

EMOTIONALLY TONED ONLINE DISCUSSIONS EVOKE SUBJECTIVELY EXPERIENCED EMOTIONAL RESPONSES

Running head: Online discussions evoke emotions

Aleksi H. Syrjämäki^{a*}, Mirja Ilves^a, Poika Isokoski^a, Joel Kiskola^b, Anna Rantasila^b, Thomas Olsson^b, Gary Bente^c, Veikko Surakka^a

^aResearch Group for Emotions, Sociality, and Computing, Faculty of Information Technology and Communication Sciences, Tampere University, Finland; ^bTechnology and Social Interaction Research Group, Faculty of Information Technology and Communication Sciences, Tampere University, Finland; ^cDepartment of Communication, Michigan State University, East Lansing, United States of America

*Corresponding author. Faculty of Information Technology and Communication Sciences, FI-33014, Tampere University, Tampere, Finland. Tel.: +358 50 477 3784. E-mail: aleksi.syrjamaki@tuni.fi.

Abstract

Emotions are an important element of all social interactions, including the increasingly frequent encounters in social media. Despite the prominent role of written verbal messages in online communication, surprisingly little is known about the effects of emotionally toned messages on the emotions of the readers. This study addresses this gap in three experiments, investigating the effects of emotionally toned comments in online news discussions.

Participants read news discussions with emotionally negative, neutral and positive tone, and then reported their subjectively experienced emotional state using scales for valence and arousal. Results showed that the negatively toned threads induced more negative emotions and higher arousal levels than the other conditions, whereas the positive threads had an opposite effect. Emotionally toned online comments evidently elicit emotional reactions in individuals reading the comments. We discuss implications for future research addressing effects of emotionality on online behavior, as well as mitigation strategies to improve online discussion quality.

Keywords: Computer mediated communication, social media, affect, incivility

Introduction

Expressions of strongly negative emotions, such as hostility, aggression and hate, are a prevalent and persistent problem in social media. Beyond individual harm, such expressions can have adverse consequences for communities, such as increasing polarization of opinions (Anderson, Brossard, Scheufele, Xenos, & Ladwig, 2014), and for online platforms, such as lowering their perceived quality (Prochazka, Weber, & Schweiger, 2018). We suspect that a key cause of these behaviors may be an underlying emotional process within online discussions. Previous research has shown that users experiencing negative emotions tend to post content with a negative tone (Cheng, Bernstein, Danescu-Niculescu-Mizil, & Leskovec, 2017). This content evoking negative emotions in other users could provoke them to post even further emotion-evoking content, possibly leading to a self-reinforcing cycle that proliferates negative emotions among users. The aim of the present research is to contribute to a deeper understanding of this putative process, starting with the question of how the emotional tone of online messages influence the emotions of recipients. By providing insight into emotional effects of online messages, the current study aims to be a step towards understanding causes of uncivil and antisocial online behavior, hopefully paving way for development of interventions to mitigate these problems.

Previous research suggests that exposure to emotional online content increases frequency of corresponding expressions within online interactions. Chmiel et al. (2011) found that emotional tone of comments predicted the tone of subsequent comments in discussion threads. Experimental research has even provided causal evidence that exposure to uncivil or emotional content causes users to post similar content (Kramer, Guillory, & Hancock, 2014; Ziegele, Weber, Quiring, & Breiner, 2018). These findings have been interpreted as indicating that the tone of online messages influence other users' emotions. Importantly,

however, this type of research measuring outward behavior (posts made by the users) provides a look on emotions from only one angle. Emotions are complex phenomena, not only consisting of emotional expressions, but also of various experiential, cognitive and physiological changes (e.g., Izard, 2006). Perhaps self-evidently, an emotional response is associated with a subjective experience. Importantly, however, it is not always possible to infer subjective emotional experience from an individual's behavior, as people tend to, for instance, mimic others' behavior (Chartrand & van Baaren, 2009), and follow social norms (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). Text-based communication in particular, could have a diminished influence on subjectively experienced emotions, as this mode of interaction lacks important emotional information such as facial expressions and speech prosody (Hatfield, Bensman, Thornton, & Rapson, 2014; Neumann & Strack, 2000). It is therefore important to also directly examine effects of online discussions on subjective emotion experiences. This can be done, for instance, by having participants rate their emotional states using dimensions of emotions, such as valence (i.e., pleasure) and arousal (i.e., activation or intensity; Bradley & Lang, 1994).

Earlier research has not yet drawn a clear picture of the subjective emotional responses to emotional online content. One study found that reading discussions on emotionally negative, neutral and positive topics differently influenced subjective emotion reports (Garcia, Kappas, Küster, & Schweitzer, 2016). However, because the negative/neutral/positive threads differed in their topic, it is unclear whether participants responded to the emotionally charged topics, or possibly to the discussers' emotional expressions. Even more relevant for the present research, studies examining effects of tone specifically, have produced mixed results. Two studies found that participants reading uncivil comments reported feeling angrier than participants reading civil comments (Gervais, 2015; Masullo Chen, & Ng, 2017). However, another study found no evidence that the tone of

comments influenced subjective emotion reports (Masullo Chen, & Lu, 2017). Comments, which disagreed with the participants' political views, evoked negative emotions when compared to a control condition with no comments, but the comments with uncivil and civil tone did not differ in their effects on self-reported emotions. Despite this, uncivil tone of the comments, compared to civil tone, increased the likelihood of participants writing an uncivil reply. In a similar vein, Rösner, Winter and Krämer (2016) found that reading uncivil comments increased hostile cognitions when compared to civil comments, but again, had no influence on subjective emotion reports. The current evidence is thus inconclusive as to how the tone of previous posters' messages influence readers' subjective emotional state.

The current study investigated whether the emotional tone of online discussions influence readers' subjective emotions. In three experiments, participants read emotionally toned news comment threads and then rated their subjective emotional experiences of valence and arousal. Three versions of each thread were created, varying in emotional tone (negative/neutral/positive). The main aim was to investigate whether the different threads would evoke corresponding subjective emotional reactions. Experiments 2-3 also investigated additional questions. Experiment 2 investigated psychophysiological responses to the threads (discussed in supplementary materials). Experiment 3 tested whether emotional reactions would be attenuated by affect labeling (naming the emotional content of the stimuli), supposedly by helping participants recognize the emotionally charged content (Torre & Lieberman, 2018). This question will only be discussed briefly in this article.

Method

In three experiments, participants read news stories, followed by an emotionally toned (negative/neutral/positive) discussion thread. Participants then rated the level of valence and arousal experienced while reading the thread.

Participants. For Experiment 1, 36 volunteers were recruited (14 males, 21 females, 1 did not want to disclose gender, $M_{\text{age}} = 30.5$ years, $SD_{\text{years}} = 10.4$). For experiment 2, the aim was to recruit 36 participants, but the data collection was cut short by the COVID-19 outbreak. Because of this, data was only collected from 24 participants. One participant was excluded from the analyses because of inconsistencies in the experimental procedure. Thus, the final sample of Experiment 2 consisted of 23 participants (1 male, 22 females, $M_{\text{age}} = 24.7$ years, $SD_{\text{age}} = 4.0$). Participants were rewarded with partial course credit and a movie ticket. For Experiment 3, 94 volunteers participated (29 male, 59 female, 2 other, 4 did not want to disclose gender, $M_{\text{age}} = 30.5$ years, $SD_{\text{age}} = 9.8$). Eight 20€ gift cards were raffled among the participants. Participants in all experiments indicated informed consent. An ethical statement for the study procedures were obtained from [name of the ethics committee anonymized for peer review].

Stimuli. The stimuli were three fabricated news articles and related discussion threads consisting of five user comments. The topics were abortion, work-related immigration, and legalization of cannabis. We edited three versions of the threads, including comments with negative, neutral and positive tone. The factual content of the three versions was similar, so they only varied in the tone of the comments (e.g., expletives and negatively charged phrases in the negative threads, and very courteous language in the positive threads). To ensure the threads differed in tone as intended, 28 participants rated each thread on the scales of valence and arousal (Bradley & Lang, 1994). Repeated-measures ANOVAs and

follow-up t-tests showed that the ratings differed in the expected direction (lowest valence and highest arousal ratings for negative threads, and highest valence and lowest arousal ratings for positive threads; all $ps < .001$). See Electronic Supplementary Material 1 for the stimulus materials and a detailed description of how the stimuli were validated.

Procedure. In all experiments, participants read the three news stories, followed by the related discussion thread. After reading each thread, participants rated how they felt while reading the thread on a 9-point (-4 ... +4) valence and arousal scales (Bradley & Lang, 1994).

Experiments 1-2 had a within-subjects design. Each participant read three threads, one of each tone and one of each topic. Tone-topic combinations and presentation order were counterbalanced across participants (full counterbalancing was not achieved in Experiment 2 due to the data collection being cut short, see above). In Experiment 1, the stimuli and the measurements were presented on a paper form, which the participants filled in alone. Experiment 2 was run in the laboratory, and the stimuli were presented on a 24" 1440×900 monitor using E-Prime® 2.0 software, and responses were collected using a keyboard. Participants were alone in the laboratory during stimulus presentation and data collection. Physiological measurements (electromyography, skin conductance) were also taken in Experiment 2 (details and results are presented in Electronic Supplementary Material 2). In Experiment 2, the discussion threads were revealed one message at a time, each successive message appearing below the previous message with 15 seconds between messages. There were short breaks before (60s) and after (30s) reading of the news stories to allow participants' physiological state to stabilize.

Experiment 3 was run online using LimeSurvey 4.1.18. In contrast to previous experiments, Emotional tone was manipulated as a between-subjects factor, with pseudo-random allocation to negative and positive groups. Neutral stimuli were not included. As

mentioned in the introduction, Experiment 3 additionally investigated whether the emotional responses would be attenuated by affect labeling, i.e., by naming the emotional tone of the messages. Each of the three negative/positive threads were presented in a different Affect labeling condition, manipulated as a within-subjects factor. The conditions were counterbalanced across participants so that each Affect labeling - discussion topic combination was presented equally often, similarly for both positive and negative tone groups (cell sizes were not perfectly equal in the final sample because not all participants completed the experiment). Presentation order was random. In the Active labeling condition, participants labeled the tone of each message, by pressing one of three radio buttons (negative/neutral/positive) below the message. In the Passive labeling condition, the label was given automatically by showing a text “The tone of this message is mostly negative/positive” (depending on the condition) below each message. In the No labeling condition, the messages were shown without labels. The latter two conditions included a button below each message with the text “I have read the message”. The threads were presented one message at a time, successive messages being revealed below the previous one after the participant pressed the button below the message. In addition to the valence and arousal items, the questionnaire presented after each thread included items regarding interest in the discussion (Garcia et al., 2016), as discussed in Electronic Supplementary Material 2.

Data analysis. The valence and arousal data of Experiments 1-2 were subjected to repeated-measures ANOVAs with Emotional tone (negative/neutral/positive) as an independent factor. Significant effects were followed with pairwise comparisons using paired-samples t-tests. The data of Experiment 3 was analyzed with 2×3 mixed-design ANOVAs with Emotional tone (negative/positive) as a between-subjects factor and Affect labeling (Active labeling/Passive labeling/No labeling) as a within-subjects factor.

Results

See Figure 1 for mean valence and arousal ratings in each condition in all three experiments. On valence ratings, a significant effect of Emotional tone was found in all experiments (Experiment 1: $F(2, 70) = 18.5, p < .001, \eta_p^2 = .35$; Experiment 2: $F(2, 44) = 5.49, p = .007, \eta_p^2 = .20$; Experiment 3: $F(1, 92) = 33.8, p < .001, \eta_p^2 = .27$). Valence ratings were lower in response to negative threads, compared to positive threads in all experiments (Experiment 1: $t(35) = 6.41, p < .001, d = 1.07$, mean difference (MD) 95% CI [1.31, 2.52]; Experiment 2: $t(22) = 2.74, p = .012, d = 0.58$, MD 95% CI [0.30, 2.14]; Experiment 3: MD 95% CI [0.84, 1.71]). Valence ratings were also lower in response to negative threads, compared to neutral threads in Experiment 1 ($t(35) = 3.27, p = .002, d = 0.60$, MD 95% CI [0.42, 1.80]), but not in Experiment 2 ($t(22) = 0.68, p = .503, d = 0.14$, MD 95% CI [-0.54, 1.06]; Experiment 3 did not include the neutral condition). In both Experiments 1 and 2, positive threads evoked higher valence ratings than neutral threads (Experiment 1: $t(35) = 2.61, p = .013, d = .048$, MD 95% CI [0.18, 1.43]; Experiment 2: $t(22) = 2.96, p = .007, d = 0.62$, MD 95% CI [0.29, 1.63]).

On arousal ratings, there was a significant effect of Emotional tone in Experiment 1 ($F(2, 70) = 7.0, p = .002, \eta_p^2 = .17$) and Experiment 3 ($F(1, 92) = 6.29, p = .014, \eta_p^2 = .06$), but not in Experiment 2 ($F(1.6, 35.0) = 3.17, p = .065, \eta_p^2 = .13$; Greenhouse-Geisser corrected). Arousal ratings were higher in response to negative than positive threads in both Experiments 1 and 3 (Experiment 1: $t(35) = 3.36, p = .002, d = 0.58$, MD 95% CI [0.54, 2.18]; Experiment 3: MD 95% CI [0.14, 1.20]). In Experiment 1, arousal ratings were also higher in response to negative than neutral threads ($t(35) = 2.48, p = .018, d = 0.41$, MD 95% CI [1.62, 2.48]), but the difference between neutral and positive threads was not significant ($t(35) = 1.36, p = .181, d = 0.24$, MD 95% CI [-0.23, 1.18]).

In Experiment 3, there were no main effects or interactions involving Affect labeling on either valence or arousal ratings (lowest p was for the main effect of Affect labeling on valence ratings; $F(2, 184) = 2.13, p = .121, \eta_p^2 = .02$).

Discussion

Although the three experiments produced slightly different results, the overall pattern is clear: emotionally toned discussion threads elicited subjectively experienced emotional reactions. In experiment 1, negative threads evoked lower self-reported valence and higher arousal than positive and neutral threads. Positive threads also evoked higher valence than neutral threads. In experiment 2, positive threads evoked higher valence than negative or neutral threads. In contrast to the first experiment, the difference between negative and neutral threads on valence ratings was not significant, and no significant effects on arousal were observed. In experiment 3, however, negative threads evoked both lower valence and higher arousal than positive threads. Overall, these studies suggest that reading negatively toned discussions leads to a slightly more negative and aroused emotional state, whereas positively toned discussions have the opposite effect.

Earlier studies have shown that the emotional tone of online posts influence subsequent users' emotional expressions, which have been interpreted to reflect emotional reactions to the prior posts (Chmiel et al., 2011; Kramer et al., 2014). The current study, together with some earlier studies (Gervais, 2015; Masullo Chen & Ng, 2017), provides direct evidence that the emotional tone of posts indeed influences subjectively reported emotions. This lends credence to the interpretation that emotional expressions in response to prior emotional content indeed reflects emotional reactions, rather than for instance

behavioral mimicry (Chartrand & van Baaren, 2009), or following of social norms within the discussion (Schultz et al., 2007).

These emotional responses could be problematic for many social media platforms and communities. Induction of negative emotions may provoke users to create further negative or even antisocial content (Cheng et al., 2017; Kramer et al., 2014), potentially leading a discussion to a self-reinforcing cycle of negative emotions and expressions. This underlying emotional process may partly explain the prevalence of negative emotional expressions on many social media platforms (Coe, Kenski, & Rains, 2014), which is harmful for the users, the social media site, and perhaps even the wider society (Anderson et al., 2014; Prochazka et al., 2018). Attempting to stop the cycle could be a useful approach to improving discussions and combating some of the negative side-effects of social media.

One proposition for improving online discussions is development of interventions to modulate users' emotions. Seering et al. (2019) found that positivity of messages was increased by introducing user interface elements that evoke positive emotions. Based on the current results, we suggest a similar, alternative approach. Rather than directly manipulating users' emotions, their emotion regulation could be promoted, thereby attenuating emotional responses to online messages. One aim of Experiment 3 was to investigate if labeling the emotional tone of messages would attenuate emotional responses, by helping participants recognize emotionally charged content (cf. Torre & Lieberman, 2018). However, no evidence for this was found. Future work could develop and test more robust methods for facilitating users' emotion regulation during social media use. This could potentially calm down emotionally charged online discussions, providing benefits for social media platforms as well as their users, without directly intervening with users' ability to express themselves, unlike typical interventions by moderators (Ruckenstein & Turunen, 2020).

Funding details

This work was supported by [name of the funding body and grant number anonymized for peer review].

Acknowledgments

The authors have not preregistered this research. The authors thank [anonymized for peer-review] for his help with data analysis.

Research transparency statement

The authors are willing to share their analytics methods and study materials with other researchers. The material will be available upon request. The research data will be archived at [anonymized for peer review] in 2022.

Electronic Supplementary Material

ESM 1. (pdf). Stimulus materials.

ESM 2. (pdf). Additional measurements.

References

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The “nasty effect:” Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, *19*, 373-387. doi:10.1111/jcc4.12009

- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49-59. doi:10.1016/0005-7916(94)90063-9
- Chartrand, T. L., & Van Baaren, R. (2009). Human mimicry. *Advances in Experimental Social Psychology*, 41, 219-274. doi:10.1016/S0065-2601(08)00405-X
- Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll: Causes of trolling behavior in online discussions. Paper presented at the *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1217-1230. doi:10.1145/2998181.2998213
- Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., & Holyst, J. A. (2011). Collective emotions online and their influence on community life. *PLoS One*, 6, e22207. doi:10.1371/journal.pone.0022207
- Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, 64, 658-679. doi:10.1111/jcom.12104
- Garcia, D., Kappas, A., Küster, D., & Schweitzer, F. (2016). The dynamics of emotions in online interaction. *Royal Society Open Science*, 3, 160059. doi:10.1098/rsos.160059
- Gervais, B. T. (2015). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics*, 12, 167-185. doi:10.1080/19331681.2014.997416
- Hatfield, E., Bensman, L., Thornton, P. D., & Rapson, R. L. (2014). New perspectives on emotional contagion: A review of classic and recent research on facial mimicry and

contagion. *Interpersona: An International Journal on Personal Relationships*, 8, 159-179. <https://doi.org/10.5964/ijpr.v8i2.162>

Izard, C. E. (2009). Emotion theory and research: Highlights, unanswered questions, and emerging issues. *Annual Review of Psychology*, 60, 1-25.
doi:10.1146/annurev.psych.60.110707.163539

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 8788-8790.
doi:10.1073/pnas.1320040111

Masullo Chen, G., & Lu, S. (2017). Online political discourse: Exploring differences in effects of civil and uncivil disagreement in news website comments. *Journal of Broadcasting & Electronic Media*, 61, 108-125. doi:10.1080/08838151.2016.1273922

Masullo Chen, G., & Ng, Y. M. M. (2017). Nasty online comments anger you more than me, but nice ones make me as happy as you. *Computers in Human Behavior*, 71, 181-188.
doi:10.1016/j.chb.2017.02.010

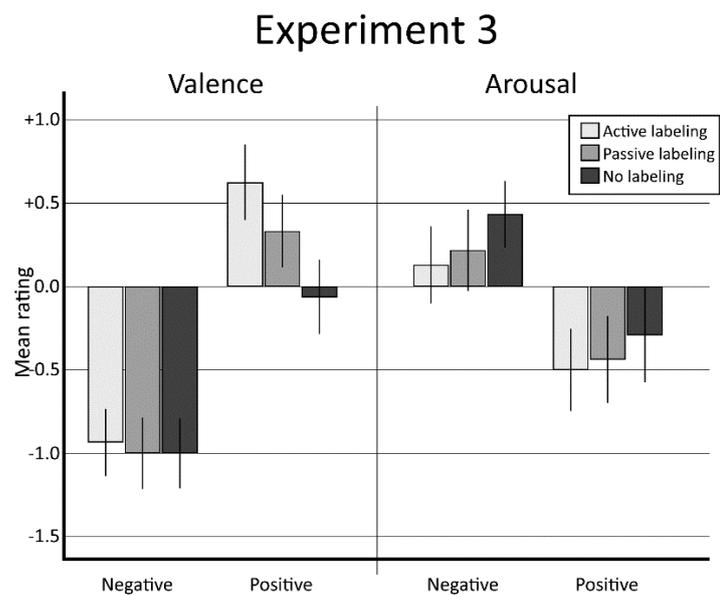
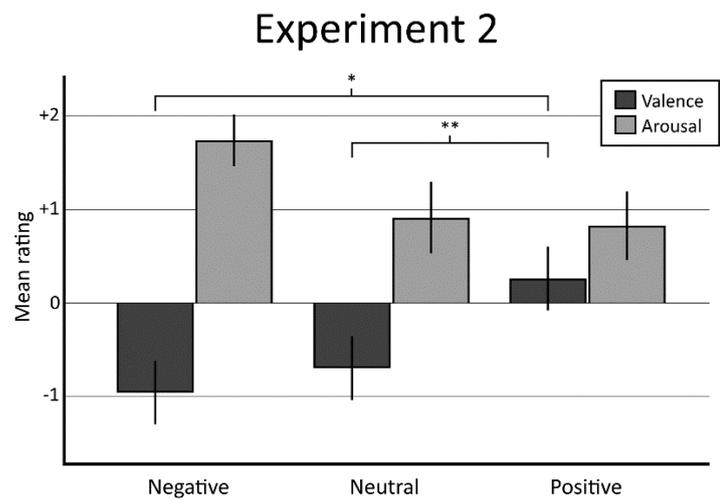
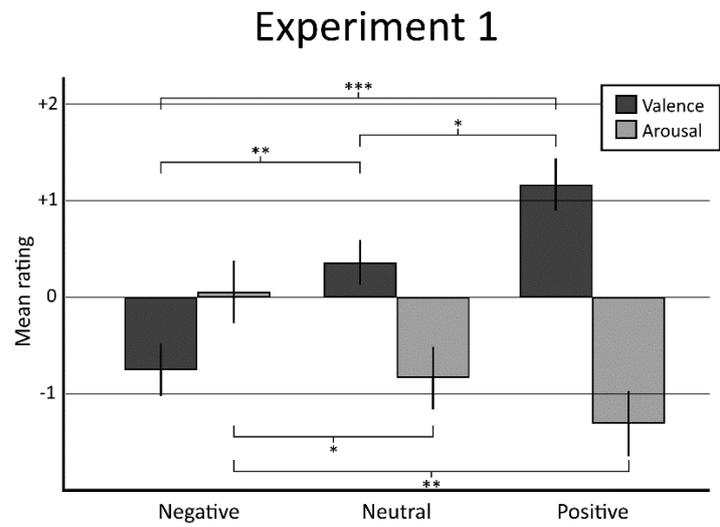
Neumann, R., & Strack, F. (2000). "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, 79, 211-223.
doi:10.1037/0022-3514.79.2.211

Prochazka, F., Weber, P., & Schweiger, W. (2018). Effects of civility and reasoning in user comments on perceived journalistic quality. *Journalism Studies*, 19, 62-78.
doi:10.1080/1461670X.2016.1161497

- Rösner, L., Winter, S., & Krämer, N. C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior, 58*, 461-470. doi:10.1016/j.chb.2016.01.022
- Ruckenstein, M., & Turunen, L. L. M. (2020). Re-humanizing the platform: Content moderators and the logic of care. *New Media & Society, 22*, 1026-1042.
doi:10.1177/1461444819875990
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18*, 429-434. doi:10.1111/j.1467-9280.2007.01917.x
- Seering, J., Fang, T., Damasco, L., Chen, M., Sun, L., & Kaufman, G. (2019). Designing user interface elements to improve the quality and civility of discourse in online commenting behaviors. Paper presented at the *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-14. doi:10.1145/3290605.3300836
- Torre, J. B., & Lieberman, M. D. (2018). Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review, 10*, 116-124.
doi:10.1177/1754073917742706
- Ziegele, M., Weber, M., Quiring, O., & Breiner, T. (2018). The dynamics of online news discussions: Effects of news articles and reader comments on users' involvement, willingness to participate, and the civility of their contributions. *Information, Communication & Society, 21*, 1419-1435. doi:10.1080/1369118X.2017.1324505

Figures

Figure 1:



Legends

Figure 1. Mean valence and arousal ratings in each condition in Experiments 1-3. Error bars denote standard error of the means. Significant differences between conditions, as indicated by paired samples t-tests, are pointed out for Experiments 1-2. $*p < .05$, $**p < .01$, $***p < .001$.