Investigating the Association Between Exposure to (Non-)like-minded Information and Use of Political Incivility on Social Media [*]

Kohei Nishi [†]

Abstract

Does exposure to a like- or non-like-minded information environment lead to political incivility on social media? While a few previous studies have investigated this question, their results have been mixed. There are two conflicting possibilities: (1) if individuals are exposed to a like-minded information environment, their preexisting beliefs are reinforced, and they are more likely to use uncivil expressions, and (2) if individuals are exposed to a non-like-minded information environment, they often feel negative emotions and therefore are more likely to use uncivil expressions. To evaluate these two competing possibilities, the present study collected Japanese political posts from social media X (ex-Twitter) via X API in April 2023 and analyzed them using the number of parties of the politicians that the user followed as a proxy for a non-like-minded information environment. The results of the statistical analysis indicated a very small and non-significant association between a like- or non-like-minded information environment and posting uncivil comments about politics.

[†] Research Fellow of Japan Society for the Promotion of Science and Ph.D. Student at the Division of Law and Political Science, the Graduate School of Law, Kobe University, Japan

Introduction

What makes people uncivil in online political discussions?[1] While the proliferation of social media seems to have made it easier for people to express their political opinions, it has also been observed that some people express their opinions in an uncivil manner or attack their political opponents online. According to Coe et al. (2014), more than one in five comments in news sites' comment sections are uncivil. Uncivil communication hinders consensus building. Hwang et al. (2018) demonstrate that uncivil discussion leads to negative emotions toward the other side and more expressions of disagreement. Popan et al. (2019) have also shown that when discussions are uncivil, individuals perceive lower levels of out-group rationality. Furthermore, several studies have revealed the negative associations between uncivil political communication and citizens' political trust as well as participation (Mutz & Reeves, 2005; Otto et al., 2020; Yamamoto et al., 2020).

Over the last two decades, a number of political communication researchers have focused on the ways in which people are exposed to political information online. Sunstein (2001) argues that on the Internet, individuals can easily form communities with other like-minded persons and are exposed to a large number of like-minded arguments offered by community members, thus reinforcing their preexisting opinions. Findings from some empirical studies support Sunstein (2001)'s argument (Stroud, 2010; Lee & Choi, 2020). However, some evidence has shown that individuals' exposure to information that is contrary to their opinions leads them to have extreme opinions. Taber and Lodge (2006) find that when individuals are exposed to a balanced set of like-minded and non-like-minded arguments, their preexisting attitudes are reinforced because they tend to process non-like-minded information skeptically. A similar tendency has also been observed on X/Twitter; Bail et al. (2018) conducted an experimental study and showed that conservative participants who were randomly assigned to follow a liberal bot account on X/Twitter became more conservative than participants in the control group.

Despite this growing concern about the relationship between exposure to both a like- and non-like-minded information environment and political attitudes, few studies have

examined the relationship between exposure to both types of information environments and the use of incivility. Lee et al. (2019) have found that an increase in the level of cyberbalkanization (the state in which contents are frequently shared within but not across communities) leads to a larger degree of political incivility on Facebook. This means that when individuals are frequently exposed to like-minded information, they are more likely to use uncivil language. This is because discussions with other like-minded persons intensify an individual's preexisting beliefs, which in turn leads to more extreme expressions (Lee et al., 2019).

On the contrary, some studies imply that exposure to non-like-minded information leads to the use of incivility. Hopp and Vargo (2019) have shown that individuals with high levels of bonded social capital are less likely to use political incivility on Facebook, using a combined method of survey and Facebook data. The mechanism for this is that individuals with high levels of bonded social capital are more likely to connect with other like-minded persons and, therefore, do not frequently experience conflicts of opinion (Hopp & Vargo, 2019). Lyu (2023) collected and analyzed the political conversation threads (sets of an original post and replies to it) on X/Twitter and found that cross-cutting interactions on X/Twitter were more likely to be offensive than in-group interaction. Furthermore, Vargo and Hopp (2017) report that low levels of partisan polarity (i.e., high levels of partisan conflict) or high levels of racial heterogeneity in districts are positively correlated with the use of incivility on X/Twitter.

Although these previous studies have made significant contributions to the literature, few studies have directly observed the ways in which individuals are exposed to information on social media to investigate the association between exposure to (non-)like-minded information environment and use of incivility. To overcome the issue and provide new findings, the present study collected data from Japanese X/Twitter data via X/Twitter API and analyzed them using the number of parties of the politicians that the user followed as a proxy for a non-like-minded information environment. The results of the statistical analysis indicated a very small and non-significant association between a like- or non-like-minded information environment and posting uncivil comments about politics.

Hypotheses

The first expectation is that exposure to a homogeneous information environment is positively associated with posting uncivil comments. According to previous studies, individuals reinforce their original opinions through selective exposure to homogeneous and like-minded information (Lee & Choi, 2020; Stroud, 2010; Sunstein, 2001). When individuals' opinions become more extreme, they perceive greater distance between their own opinions and the out-group's opinion, and thus, they might be more likely to perceive the out-group as a threat. When individuals perceive the out-group as a threat, they attack it to protect the in-group from the threat (Böhm et al., 2016). Thus, a positive association is expected between exposure to a homogeneous information environment and posting uncivil comments (i.e., a negative association is expected between exposure to a heterogeneous information environment and posting uncivil comments). Suhay et al. (2015) have shown that individuals with extreme opinions are more likely to use uncivil expressions, which indirectly supports the aforementioned mechanism.

Hypothesis 1: Those who are exposed to a more heterogeneous information environment are less likely to post uncivil comments about politics on social media.

On the contrary, the alternative expectation is that exposure to a heterogeneous information environment is positively associated with posting uncivil comments. Individuals who are more exposed to non-like-minded information might have more opportunities to experience opinion conflicts with out-groups, and thus, they might more frequently perceive out-group threats. As mentioned above, perceiving the out-group as a threat motivates individuals to attack it (Böhm et al., 2016). Thus, this time, a positive association is expected between exposure to a homogeneous information environment and posting uncivil comments.

Hypothesis 2: Those who are exposed to a more heterogeneous information environment

are more likely to post uncivil comments about politics on social media.

These two conflicting expectations are both theoretically plausible. Therefore, the present study adopts both of them as hypotheses and aims to clarify if either or neither one is correct through an empirical analysis.

Methods

Collecting Social Media Data

First, the present study collected Japanese political posts on social media X (ex Twitter) using X/Twitter API. More specifically, posts that contained the name of a political party or its leader and were posted during April 3 (Monday) and 9 (Sunday), 2023, were collected (280,628 posts).

From the collected posts, 3,000 users (hereinafter referred to as target users) were randomly selected, and their lists of followees (other users that the target user follows), count of past posts (the number of posts that the target user posted in the past), and past posts created by the target user within the first 30 days of creating their accounts were collected using X/Twitter API for Academic Research (74,104 posts in total).

Then, posts that contained the name of a political party or its leader and were posted during April 10 (Monday) and 16 (Sunday), 2023, were collected, and posts authored by the target users were extracted from them (7,537 posts). The reason why only posts that contained the name of a political party or its leader were collected was that the present study aimed to collect posts about politics.

Dependent Variable

The dependent variable was a dummy variable for posting uncivil comments. More specifically, it is a binary variable that takes a value of 1 if the user has posted one or more uncivil posts about politics during April 10 (Monday) and 16 (Sunday), 2023, and

otherwise, it takes a value of 0. The present study classified the collected posts as uncivil or not using Google Jigsaw's Perspective API, which is a machine learning-based API trained to detect toxic texts (Google Jigsaw, 2024). As its definition of toxic is very close to the definition of incivility in the present study, texts classified as toxic are regarded as uncivil herein. The threshold was set to 0.6. The doNotStore option was used to ensure that the texts would not be stored and used for training by Jigsaw. Consequently, 775 out of 7,537 posts (10.28%) were classified as uncivil.

Independent Variable

The independent variable was the number of parties of the politicians that the target user followed.[2] The present study assumes that this is a good proxy for the degree to which a user is exposed to a politically heterogeneous information environment. Some politicians had no party affiliations, and for convenience, such politicians were treated as belonging to a party labeled "independent." As 10 parties had seats in the Japanese parliament at the time of the data collection, the independent variable theoretically ranged from 1 to 11 (10 parties + independent). Those who did not follow politicians were excluded from the dataset because the independent variable could not be computed for them (N = 918 out of 3,000).

Control Variables

As the present study uses observational data at one time point, it is difficult to infer causal relationships from the data; thus, it aims to investigate the association instead of causal relationships. However, to remove endogeneity bias as much as possible, the present study controlled some variables.

To remove the bias due to the target users' initial level of incivility when they created the account being associated with both independent and dependent variables, a dummy variable for posting uncivil comments during the first 30 days of creating the X/Twitter account was controlled. The variable was computed by applying the same approach as that

of the dependent variable. Consequently, 1,456 out of 74,104 posts (1.96%) were classified as uncivil.

In addition, to remove bias due to the target users' activity levels on X being associated with both independent and dependent variables, the count of followees (the number of other users that the target user follows), count of past posts (the number of posts that the target user posted in the past), and count of politician followees (the number of politicians that the target user follows) were controlled.

Statistical Analysis

A logistic regression model was estimated using the variables described above.[3]

Results

The results of the statistical analysis indicated a very small and non-significant association between exposure to a heterogeneous information environment and the use of political incivility. The results of the estimation of the logistic regression model are shown in Table 1. As shown in the table, the coefficient of the independent variable is 0.002, and it is not statistically significant at the 5% level ($p = .961$). These results do not support either Hypothesis 1 or 2. For easy interpretation, the predicted probabilities of posting uncivil comments conditioned by the independent variable were plotted with their 95% confidence intervals in Figure 1.[4] It suggests a negligible association between the number of parties of the politicians that the target user follows and the probability of posting uncivil comments.
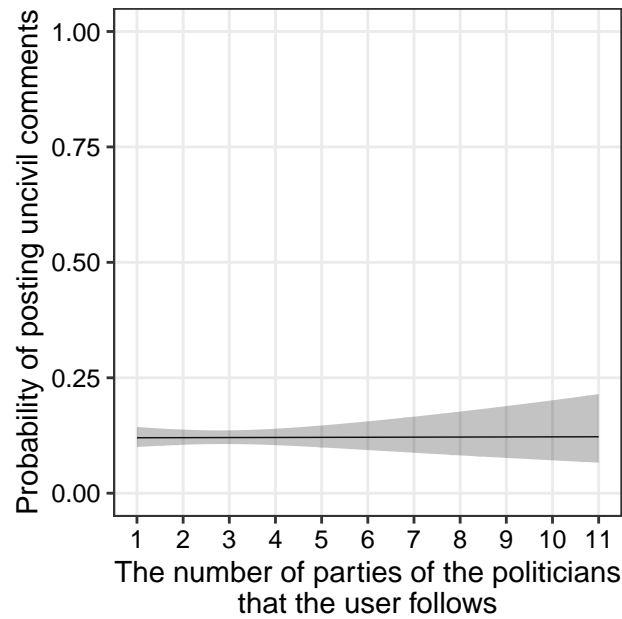
Discussion

The question of whether exposure to a like- or non-like-minded information environment on social media leads to political incivility has not been adequately studied, with previous studies suggesting conflicting answers. The present study collected and analyzed data from

Table 1: Results of Logistic Regression

	Coefficient
Number of parties of the politicians	0.002
that the target user follows	*SE* = 0.041
	p = .961
Early uncivil posts	0.955
	SE = 0.176
Count of followees	0.000
	SE = 0.000
Count of past posts	0.000
	SE = 0.000
Count of politician followees	0.006
	SE = 0.004
Intercept	-2.070
	SE = 0.121
N of observations	2,082

The dependent variable is the dummy variable for posting
uncivil comments.

Figure 1: Predicted Probability of Posting Uncivil Comments

Japanese X/Twitter data. The results of the statistical analysis indicated a very small and non-significant association between exposure to a heterogeneous information environment and posting uncivil comments. One possible reason for these unexpected results is that the two mechanisms (H1 and H2) cancel each other out, resulting in only a very small and non-significant association. However, this cannot be confirmed by the data in the present study, and additional research is needed.

The present study makes important contributions to the literature. First, while two conflicting hypotheses were set, the results of the statistical analysis do not support either. This suggests that the relationship between exposure to a (non-)like-minded information environment and using incivility may be more complex than expected, and further research is required. Second, the present study investigates the association between exposure to a (non-)like-minded information environment and use of incivility by directly observing the ways in which individuals are exposed to information on social media. Third, it expands the regional scope of the study of political incivility, as the number of the study of online political incivility focusing on Japan is limited.

The present study has some limitations. First, it is unclear whether similar results can be found in other languages, cultures, and social media platforms. Second, as the present study uses observational data at one time point, it is difficult to infer causal relationships from the data. Therefore, further research is required to overcome these issues in the future.

Notes

[1] The present study defines political incivility as "a disrespectful or insulting expression that attacks an individual or group in political communication."

[2] Here, politicians mean members of parliament in Japan. The usernames of the X/Twitter accounts of politicians were collected from parties' official websites. For politicians with no party affiliations, usernames were collected from politicians' websites or by searching in a search engine.

[3] Data process was conducted using Python version 3.10.12 (Python Software Foundation, 2024) and R version 4.4.0 (R Core Team, 2024). The pandas package (The Pandas Development Team, 2024) in Python and the tidyverse package (Wickham et al., 2019) in R were used for data wrangling. A logistic regression model was estimated using the glm function in R.

[4] The predictions function in the marginaleffects package (Arel-Bundock, 2024) and the ggplot function in the tidyverse package (Wickham et al., 2019) in R were used for the visualization. When

computing predicted values, control variables were set to mean (in case of a quantitative variable) or mode (in case of a dummy variable) values.

Funding

References

Arel-Bundock, V. (2024). *marginaleffects: Predictions, comparisons, slopes, marginal means, and hypothesis tests*. https://marginaleffects.com

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., . . . Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(37), 9216–9221. https://doi.org/10.1073/pnas.1804840115

Böhm, R., Rusch, H., & Gürerk, Ö. (2016). What makes people go to war? Defensive intentions motivate retaliatory and preemptive intergroup aggression. *Evolution and Human Behavior*, *37*(1), 29–34. https://doi.org/10.1016/j.evolhumbehav.2015.06.005

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *The Journal of Communication*, *64*(4), 658–679. https://doi.org/10.1111/jcom.12104

Google Jigsaw. (2024). *Perspective API*. https://developers.perspectiveapi.com

Hopp, T., & Vargo, C. J. (2019). Social capital as an inhibitor of online political incivility: An analysis of behavioral patterns among politically active Facebook users. *International Journal of Communication*, *13*(0), 4313-4333. https://ijoc.org/index.php/ijoc/article/view/10523

Hwang, H., Kim, Y., & Kim, Y. (2018). Influence of discussion incivility on deliberation: An examination of the mediating role of moral indignation. *Communication Research*, *45*(2), 213–240. https://doi.org/10.1177/0093650215616861

Lee, F. L. F., Liang, H., & Tang, G. K. Y. (2019). Online incivility, cyberbalkanization, and the dynamics of opinion polarization during and after a mass protest event. *International Journal of Communication*, *13*(0), 4940-4959. https://ijoc.org/index.php/ijoc/article/view/11666

Lee, J., & Choi, Y. (2020). Effects of network heterogeneity on social media on opinion polarization among South Koreans: Focusing on fear and political orientation. *International Communication Gazette*, *82*(2), 119–139. https://doi.org/10.1177/1748048518820499

Lyu, Z. (2023). Cross-cutting interaction, inter-party hostility, and partisan identity: Analysis of offensive speech in social media. *New Media & Society*. https://doi.org/10.1177/14614448231180654

Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, *99*(1), 1–15. https://doi.org/10.1017/s0003055405051452

Otto, L. P., Lecheler, S., & Schuck, A. R. T. (2019). Is context the key? The (non-)differential effects of mediated incivility in three European countries. *Political Communication*, 1–20. https://doi.org/10.1080/10584609.2019.1663324

Popan, J. R., Coursey, L., Acosta, J., & Kenworthy, J. (2019). Testing the effects of incivility during internet political discussion on perceptions of rational argument and evaluations of a political outgroup. *Computers in Human Behavior*, *96*, 123–132. https://doi.org/10.1016/j.chb.2019.02.017

Python Software Foundation. (2024). *The Python language reference*. https://docs.python.org/3/reference/index.html

R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Stroud, N. J. (2010). Polarization and partisan selective exposure. *The Journal of Communication*, *60*(3), 556–576. https://doi.org/10.1111/j.1460-2466.2010.01497.x

Suhay, E., Blackwell, A., Roche, C., & Bruggeman, L. (2015). Forging bonds and burning bridges: Polarization and incivility in blog discussions about Occupy Wall Street *American Politics Research*, *43*(4), 643–679. https://doi.org/10.1177/1532673x14553834

Sunstein, C. R. (2001). *Republic.com*. Princeton, NJ: Princeton University Press.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

The Pandas Development Team. (2024). *pandas-dev/pandas: Pandas*. https://doi.org/10.5281/ZENODO.3509134

Vargo, C. J., & Hopp, T. (2017). Socioeconomic status, social capital, and partisan polarity as predictors of political incivility on Twitter. *Social Science Computer Review*, *35*(1), 10–32. https://doi.org/10.1177/0894439315602858

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Yamamoto, M., Dalisay, F., & Kushin, M. J. (2020). An examination of uncivil and reasoned comments and perceived civility in politics. *International Journal of Communication*, *14*(0), 279-298. https://ijoc.org/index.php/ijoc/article/view/12036