



Identifiability as an “Antidote”: Exploring Emotional Contagion and the Role of Anonymity in Twitter Discussions on Misinformation

Chen (Crystal) Chen^(✉) , Hao Yuan, and Mike Z. Yao

University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA
{chenc4, haoyuan4, mzyao}@illinois.edu

Abstract. Misinformation carries both distorted facts and sophisticated emotional signals. Comparing to facts that could be labeled as true or false, we are more concerned about contaminative negative emotions transferring digitally among users. In this study, we explored an emotional contagion effect among misinformation discussion participants on Twitter. We analyzed the sentiment of 573 tweets in 192 discussion threads. Our result revealed that highly emotional tweets do not have a universal effect on the online discussions, but it affects those individuals with limited social and personal identity cues (i.e., being anonymous). We found that anonymous members of the online discussion are more susceptible to emotional contagions than those are not. We also suggest coping strategies that protect social media users’ emotional well-being during the era COVID-19.

Keywords: Emotion contagion · Anonymity · SIDE · Sentiment analysis

1 Introduction

The proliferation of social media communication allows users to engage in conversations and exchange opinions related to various topics, but little validation of online information and the careless use of social media also lowers the bar for misinformation propagation. The spread of misinformation has been a concern since the beginning of coronavirus pandemic (COVID-19). Distorted facts and misinformation not only deepen users’ misunderstandings of the virus and the pandemic, but they also often carry negative emotional signals that may affect Internet users’ emotional well-being. Online discussions that contain misinformation or involve accusations of misinformation are often contentious and heated. Such discussions divide the online community and spread negative sentiment. There has been initial evidence indicating global emotional contagion in the era of COVID-19. For instance, Kabir and Madria [11] identified an increasing number of aggressive sentiments in tweet threads compared to positive counterparts since the beginning of COVID-19. Medford and colleagues’ [17] exploratory study showed that anger and fear are the top two emotions identified in users’ COVID-19 twitter posts.

The current study starts with exploring emotional contagion within discussion threads about misinformation on Twitter. Emotional contagion refers to individuals’

tendency to automatically mimic and synchronize others’ emotions [9] While different emotions can transfer digitally among users, negative emotions are particularly contaminative [13]. Previous studies found that when reducing positive content in Facebook’s News Feed, participants posted more negative words in their own posts. In a similar way, we expect to observe an emotional contagion effects in discussions threads about COVID-19 and misinformation on Twitter.

In this study, we first examine whether the emotional signals carried by an initial tweet would influence the valence and sentiment of subsequent replies. We then explore the moderating effect of online anonymity on the emotional contagion effect. Following principles of the Social Identity model of Deindividuation Effects (SIDE) [22, 24, 25], we speculate that anonymous participants in an online discussion group might be particularly susceptible to negative emotion contagions. The SIDE model proposes that online anonymity obscures social and personal identity cues and interpersonal differences, thereby diminishes the relative importance of interpersonal concerns in online groups [24]. Under such conditions, depersonalized individuals become susceptible to group influence [22, 33]. For instance, past research has found that when participants were exposed to an aggressive group discussion in which their peers spoke aggressively to each other, anonymous participants were more susceptible to the influence of negative emotions and behave in a similar fashion; in contrast, identified participants seemed to be immune from such influence [26]. As such, we are particularly interested in exploring identifiability as an “antidote” for the emotional contagions in negative online discussion groups.

We collected twitter posts and replies discussing COVID-19 misinformation from January to June using a web crawler and applied a pre-trained sentiment analysis model “VADER” [10] for language analysis. By applying the SIDE model to a more natural user generated content on Twitter, this exploratory study aims to contribute to a better understanding of factors that affect users’ emotional well-being during COVID-19 and come up with a proper coping strategy that sets the stage for a better mediated communication environment.

2 COVID-19 Misinformation and Negative Emotion

Since the beginning of the quarantine and execution of social distance, individuals’ increasing dependence on social media also raises concerns of misinformation exposure. Previous studies have identified an enormous amount of misinformation on twitter. For instance, Sharma and colleagues [27] using keywords related to COVID-19 collected tweets from 182 countries, resulting in a subset of 4.58M English tweets with external links. Among this dataset, 150.8K (3.29%) tweets were identified as misinformation source tweets, carrying unreliable, conspiracy, or biased news sources. Although different websites used fact-checkers and took quick actions to remove or attach warnings to these pieces of misinformation, the effectiveness of these actions varies among companies [2]. According to Brennen and colleagues’ [2] study, about 24% of the COVID-19 misinformation remains up without warning labels on Facebook. On YouTube, this number rises about 3%, with 27% of the misinformation remaining unchecked. Twitter has the worst condition, with about 59% of misinformation remaining unchecked.

Misinformation creates misleading guidance in the face of COVID-19, but it also carries sophisticated emotional signals, particularly negative emotions such as anger and fear [4]. The worst part is that the more anger a fake news source carries, the more likely it becomes viral, leading to a higher exposure to normal social media users. Recent research collected the fake news from both Weibo and Twitter and investigated the emotion carried by misinformation across platforms. When compared to real news, fake news from both platforms contained significantly higher proportions of anger and lower proportions of joy. Also, researchers found a positive association between proportion of anger in the fake news and the number of retweets. If a piece of misinformation contains more anger, it has a higher chance to spread out and infect the vulnerable general public. We are concerned that the prevalence of misinformation that carries negative emotion may contaminate social media users' emotional well-being and mental health during COVID-19. This concern is related to emotional contagion, a well-established finding from social psychology.

2.1 Emotional Contagion

Emotional contagion refers to individuals' tendency to automatically mimic and synchronize others' expressions, vocalizations, postures, and movement, and consequently, to converge emotionally [9]. For instance, when showing human pictures with varying facial expressions to viewers, happiness and sadness expressions significantly and repeatedly evoke the same emotion among viewers, even when the presentation duration lasts only 500 million seconds [36].

Recent work has indicated that emotional contagion does not always require in-person interaction and non-verbal cues; rather, emotional contagion can happen through social networks [13]. An experiment conducted on Facebook showed that, when reducing positive content in the News Feed, participants posted more negative words and less positive words in their own posts. When reducing negativity in the News Feed, the opposite pattern occurred [13]. Further, another study pointed out that when one user gets affected by one particular emotion on social media, this emotion is likely to transmit to friends in their social network [5]. In their study, researchers collected data from millions of Facebook users and found that bad weather negatively influences the emotional content of posters who experienced bad weather, but also, that this negative emotion additionally contaminates the status content of posters' friends who did not experience bad weather.

2.2 Emotional Contagion in the Era of COVID-19

Different emotions appear to be contagious, including anger, happiness, and anxiety [1], but negative emotions appear to be the most influential [21]. The effect of negative emotional contagion has received some evidence in studies conducted during the recent pandemics (i.e., during SARS: [37]; during H1N1:[31]; during COVID-19: [20, 35]). For instance, Wheaton and colleagues [35] conducted a survey among 600 students

during COVID-19 and found high COVID-19 media consumption was associated with high anxiety among participants. For those who are more susceptible to emotional contagion, COVID-19 media consumption had a higher impact on their elevated obsessive-compulsive disorder (OCD) symptoms. In this case, reading misinformation containing anger and fear may also elicit more negative emotions among users.

Indeed, when facing COVID-19 misinformation, not everyone is a passive recipient. Social media empowers users to debunk misinformation and clarify misleading content. However, those who intend to debunk misinformation also have a high chance to get infected with negative emotion [18]. In a recent study, researchers found that susceptibility to and severity of misinformation induced negative emotions among vaccine supporters, including anticipated guilt and anger [30]. In this case, even though misinformation per se did not distort users’ understanding of vaccines, it negatively impacted users’ emotions.

That being said, when it comes to discussion related to negatively emotional misinformation, regardless of communicators’ role as a misinformation disseminator, debunker, or reader, their emotions are likely to be contaminated during the interaction. Their negative emotions will be reflected in the linguistic valence of their responses and further influence more users. Emotional contagion happening digitally and transmitting via social networks may magnify the intensity of emotional synchronization, which is particularly sensitive to social media users’ emotional well-being and mental health in the era of COVID-19 [32]. As such, we propose the following:

H1. Linguistic valence of the lead post in an online discussion thread will transmit to its replies, such that a negative post will more likely be followed by negative replies.

2.3 The SIDE Model: Depersonalization Versus Deindividuation

Following the SIDE model, we speculate that that users’ online anonymity may modify the impact of negative emotion contagions. The SIDE model states that online anonymity obscures *personal features* and *interpersonal differences* but heightens one’s sensitivity to the situational and group norms [24, 33]. More specifically, the model attributes individuals’ heightened norm sensitivity to the influence of depersonalization, a tendency to perceive the self and others not as individuals with a range of idiosyncratic characters and ways of behaving but as representatives of social groups or wider social categories that are made salient during interaction [22].

The idea of depersonalization is different from deindividuation, a concept that was developed to explain antisocial behavior in crowds [6, 39]. Deindividuation theory proposes that, with reduced social cues (i.e., high visual anonymity or uniform representation of group members) in a crowd, individuals have a *decreased sense of self identity*, and consequently, exhibit a deregulation of social behavior [12, 39]. On the other hand, depersonalization from the SIDE model suggests conformity happens not because individuals lose self-identity. Rather, it happens because of a *reduced awareness of group member differences*, an increased salience of social identity among group members, and consensus of group norms. Further, crowd behavior is not always antinormative. Rather, depersonalization is the result of conformity to norms, which could be either positive or negative, depending on the nature of the group. Postmes and Spears [25] conducted

a meta-analysis with 60 empirical studies and found weak support for the deindividuation theory, suggesting that being anonymous is not the only factor that contributes to individuals' anti-normative behavior.

Although the SIDE model rejects the process account of antisocial behavior in the deindividuation theory, it builds on previous theory's distinctions between two aspects of anonymity, high visual anonymity and uniform representation of group members, and formed cognitive and strategic dimensions of the SIDE model [29]. The cognitive dimension of the SIDE model refers to "anonymity of" and how the anonymity of/within the group can enhance the salience of group identity. The strategic dimension of the SIDE model refers to "anonymity to", the reduced accountability to others. The reduced accountability allows group members to nominate in group behaviors that may otherwise be sanctioned by the outgroup.

2.4 The Influence of Anonymity Based on the Cognitive Dimension of the SIDE Model

Previous research has collected empirical evidence in favor of the cognitive dimension of the SIDE model (i.e. [14, 16, 26, 28, 38]). For instance, Spears, Lea, and Lee [28] conducted a study in which participants were directed to discuss a range of controversial topics using a text-based synchronous CMC system. Researchers manipulated participants' visual anonymity by arranging them to co-present in the same room or separating them into different rooms. Researchers also provided participants with norm references where participants could know their groups' preferred solution for each topic. The results revealed that participants show shifts in the direction of group norms when their social identities are salient, and they are isolated (and therefore anonymous).

In addition, when aggressiveness becomes the group norm, researchers found participants in an anonymous condition are more likely to adopt similar ways of communication styles, when compared to participants in a non-anonymous condition. For instance, Rosner and Kramer [26] conducted a lab experiment investigating the moderating role of anonymity on the effect of aggressive group norms on individuals' commenting behavior. Researchers manipulated anonymity by two measures: on one hand, participants could reply using their registered Facebook account (identified) or reply without registration. On the other hand, participants could see peer comments either with or without the author's identifiable information (i.e., name, profile). The result revealed that, when peer comments include more aggressive wording, participants who perceive commenters have a high anonymity also use more aggressive expressions in their comments, but those who perceive commenters have a low anonymity are not influenced by aggressive comments. Further, there were studies zoomed in on specific information process [15]. For instance, Lee [15] extended the depersonalized effect by showing that identified users paid more attention to and influence of the strengths and weakness of specific arguments, while among those who were not individually identifiable, they paid less attention to the argument quality, but more were susceptible to the group influence.

These studies shed light on our current situation and suggest that when joining an online discussion thread, if users are anonymous or perceive other commenters as anonymous, they may be vulnerable to emotional contagion and will reply in a similar fashion to the lead post. However, if they are identified or perceive other commenters

as more identified, their personal identity may remain salient. In this case, they will be less susceptible to the influence of the affective comments in misinformation discussion. The current study proposed the following hypotheses:

H2. In an online discussion thread, a) when repliers are anonymous, they are more susceptible to the emotional contagions in the environment. b) On the other hand, the effect of emotional contagion from the lead post will diminish among identified repliers.

3 Method

3.1 Data Collection

We started with an existing COVID-19 misinformation tweet dataset (CMU-MisCov19) [19]. The original dataset was collected based on hashtags and keywords in conjunction with “coronavirus” and “covid” on three days: 29th March 2020, 15th June 2020, and 24th June 2020 (Table 1). Each of these collections extracted a set of tweets from their corresponding week. The original dataset contained 4573 tweets, comprising 3629 users.

This dataset helped us to identify a list of tweets that posted topics related to COVID-19 misinformation. We collected the tweet IDs from this dataset and further scraped the webpage. We applied the python Selenium package and mimicked real visits to each tweet in this dataset. During the web scrape, we found that twitter blocked some original posts and replies due to policy violation and misinformation dissemination, so some content was not retrievable. In addition, we did not collect posts that do not have replies. The final dataset contained a total of 573 tweets, including 192 posts and 381 replies. We further used the python BeautifulSoup Soup package to parse and extract relevant variables, including posters’ and repliers’ profile images, usernames, ids, posting timestamps, and their tweets.

Table 1. Keywords and Hashtags used in the data collection [19]

Type	Terms
Keywords	Bleach, vaccine, acetic acid, steroids, essential oil, saltwater, ethanol, children, kids, garlic, alcohol, chlorine, sesame oil, conspiracy, 5G, cure, colloidal silver, dryer, bioweapon, cocaine, hydroxychloroquine, chloroquine, gates, immune, poison, fake, treat, doctor, senna makki, senna tea
Hashtags	#nCoV2019, #CoronaOutbreak, #CoronaVirus, #CoronavirusCoverup, #CoronavirusOutbreak, #COVID19, #Coronavirus, #WuhanCoronavirus, #coronaviris, #Wuhan

3.2 Measure

Sentiment Analysis. In this study, we intended to use sentiment analysis to *estimate* individual posts’ and replies’ emotions (as indicated by language valence). This is because affective comments often carry emotion-expressive functions, which are reflected in

their language valence [7]. In this study, we chose a pre-trained model “VADER” over human raters to measure language valence due to its convenience and accuracy. The training of the model started with examining existing sentiment word banks (LIWC, ANEW, and GI) and incorporating lexical features common to emotional expressions in CMC, including a full list of Western-style emoticons. Then researchers collected intensity ratings on each lexical feature candidate from 10 independent coders, ranging from “−4 extremely negative” to “4 extremely positive.” In addition, the model also incorporated five generalizable heuristics based on grammatical and syntactic cues that further changed the sentiment intensity. The heuristics included punctuation (i.e., the exclamation mark “!”), capitalization, degree modifiers (i.e., degree adverbs such as “extremely”), contrastive conjunctions (i.e., “but” stands for a shift in sentiment polarity), and tri-grams preceding sentiment-laden lexical features (i.e., “The food here isn’t really all that great”). Researchers also conducted two quantitative studies and found that the trained model outperformed individual human raters [10].

The VADER provided an output in the following form: { ‘neg’: 0.267, ‘neu’: 0.662, ‘pos’: 0.072, ‘compound’: −0.9169}. This means there is a 26.7% possibility that the sentence is negative, a 66.2% possibility that the sentence is neutral, and a 7.2% possibility that the sentence is positive. The final “compound” score is a normalized, weighted, and composite score, indicating the sentence carries negative emotion. Our calculation of valence is based on the compound score. We grouped posts with their replies and ended up with a total of 192 groups.

Post/Reply Valence. The posts’ and replies’ emotions were estimated by sentiment analysis and indicated by their valence. If the post or reply had a negative valence (valence < 0), we regarded it as having a negative emotion. If the post or reply had a positive or neutral valence (valence ≥ 0), we regarded it as having a neutral or positive emotion.

Anonymity. Individuals’ anonymity was measured based on the user’s first and last name identification. First, username identification was completed by matching usernames with an existing open-source name database¹. The database included approximately 160k first names and 100k last names. We coded 0 for each identified first name and last name and 1 for cells with no name. For each reply, we combined first name and last name scores and formed their name anonymity scores, with a higher score indicating a higher level of anonymity (N[both names|0] = 79, N[one name|1] = 126, N[no name|2] = 176). We regarded replies with no name as an anonymous group and combined replies with one name or both names to form an identified group (N[anonymous] = 176, N[identified] = 205).

4 Result

Hypothesis 1 predicted that the emotion of a post, as reflected by its language valence, will transmit to its replies, such that a negative valence post will be followed by negative valence replies. A simple linear regression model was conducted to test this hypothesis, with the valence of the post as the independent variable and valence of the reply as the dependent variable. The result showed that the valence of the post does not significantly

¹ <https://github.com/philipperemy/name-dataset>.

influence the valence of the reply ($b = 0.01$, $t = 0.15$, $p > .05$). As a result, H1 was not supported.

Hypothesis 2 proposed the valence of the post by the replier’s anonymity interaction. We proposed different patterns of commenting behavior among repliers classified as anonymous or identified when they are exposed to posts with positive or negative valence. Because we believe that the valence of the replies is influenced by factors from an individual level (between person difference: anonymous or identified) and factors from a group level (between group/thread difference: the valence of the post), we fitted a multilevel model with cross-level interaction. At Level 1, we modeled the valence of the post as a predictor of the valence of the reply across groups/threads. At Level 2, we examined the individual anonymity (anonymous = 0, identified = 1) as predictors of the intercept of repliers’ language valence (i.e., average replies’ language valence across all misinformation threads) and the level 1 random slope. The cross-level interaction between the estimated slope at level 1 and repliers’ anonymity (anonymous or identified) assessed whether the association between the valence of the reply and the valence of the post differed as a function of the replier’s individual anonymity. All variables were mean-centered (i.e., individual’s raw score minus group mean) prior to model tests.

To test our hypotheses, we fit a series of models in SPSS using maximum likelihood estimation. The intercept-only model indicated significant variability in replies’ language valence at both the level 1 (.16) and level 2 (.03); the ICC was 16.93%, indicating that 17% of the variation in reply valence was between persons, whereas 83% was between groups (threads). Results from the conditional multilevel model are summarized in Table 2. At level 1, the random linear slope was not significant, such that the valence of the posts did not predict the valence of replies ($B = .03$, $SE = .04$, $t = 0.64$, $p > .05$). At level 2, the cross-level interaction between anonymity (level 2) and valence of the post (level 1) in the prediction of the valence of replies was significant, indicating that the influence of group norm on valence of replies differed depending on users’ anonymity ($B = -.16$, $SE = .08$, $t = -1.97$, $p < .05$) (See Fig. 1).

To probe this significant cross-level interaction, we examined the influence of posts’ valence for the identified and the anonymous groups respectively. Because the anonymous group was coded as 0 in the main model reported in Table 2, the level 1 slope represents the simple slope for the anonymous group. To obtain the simple slope for the identified group, we fit an identical multilevel model with the anonymity variable reverse coded (i.e., identified = 0, anonymous = 1). As shown in Fig. 1, simple slope tests revealed a significant positive association between the valence of anonymous reply and the valence of the post ($B = 0.11$, $SE = .06$, $t = 1.92$, $p = .05$), and an insignificant negative association between the valence of identified reply and the valence of the post ($B = -0.05$, $SE = .04$, $t = -.80$, $p > .05$). The result suggested that individual’s anonymity modified replies’ language valence in the twitter thread. For anonymous users, the emotions in posts have a positive effect on their replies, and this effect is significant (H2a supported). For identified users, the emotion carried in the post has a negative effect on their replies, but the influence is not significant (H2b supported).

Table 2. Results of hierarchical linear modeling

Variables	Null	Model 1: adding Level 1 predictors		Model 2: effect of level 1 and level 2 predictors		Model 3: cross-level interaction effects	
		Coefficient	SE	Coefficient	SE	Coefficient	SE
Level 1							
Intercept		.00	.02	.03	.03	.04	.03
Post valence		.03	.04	.03	.04	.11	.06
Level 2							
Anonymity				-.06	.04	-.06	.04
Cross-level							
Post valence* Anonymity						-.16*	.08
Additional information							
ICC	0.17						

Note. *p < .05

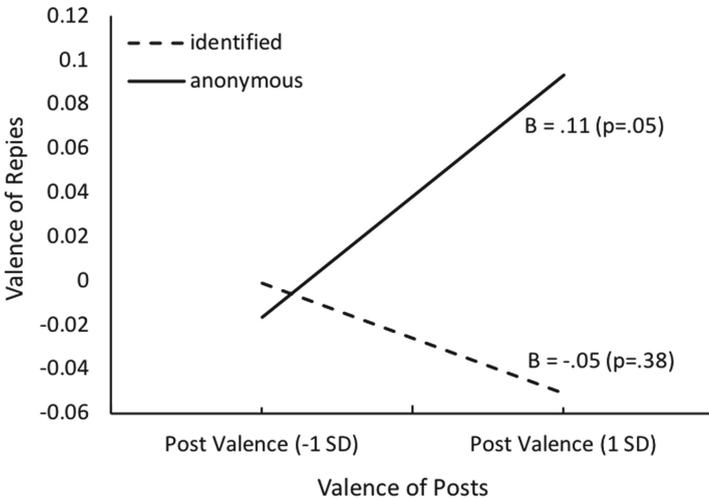


Fig. 1. Associations between valence of the posts and valence of the replies given identified or anonymous users.

5 Discussion

Misinformation is more likely to transmit among social media users in highly uncertain emergencies, such as disease outbreaks and pandemics [3]. Increasing exposure to misinformation may not only deepen users’ understanding of the disease but also immerse users in unnecessary anxiety, anger, and fear. A recent survey conducted by the National

Center for Health Statistics (NCHS) in January 2021 estimated that 35.8% of adults have symptoms of anxiety disorder, which is 27.6% higher than estimates in 2019. Empirical evidence suggested that these negative emotions are associated with excessive social media exposure [35], and that negative emotions could transmit digitally across social networks [5]. As a result, the current study starts with exploring emotional contagion among misinformation discussion participants on Twitter. We then investigated the moderating effect of online anonymity on the emotional contagion effect under the guidance of the SIDE model.

The result showed that highly emotional tweets do not have a universal effect on online discussions, but it affects those individuals with limited social and personal identity cues (i.e., being anonymous). We found that anonymous members of the online discussion are more susceptible to emotional contagions than those are not. When users being anonymous, they are more susceptible to emotional contagion. When users being identified, they are more immune to emotional contagion. This finding is aligned with the prediction of the SIDE model that users tend to be more susceptible to group influence when they have less personal identity cues.

By further looking into group differences, we found that anonymous and identified users responded to posts carrying positive emotion differently, with anonymous users replying in a more positive tone. Anonymous and identified users did not differ significantly when replying to a post carrying negative emotion. However, the result revealed a trend that anonymous users reply in a negative tone to posts with negative emotion. The reason that we did not see a significant difference between groups when they replying to a negative post may be due to the sample bias. Twitter has enforced a stricter regulation to remove potentially harmful information since last March, and the automated systems have challenged more than 1.5 million accounts targeting COVID-19 discussions with spammy or manipulative behaviors [34]. As a result, a number of posts and replies containing sensitive or improper contents may have been removed by Twitter.

This study also suggests coping strategies that protect social media users’ emotional well-being during the era of COVID-19. The result showed that users are more immune to the influence of emotional contagion when they are identified with names. This finding complies with the prediction of the SIDE model, except that our goal is to reduce (negative) social influence by maintaining one’s personal identity and refusing the social influence (the null hypothesis in the SIDE model). Our study suggests that being identified with names in social media may be one way to reduce negative social influence, but it is not the only way. Previous experimental findings suggested that visual identification, such as unique profile pictures/avatars, also highlights one’s personal identity and inhibit social identity [16]. That being said, we encourage users to find their own way (i.e., create unique profile avatar, use meaningful pseudonyms) to highlight their personal identities and reduce negative social influence.

6 Limitations and Future Research

This study is a timely piece of research that investigates the effect of emotional contagion and suggests coping strategies for users during the era of COVID-19, but it is limited in its nature as a cross-sectional design that precludes causal inferences. For example,

we cannot determine if there is a third variable outside of this observational dataset that contributes to users' sensitivity to emotional contagion. A more powerful design would be a controlled lab experiment that manipulates the valence of the lead post and compares group differences given different levels of individual identity salience and group uniformity. Also, as a study applying the SIDE model, it does not have a strict manipulation of the group or a salient group identity. In the current study, a group is regarded as nodes in communication networks [23], such that a post and its subsequent replies are treated as one group. Although this approach defines a group structurally rather than psychologically, users' posting and reacting to COVID-19 misinformation topics indicated a common interest and some shared social identity. Previous research also indicated that a conformity to a form of communication (such as humor, verbal aggression) is an indication of social identity influences [8]. Regardless, a better measure of the group would be to apply network analysis and identifying users who interact most within all misinformation conversations.

7 Conclusion

Notwithstanding these limitations, the present observational study highlights the group difference between users who are anonymous versus identified in their sensitivity to the emotional contagion coming from misinformation discussions. The findings indicated that users with a reduced sense of depersonalization through having an identified name are more immune to the emotional contagion. Based on the finding, we proposed coping strategies for inhibiting global emotional contagion.

References

1. Behnke, R.R., Sawyer, C.R., King, P.E.: Contagion theory and the communication of public speaking state anxiety. *Commun. Educ.* **43**, 246–251 (1994)
2. Brennen, J.S., Simon, F., Howard, P.N., Nielsen, R.K.: Types, sources, and claims of COVID-19 misinformation. *Reuters Inst.* **7**, 3–1 (2020)
3. Carey, J.M., Chi, V., Flynn, D.J., Nyhan, B., Zeitzoff, T.: The effects of corrective information about disease epidemics and outbreaks: evidence from Zika and yellow fever in Brazil. *Sci. Adv.* **6**(5), eaaw7449 (2020)
4. Chuai, Y., Zhao, J.: Anger makes fake news viral online. [arXiv:2004.10399](https://arxiv.org/abs/2004.10399) (2020)
5. Coviello, L., et al.: Detecting emotional contagion in massive social networks. *PLoS ONE* **9**(3), e90315 (2014)
6. Festinger, L., Pepitone, A., Newcomb, T.: Some consequences of de-individuation in a group. *J. Abnorm. Soc. Psychol.* **47**, 382–389 (1952)
7. Foolen, A.: Word valence and its effects. In: Lüdke, U. (ed.) *Emotion in Language*, pp. 241–256. John Benjamins, Amsterdam (2015)
8. Giles, H., Coupland, N., Coupland, J.: Accommodation theory: communication, context, and consequence. In: Giles, H., Coupland, J. (eds.) *Contexts of Accommodation: Developments in Applied Sociolinguistics*, pp. 1–68 (1991)
9. Hatfield, E., Cacioppo, J.T., Rapson, R.L.: Emotional contagion. *Curr. Dir. Psychol. Sci.* **2**, 96–100 (1993)

10. Hutto, C., Gilbert, E.: Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and Social Media, vol. 8, No. 1 (2014)
11. Kabir, M., Madria, S.: Coronavis: a real-time covid-19 tweets analyzer. [arXiv:2004.13932](https://arxiv.org/abs/2004.13932) (2020)
12. Kiesler, S., Siegel, J., McGuire, T.W.: Social psychological aspects of computer-mediated communication. *Am. Psychol.* **39**, 1123–1134 (1984)
13. Kramer, A.D., Guillory, J.E., Hancock, J.T.: Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci.* **111**(24), 8788–8790 (2014)
14. Kugihara, N.: Effects of aggressive behavior and group size on collective escape in an emergency: a test between a social identity model and deindividuation theory. *Br. J. Soc. Psychol.* **40**(4), 575–598 (2001)
15. Lee, E.J.: When are strong arguments stronger than weak arguments? Deindividuation effects on message elaboration in computer-mediated communication. *Commun. Res.* **35**(5), 646–665 (2008)
16. Lee, E.J.: Effects of visual representation on social influence in computer-mediated communication: experimental tests of the social identity model of deindividuation effects. *Hum. Commun. Res.* **30**(2), 234–259 (2004)
17. Medford, R.J., Saleh, S.N., Sumarsono, A., Perl, T.M., Lehmann, C.U.: An “infodemic”: Leveraging high-volume Twitter data to understand early public sentiment for the coronavirus disease 2019 outbreak. *Open Forum Infect. Dis.* **7**(7), 258 (2020)
18. Melissa, H.: Misinformation about the coronavirus abounds, but correcting it can backfire”. *Los Angeles Times* (2020). <https://www.latimes.com/science/story/2020-02-08/coronavirus-outbreak-false-information-psychology>
19. Memon, S.A., Carley, K.M.: Characterizing covid-19 misinformation communities using a novel twitter dataset. [arXiv:2008.00791](https://arxiv.org/abs/2008.00791) (2020)
20. National Center for Health Statistics: Mental health household pulse survey (2020). <https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm>
21. Paukert, A.L., Pettit, J.W., Amacker, A.: The role of interdependence and perceived similarity in depressed affect contagion. *Behav. Ther.* **39**, 277–285 (2008)
22. Postmes, T., Spears, R., Lea, M.: Breaching or building social boundaries? SIDE-effects of computer-mediated communication. *Commun. Res.* **25**(6), 689–715 (1998)
23. Postmes, T., Spears, R., Lea, M.: The formation of group norms in computer-mediated communication. *Hum. Commun. Res.* **26**, 341–371 (2000)
24. Postmes, T., Spears, R., Sakhel, K., de Groot, D.: social influence in computer-mediated communication: the effects of anonymity on group behavior. *Pers. Soc. Psychol. Bull.* **27**(10), 1243–1254 (2001)
25. Postmes, T., Spears, R.: Deindividuation and antinormative behavior: a meta-analysis. *Psychol. Bull.* **123**(3), 238–259 (1998)
26. Rösner, L., Krämer, N.C.: verbal venting in the social web: effects of anonymity and group norms on aggressive language use in online comments. *Soc. Media + Soc.* **2**(3), 1–13 (2016)
27. Sharma, K., et al.: Coronavirus on social media: analyzing misinformation in Twitter conversations. [arXiv:2003, p. 12309](https://arxiv.org/abs/2003.12309) (2020)
28. Spears, R., Lea, M., Lee, S.: De-individualization and group polarization in computer-mediated communication. *Br. J. Soc. Psychol.* **29**(2), 121–134 (1990)
29. Spears, R., Postmes, T.: Group identity, social influence, and collective action online. In: Sundar, S.S. (ed.) *The Handbook of the Psychology of Communication Technology* (2015)
30. Sun, Y., Chia, S.C., Lu, F., Oktavianus, J.: The battle is on: factors that motivate people to combat anti-vaccine misinformation. *Health Commun.* 1–10 (2020)

31. Tausczik, Y., Faasse, K., Pennebaker, J.W., Petrie, K.J.: Public anxiety and information seeking following the H1N1 outbreak: blogs, newspaper articles, and Wikipedia visits. *Health Commun.* **27**(2), 179–185 (2012)
32. Trautmann, S., et al.: Susceptibility to others' emotions moderates immediate self-reported and biological stress responses to witnessing trauma. *Behav. Res. Ther.* **110**, 55–63 (2012)
33. Turner, J. C.: The analysis of social influence. In: Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D., Wetherell, M.S. (eds.) *Rediscovering The Social Group: A Self-Categorization Theory*, pp. 68–88 (1987)
34. Twitter Inc. Staying safe and informed on Twitter (2021). https://blog.twitter.com/en_us/topics/company/2020/covid-19.html#definition
35. Wheaton, M.G., Prikhidko, A., Messner, G.R.: Is fear of COVID-19 contagious? The effects of emotion contagion and social media use on anxiety in response to the coronavirus pandemic. *Front. Psychol.* **11** (2020)
36. Wild, B., Erb, M., Bartels, M.: Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry Res.* **102**(2), 109–124 (2001)
37. Xie, X.F., Stone, E., Zheng, R., Zhang, R.G.: The 'typhoon eye effect': Determinants of distress during the SARS epidemic. *J. Risk Res.* **14**(9), 1091–1107 (2011)
38. Yao, M.Z., Flanagin, A.J.: A self-awareness approach to computer-mediated communication. *Comput. Hum. Behav.* **22**(3), 518–544 (2004)
39. Zimbardo, P.G.: The human choice: Individuation, reason, and order vs. deindividuation, impulse and chaos. In: Arnold, W.J., Levine, D. (eds.) *Nebraska Symposium on Motivation*, vol. 17, pp. 237–307. University of Nebraska Press, Lincoln, NE (1969)