got Injured in the Olympic Games
Xiang Liu, China, London, Olympic Games, cry, fail down, referee, hero, fighting
H: Sep 30, 2012, The final of Voice of China
Voice of China, Mochou Wu, ads, Zhiwen Jin, shady deal, Bo Liang, Sherry Chang Huei-mei

doi:10.1371/journal.pone.0147079.t002

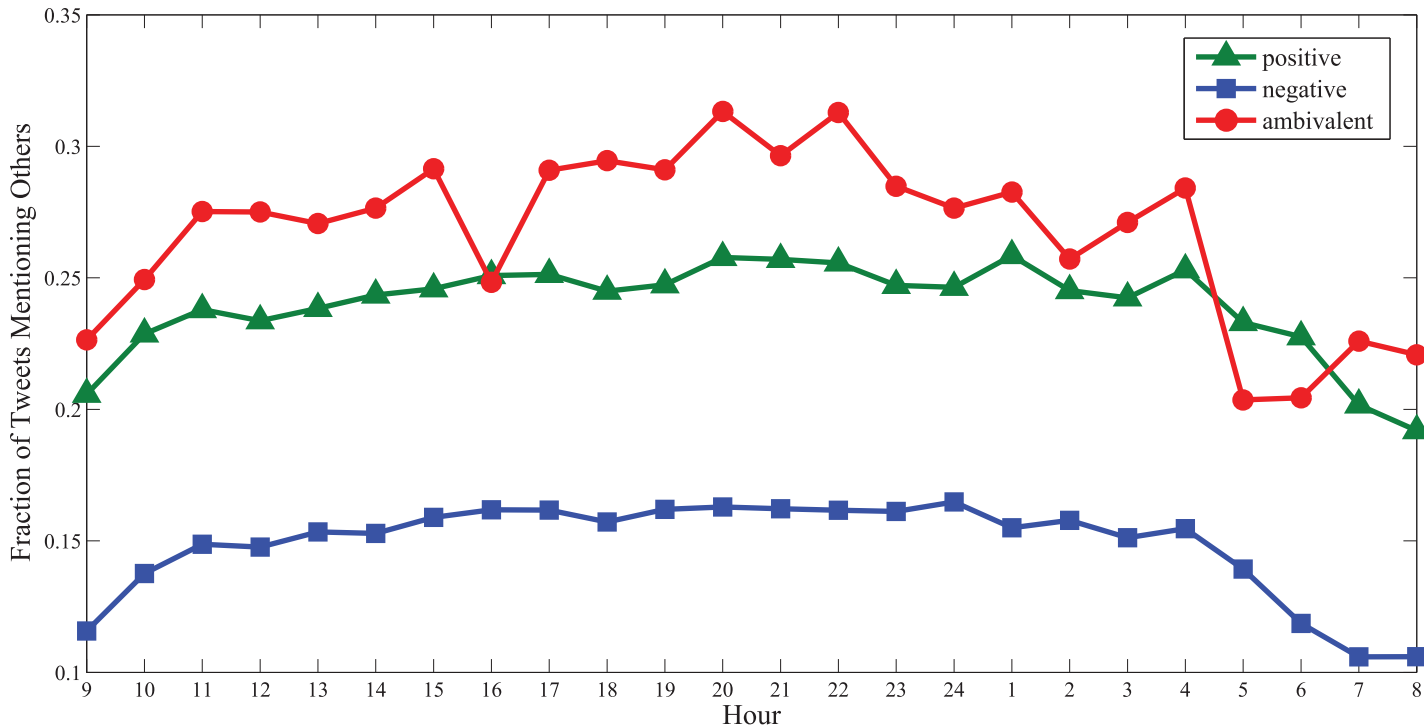


Fig 4. Fraction of tweets mentioning others at different hours. The ambivalent tweets contain significantly more mentions, especially at 20:00 and 22:00.

doi:10.1371/journal.pone.0147079.g004

Discussion

The sentiment shift of ambivalent tweets gives an important clue that online social media might affect or even manipulate users' mood status, which indeed has attracted special interests from researchers of various domains. For example, it has been found that socializing, information-seeking, and entertainment in online social media significantly influence user's positive emotion [33], persistent emotional expressions for individual users and channels are revealed in online chatting rooms [34], and positive terms in news feed may trigger online friends to express similar feelings [35]. From the dark side, however, disordered online social networking use leads to difficulties with emotion regulation [36], and daily time spent on social networking is related to depression [37]. These pilot studies, though being very interesting, have not touched the special group of ambivalent users hidden inside online social media. Nor did they provide adequate explanations to the abnormal behavioral patterns of these users. In particular, the sentiment shift from negative to positive implies that ambivalent users "actively" leverage the Weibo platform to realize self-correcting, which has clear contrast to the "unconscious" emotional influences discussed in the above-mentioned literature.

Nevertheless, we can find some evidence for the self-regulation of negative emotions from earlier sociological research. Indeed, based on offline survey studies, these research reveals that people could tune their moods through music, social interactions, enjoyable activities, shopping, religion, distraction [29] and cognitive reappraisal [30]. These results, though being widely accepted in the field of sociology, have hardly been tested in an online environment. Therefore, it would be very interesting if our findings on ambivalent users in Weibo could resonate with these offline findings. Indeed, as illustrated below, we can really explain the complicated behavioral patterns of ambivalent users from a sentiment self-regulation perspective.

First, as can be found in Section “Ambivalent Expression” in “Results”, ambivalent users are apt to express passive moods and post ambivalent tweets in the midnights or at the weekends. This implies that ambivalent users have the objective need for self-regulation of negative emotions, especially when they are not occupied by work. Hence, we can conjecture that publishing tweets should be an ideal way to distract them from negative mentality, especially the ambivalent tweets that can help them realize cognitive reappraisal. To understand the last point, recall the finding in Section “Sentiment Shift” in “Results” that ambivalent users demonstrate an unusual sentiment-shift behavior, most evidently from negative mood to positive mood. This pattern implies that ambivalent users try to reduce their passive feelings through building new cognitions to the context, including opinions, attitudes, sentiments and etc.

Second, as illustrated in Section “Topic Preference” in “Results”, ambivalent users are particularly fond of topics like entertainment and sports. These enjoyable activities could arouse relaxed feelings and thus might liberate them from depression or stress. Moreover, the special social interaction patterns of ambivalent users given in Section “Social Interaction” in “Results” further hint the existence of self-regulation. That is, ambivalent users prefer to mention their close friends in tweets and get encouraging feedbacks from those friends. This coincides with previous findings from surveys that supportive interactions with friends produce positive affect [38] and positive feedbacks on the profile enhance social self-esteem and well-being in online social networking [39]. More interestingly, our findings suggest that interacting with close friends through more private actions like “@” and “reply” in online social networking can help boost self-regulation. This subtle detail exhibits the possible uniqueness of self-regulation theory in an online environment, which cannot be touched in traditional offline studies.

In a nutshell, the self-regulation theory in sociology indeed well explain the distinct behaviors of ambivalent users in Weibo, and social media big data even enable us to find more distinct details in an online environment. It is also noteworthy that the self-regulation of ambivalent users seems to be conscious, which makes our finding an important supplement to existing online studies. Along this way, one interesting corollary according to the self-regulation theory is that ambivalent users would like to use shopping to self-correct negative mentality. We test this point by finding the shopping behaviors of ambivalent users around the Singles’ Day (or called Double Eleven, a famous day for promotion sales in China advocated by Taobao inc.) in 2012 and observing their emotional shifts (see [Methods](#)). It is interesting that similar to compulsive buyers [40], ambivalent users experience an evident increment in positive mood after shopping. Specifically, around 20.5% of ambivalent users post shopping tweets during the promotion sales while the ratio for ordinary users is just 7.8%. Moreover, as shown in [Fig 5](#), in time window of one day before and after shopping, the fraction of positive tweets posted by ambivalent users grows from 65.2% to 71.5% (negative tweets drop from 30.2% to 24.9%), while the value for ordinary users seems dropped slightly. This reversed shift of sentiments suggests that ambivalent users indeed regulate their negative feelings actively through shopping, which makes them ideal targets for online marketing.

Conclusion

In this paper, we take ambivalent expression in Weibo as an important clue for tracing unusual online behaviors of ambivalent users. Our findings first confirm traditional sociological studies by showing that the ambivalent expression is preferred by the east, especially the female. Further investigation reveals various interesting behaviors of ambivalent users, including the topic preference in entertainment and sports events, the increasing communication desire when relaxed, and the private contact to close friends via @ symbol in ambivalent tweets as well as the replies as feedback. Moreover, we find a clear negative to positive mode in sentiment shift

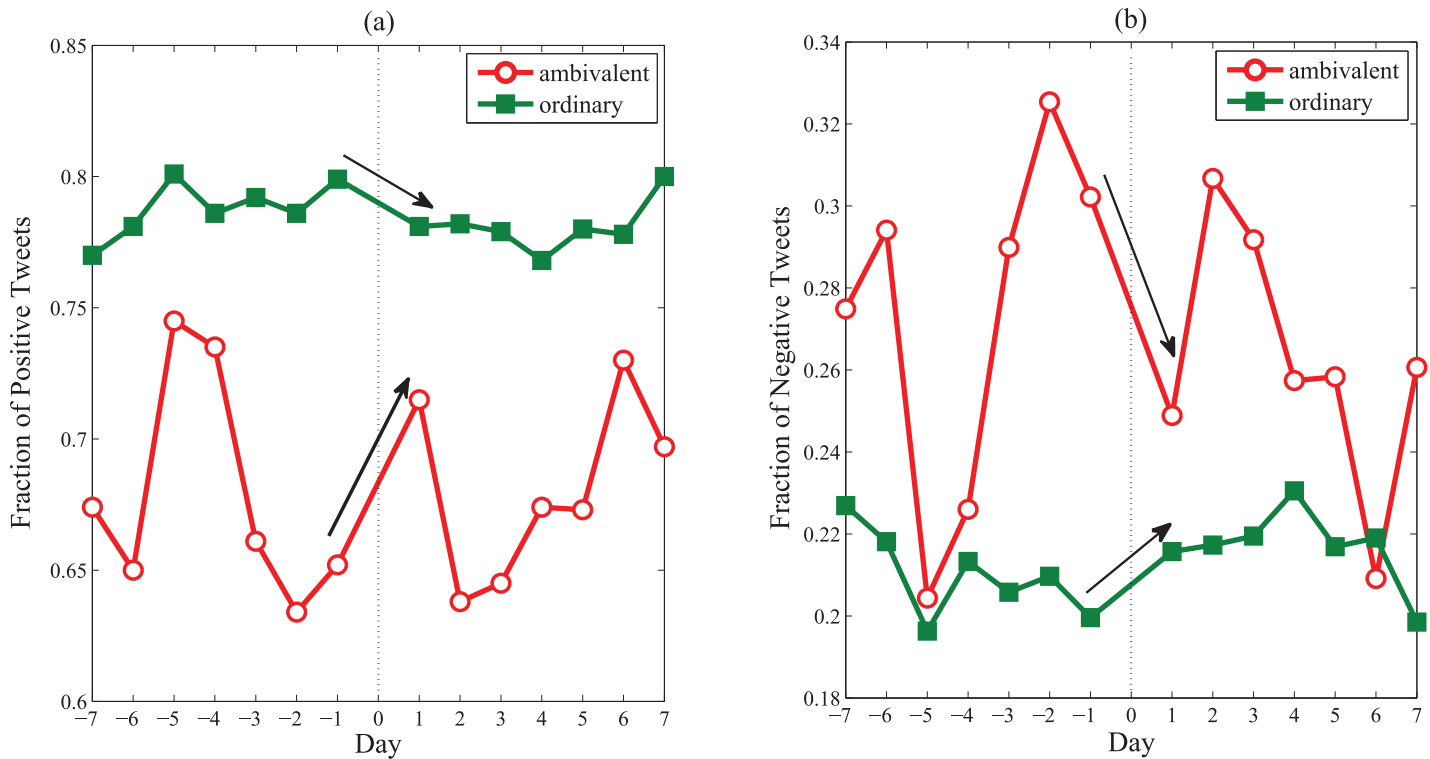


Fig 5. Fractions of tweets posted in three days before and after shopping. For x axis, the unit is a day and 0 stands for the shopping date. (a) Positive tweets. (b) Negative tweets.

doi:10.1371/journal.pone.0147079.g005

in ambivalent tweets, which guides us to find reasonable explanations to the above behaviors via the conventional theory of emotion management in sociology. As a result, we conclude that ambivalent expression is an effective way to self-regulate negative feelings in online social media, which is an important supplement to traditional offline sociological studies. Extended studies on the self-correcting effect of shopping suggest that our findings could shed light on applications like online marketing.

This study inevitably has limitations. First, for the online users investigated, the demographics other than gender, like age, occupation and geo-distribution, are not well discussed for data limitation, and their correlations with ambivalence expression thus remain unclear. Second, because of missing a complete online following graph, this study is not embedded into the context of a social network, which otherwise might help us structurally characterize the detailed mechanism of emotional interplay between different users. Therefore, further explorations focusing on the above points would be interesting directions in future work.

Methods

Weibo Data Weibo is a Twitter-like service in China, which has accumulated more than 500 million registered users since founded in 2009. Under the permission granted by Weibo, through its open API (<http://open.weibo.com>), we select 137,981 users by breadth-first search starting from several Weibo-verified seeds and then crawl each user's profile page. We collect totally 68 million tweets posted by these users in 2012, among which 22.7% are emoticonized tweets. In this paper, we only conduct the experiment on the data set of 2012, because from April of 2013, Alibaba has been a strategic cooperative partner of Weibo and Weibo therefore

experiences a sudden growth in marketing tweets, which might badly contaminate the analysis. We select emoticons appeared more than 1,000 times and manually label them as positive or negative according to their images and descriptive words. Specifically, three active Weibo users that are familiar with emoticon usage are asked to label each emoticon into positive, negative or neutral, and then their votes are summarized to determine the final label. To further ensure the consistency of an emoticon’s sentiment orientation, we also check the top-200 frequently used keywords extracted from the 1,000 tweets containing that emoticon. The results show that many popular emoticons are strongly and consistently bonded with one specific emotion. Finally, we have 115 verified emoticons, among which 79 are positive and 36 are negative, which are listed in [S1 Table](#). They appear in more than 84.3% of the emoticonized tweets, which consist of 312,456 ambivalent tweets, 13,430,096 positive tweets and still 2,858,413 negative ones. Our data sets’ download sources can be found in [S1 Datasets](#) and all the users’ ids and screen names are anonymized to protect the privacy. The authors also declare that there is no identifying personal information is collected or analyzed in this paper.

Twitter Data For the comparison purpose, 467 million Twitter posts published by 20 million users are employed in this study [41], covering about 30% of all public tweets posted in a 7-month period from June 1, 2009 to Dec. 31, 2009. We use 15 positive smileys (i.e., “:”), “:D”, “=D”, “=)”, “:]”, “=]”, “:-)”, “:-D”, “:-]”, “;)”, “;D”, “;]”, “;-)”, “;-D” and “;-]”) and 9 negative smileys (i.e., “:(”, “=(”, “:[”, “=[”, “:-(”, “:-[”, “:(’”, “:[’” and “D:’”), which are also adopted in [42] for polarity classification. Finally, we find 34 million tweets containing at least one smiley, among which 28 million are positive, 6 million are negative, and only 385,145 are ambivalent.

Computation of Topic Preference We draw everyday tweets published in 2012 by ambivalent and ordinary users, respectively, for topic inference. Let $p_u^e(t)$ be the number of tweets published by user u at day t with emotion e (either positive or negative). Let \bar{p}_u be the average number of tweets published daily by u . Then for a group of users G , the emotional intensity of e at day t is given by $p_G^e(t) = \sum_{u \in G} p_u^e(t) / \bar{p}_u$. We trace $p_G^e(t)$ continuously along the year 2012, with G being the group of ambivalent and ordinary users and e being the positive and negative emotions, respectively, to find emotional peaks that might indicate real events. We finally obtain 16 and 8 peaks with clear event semantics from the top-20 peaks in positive and negative emotion lines of ambivalent users, respectively, as shown in [Fig 3](#). The topics discussed by ambivalent users with extracted frequent terms are reported in [Tables 1](#) and [2](#) and the original Chinese terms and translation can be further found in [S2](#) and [S3](#) Tables.

Computation of Emotional Shift We take all the emoticonized tweets containing at least one of the 115 verified emoticons to study the emotional shift inside a clause of ambivalent tweets. We first cut ambivalent tweets into short clauses by punctuation characters or whitespaces, and label the clauses containing only positive emoticons as positive clauses. We then extract the top-2,000 frequent keywords from the positive clauses, denoted as W_{pa} . Similarly, we extract top-2,000 keywords from purely positive tweets and form the keyword set W_{pp} . The different set $\tilde{W}_p = W_{pa} \setminus W_{pp}$ with $|\tilde{W}_p| = 432$ then indicates the distinctive keywords used in ambivalent tweets to express positive feeling, which are subject to the emotional orientation test below. That is, for any keyword $t \in \tilde{W}_p$, we compute the occurrence frequencies of t in positive and negative tweets, denoted as $f_p(t)$ and $f_n(t)$, respectively. Let $f_p(\tilde{W}_p) = \sum_{t \in \tilde{W}_p} f_p(t) / |\tilde{W}_p|$ denote the positive intensity of \tilde{W}_p and $f_n(\tilde{W}_p) = \sum_{t \in \tilde{W}_p} f_n(t) / |\tilde{W}_p|$ denote the negative intensity. It is interesting to find that $f_p(\tilde{W}_p) : f_n(\tilde{W}_p) = 1.72\% : 3.23\%$, indicating that the positive clauses of ambivalent tweets actually convey more negative feelings than purely positive tweets. We also find this subtle mismatch from the negative clauses of ambivalent tweets. Let \tilde{W}_n denote the distinctive keywords used in ambivalent tweets to express negative feelings with $|\tilde{W}_n| = 423$, we have

$f_p(\tilde{W}_n) : f_n(\tilde{W}_n) = 4.17\% : 1.67\%$, indicating the stronger positive feelings in ambivalent tweets' negative clauses.

Computation of Shopping Behaviors We collect all the tweets published by ambivalent users from Nov. 4 to Nov. 18, 2012 for identifying their shopping behaviors around the Singles' Day. The tweets containing keywords like "buy", "shopping" and "taobao" are then selected as candidates and subject to manual labeling for finding shopping tweets depicting users' real shopping behaviors. For each shopping tweet, we trace back to its author's tweeting history and calculate the fraction of positive and negative tweets, respectively, in the specified time window before and after shopping. The results of corresponding ambivalent or ordinary users are then averaged to get the final shift patterns. Note that we here neglect neutral tweets without emoticons. Statistical significance of sentiment shift is also testified by randomly shuffling the posting time of each user's tweets, after which the sentiment shift produced by the shopping behavior disappears.

Supporting Information

S1 Datasets. The datasets download location. All the datasets collected from Weibo can be freely downloaded from the permanent location in figshare.com: <https://dx.doi.org/10.6084/m9.figshare.2060046>.

(PDF)

S1 Table. The entire list of emoticons and their descriptions.

(PDF)

S2 Table. Key terms for positive events.

(PDF)

S3 Table. Key terms for negative events.

(PDF)

Author Contributions

Conceived and designed the experiments: JW JZ YH. Performed the experiments: YH JZ. Analyzed the data: YH JZ JW. Wrote the paper: JW JZ.

References

1. Valenza G, Citi L, Lanatá A, Scilingo EP, Barbieri R. Revealing Real-Time Emotional Responses: a Personalized Assessment based on Heartbeat Dynamics. *Sci Rep*. 2014; 4:4998. doi: [10.1038/srep04998](https://doi.org/10.1038/srep04998) PMID: [24845973](https://pubmed.ncbi.nlm.nih.gov/24845973/)
2. Liu B. Sentiment analysis and subjectivity. In: *Handbook of Natural Language Processing*, Second Edition. Taylor and Francis Group, Boca; 2010.
3. Tossell CC, Kortum P, Shepard C, Barg-Walkow LH, Rahmati A, Zhong L. A longitudinal study of emoticon use in text messaging from smartphones. *Computers in Human Behavior*. 2012; 28(2):659–663. doi: [10.1016/j.chb.2011.11.012](https://doi.org/10.1016/j.chb.2011.11.012)
4. Ptaszynski M, Rzepka R, Araki K, Momouchi Y. Research on Emoticons: Review of the Field and Proposal of Research Framework. In: *Proceedings of 17th Association for Natural Language Processing*; 2011. p. 1159–1162.
5. Zhao J, Dong L, Wu J, Xu K. MoodLens: An Emoticon-based Sentiment Analysis System for Chinese Tweets. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD'12*. New York, NY, USA: ACM; 2012. p. 1528–1531.
6. Purver M, Battersby S. Experimenting with Distant Supervision for Emotion Classification. In: *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics. EACL'12*. Stroudsburg, PA, USA: Association for Computational Linguistics; 2012. p. 482–491.

7. Liu KL, Li WJ, Guo M. Emoticon Smoothed Language Models for Twitter Sentiment Analysis. In: AAAI; 2012.
8. Hu X, Tang J, Gao H, Liu H. Unsupervised Sentiment Analysis with Emotional Signals. In: Proceedings of the 22Nd International Conference on World Wide Web. WWW'13; 2013. p. 607–618.
9. Liu S, Li F, Li F, Cheng X, Shen H. Adaptive Co-training SVM for Sentiment Classification on Tweets. In: Proceedings of the 22Nd ACM International Conference on Conference on Information & Knowledge Management. CIKM'13; 2013. p. 2079–2088.
10. Derks D, Bos AER, von Grumbkow J. Emoticons and social interaction on the Internet: the importance of social context. *Computers in Human Behavior*. 2007; 23(1):842–849. doi: [10.1016/j.chb.2004.11.013](https://doi.org/10.1016/j.chb.2004.11.013)
11. Schnoebelen T. Do You Smile with Your Nose? Stylistic Variation in Twitter Emoticons. University of Pennsylvania Working Papers in Linguistics. 2012; 18:117–125.
12. Park J, Barash V, Fink C, Cha M. Emoticon Style: Interpreting Differences in Emoticons Across Cultures. In: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media; 2013.
13. Wood SA, Guerry AD, M SJ, Lacayo M. Using social media to quantify nature-based tourism and recreation. *Sci Rep*. 2013; 3:2976. doi: [10.1038/srep02976](https://doi.org/10.1038/srep02976) PMID: [24131963](https://pubmed.ncbi.nlm.nih.gov/24131963/)
14. González MC, Hidalgo CA, Barabási AL. Understanding individual human mobility patterns. *Nature*. 2008; 453:779–782. doi: [10.1038/nature06958](https://doi.org/10.1038/nature06958) PMID: [18528393](https://pubmed.ncbi.nlm.nih.gov/18528393/)
15. Zhao ZD, Yang Z, Zhang Z, Zhou T, Huang ZG, Lai YC. Emergence of scaling in human-interest dynamics. *Sci Rep*. 2013; 3:3472. doi: [10.1038/srep03472](https://doi.org/10.1038/srep03472) PMID: [24326949](https://pubmed.ncbi.nlm.nih.gov/24326949/)
16. Zhao ZD, Huang ZG, Huang L, Liu H, Lai YC. Scaling and correlation of human movements in cyberspace and physical space. *Phys Rev E*. 2014; 90:050802. doi: [10.1103/PhysRevE.90.050802](https://doi.org/10.1103/PhysRevE.90.050802)
17. Preis T, Moat HS, Bishop SR, Treleaven P, Stanley HE. Quantifying the Digital Traces of Hurricane Sandy on Flickr. *Sci Rep*. 2013; 3:3141. doi: [10.1038/srep03141](https://doi.org/10.1038/srep03141) PMID: [24189490](https://pubmed.ncbi.nlm.nih.gov/24189490/)
18. Zhao ZD, Cai SM, Lu Y. Non-Markovian character in human mobility: Online and offline. *Chaos*. 2015; 25:063106. doi: [10.1063/1.4922302](https://doi.org/10.1063/1.4922302) PMID: [26117100](https://pubmed.ncbi.nlm.nih.gov/26117100/)
19. Abelson RP. Whatever Became of Consistency Theory? *Personality & Social Psychology Bulletin*. 1983; 9:37–54. doi: [10.1177/0146167283091006](https://doi.org/10.1177/0146167283091006)
20. Bagozzi RP, Wong N, Yi Y. The Role of Culture and Gender in the Relationship between Positive and Negative Affect. *COGNITION AND EMOTION*. 1983; 13(6):641–672. doi: [10.1080/026999399379023](https://doi.org/10.1080/026999399379023)
21. Thelwall M, Buckley K, Paltoglou G, Cai D. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*. 2010; 61:2544–2558. doi: [10.1002/asi.21416](https://doi.org/10.1002/asi.21416)
22. Bollen J, Gonçalves B, Ruan G, Mao H. Happiness is assortative in online social networks. *Artif Life*. 2011; 17(3):237–251. doi: [10.1162/artl_a_00034](https://doi.org/10.1162/artl_a_00034) PMID: [21554117](https://pubmed.ncbi.nlm.nih.gov/21554117/)
23. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS ONE*. 2011; 6:e26752. doi: [10.1371/journal.pone.0026752](https://doi.org/10.1371/journal.pone.0026752) PMID: [22163266](https://pubmed.ncbi.nlm.nih.gov/22163266/)
24. Bliss CA, Kloumann IM, Harris KD, Danforth CM, Dodds PS. Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *Journal of Computational Science*. 2012; 3(5):388–397. doi: [10.1016/j.jocs.2012.05.001](https://doi.org/10.1016/j.jocs.2012.05.001)
25. S%vuvakov M, Mitrović M, Gligorijević V, Tadić B. How the online social networks are used: dialogues-based structure of MySpace. *J R Soc Interface*. 2012; 10:20120819. doi: [10.1098/rsif.2012.0819](https://doi.org/10.1098/rsif.2012.0819)
26. Mitchell L, Frank MR, Harris KD, Dodds PS, M DC. The Geography of Happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place. *PLoS ONE*. 2013; 8(5):e64417. doi: [10.1371/journal.pone.0064417](https://doi.org/10.1371/journal.pone.0064417) PMID: [23734200](https://pubmed.ncbi.nlm.nih.gov/23734200/)
27. R Frank M, Mitchell L, Dodds PS, Danforth CM. Happiness and the Patterns of Life: A Study of Geolocated Tweets. *Sci Rep*. 2013; 3:2625. doi: [10.1038/srep02625](https://doi.org/10.1038/srep02625)
28. Fan R, Zhao J, Chen Y, Xu K. Anger is more influential than joy: sentiment correlation in Weibo. *PLoS ONE*. 2014; 9:e110184. doi: [10.1371/journal.pone.0110184](https://doi.org/10.1371/journal.pone.0110184) PMID: [25333778](https://pubmed.ncbi.nlm.nih.gov/25333778/)
29. Gallup JG, Castelli J. The people's religion: American faith in the Nineties. New York Macmillan Publishing Company; 1989.
30. Dandoy AC, Goldstein AC. The use of cognitive appraisal to reduce stress reactions: A replication. *Journal of Social Behavior and Personality*. 1990; 5:275–285.
31. Mathiesen J, Angheluta L, Ahlgren PTH, Jensen MH. Excitable Human Dynamics Driven by Extrinsic Events in Massive Communities. *PNAS*. 2013; 110:17259–17262. doi: [10.1073/pnas.1304179110](https://doi.org/10.1073/pnas.1304179110) PMID: [24101482](https://pubmed.ncbi.nlm.nih.gov/24101482/)

32. Granovetter MS. The Strength of Weak Ties. University of Chicago Press; 1974.
33. Apaolaza V, He J, Hartmann P. The effect of gratifications derived from use of the social networking site Qzone on Chinese adolescents positive mood. *Computers in Human Behavior*. 2014; 41(0):203–211. doi: [10.1016/j.chb.2014.09.029](https://doi.org/10.1016/j.chb.2014.09.029)
34. Garas A, Garcia D, Skowron M, Schweitzer F. Emotional persistence in online chatting communities. *Sci Rep*. 2012; 2:402. doi: [10.1038/srep00402](https://doi.org/10.1038/srep00402) PMID: [22577512](https://pubmed.ncbi.nlm.nih.gov/22577512/)
35. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014; 111(24):8788–8790. doi: [10.1073/pnas.1320040111](https://doi.org/10.1073/pnas.1320040111)
36. Hormes JM, Kearns B, Timko CA. Craving Facebook? Behavioral addiction to online social networking and its association with emotion regulation deficits. *Addiction*. 2014; 109:2079–2088. doi: [10.1111/add.12713](https://doi.org/10.1111/add.12713) PMID: [25170590](https://pubmed.ncbi.nlm.nih.gov/25170590/)
37. Pantic I, Damjanovic A, Todorovic J, Topalovic D, Bojovic-Jovic D, Ristic S, et al. Association between online social networking and depression in high school students: Behavioral physiology viewpoint. *Psychiatria Danubina*. 2012; 24(1):90–93. PMID: [22447092](https://pubmed.ncbi.nlm.nih.gov/22447092/)
38. Oh HJ, Ozkaya E, Larose R. How Does Online Social Networking Enhance Life Satisfaction? The Relationships Among Online Supportive Interaction, Affect, Perceived Social Support, Sense of Community, and Life Satisfaction. *Comput Hum Behav*. 2014 Jan;30:69–78. doi: [10.1016/j.chb.2013.07.053](https://doi.org/10.1016/j.chb.2013.07.053)
39. Valkenburg P, Jochen P, Alexander SP. Friend Networking Sites and Their Relationship to Adolescents' Well-Being and Social Self-Esteem. *CYBER PSYCHOLOGY & BEHAVIOR*. 2006; 9(5):584–590. doi: [10.1089/cpb.2006.9.584](https://doi.org/10.1089/cpb.2006.9.584)
40. Faber RJ, Christenson GA. In the mood to buy: Differences in the mood states experienced by compulsive buyers and other consumers. *Psychology & Marketing*. 1996; 13:803–819. doi: [10.1002/\(SICI\)1520-6793\(199612\)13:8%3C803::AID-MAR6%3E3.0.CO;2-J](https://doi.org/10.1002/(SICI)1520-6793(199612)13:8%3C803::AID-MAR6%3E3.0.CO;2-J)
41. Yang J, Leskovec J. Patterns of Temporal Variation in Online Media. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. WSDM'11*. New York, NY, USA: ACM; 2011. p. 177–186.
42. Speriosu M, Sudan N, Upadhyay S, Baldrige J. Twitter Polarity Classification with Label Propagation over Lexical Links and the Follower Graph. In: *Proceedings of the First workshop on Unsupervised Learning in NLP*. Edinburgh, Scotland: Association for Computational Linguistics; 2011. p. 53–63.