# The Effect of Exposure to (Non-)Like-Minded Information on the Use of Political Incivility on Twitter

Kohei Nishi*

*Graduate School of Law, Kobe University

koheinishi46@gmail.com

**Abstract**

Does exposure to like-minded/non-like-minded information lead to the use of political incivility? Few studies have investigated this question, and the results have been mixed. There are two conflicting possibilities: (i) if individuals are frequently exposed to like-minded political information, they reinforce their pre-existing beliefs and are, thus, more likely to use uncivil language, and (ii) if individuals are frequently exposed to non-like-minded information, they often feel negative emotions and are, therefore, more likely to use incivility. To evaluate these two competing hypotheses, I analyze Japanese Twitter data using a semi-supervised learning method. The results show that individuals who are exposed to non-like-minded information are more likely to use political incivility.

# Introduction

## Political Incivility

What makes people uncivil[1] in online political discussions? While it has become much easier for people to express their political opinions due to the proliferation of social network services, it has also been observed that people tend to express their opinions in an uncivil manner or attack their political opponents online.[2] Since respect for each other is essential in a democracy, this is a critical problem. The empirical literature has shown that uncivil communication hinders consensus building. For example, Hwang et al. (2018) demonstrated that uncivil discussion leads to negative emotions toward the other side and more expressions of disagreement. Popan et al. (2019) has also shown that when discussions are uncivil, individuals perceive lower levels of out-group rationality. Furthermore, several studies have revealed that uncivil political communication has direct and indirect negative effects on citizens' political trust and participation (Mutz & Reeves, 2005; Otto et al., 2020; Yamamoto et al., 2020).

In recent years, a large body of literature has investigated political incivility (for example, Coe et al., 2014; Sobieraj & Berry, 2011), and some

[1] Coe et al. (2014) defined incivility as "features of discussion that convey an unnecessarily disrespectful tone toward the discussion forum, its participants, or its topics."

[2] According to Coe et al. (2014), more than one in five comments in online discussions are uncivil.

researchers have focused on platform design. For example, Sydnor (2018) have shown that levels of perception of incivility depend on the type of platform (e.g., audio, video, and text). Maia & Rezende (2016) and Rowe (2015) have found that individuals more frequently use incivility on platforms that permit anonymity than on the ones that identify users. Oz et al. (2018) has investigated the difference in the level of incivility between Twitter and Facebook users. Both platforms have different features in terms of anonymity and character limits. Similar to these studies, the present study also focuses on the features of platforms. Specifically, I focus on users' ability to choose who they follow on Twitter.

## Exposure to (Non-)Like-Minded Information

In an environment where people have the freedom to choose their information sources, they tend to select information that reinforces their pre-existing beliefs. For instance, right-leaning individuals frequently read right-leaning blogs, and do not often read left-leaning blogs. This is true for left-leaning individuals as well. Such tendency is called selective exposure, a subject that has been studied for a long time (for example, Festinger, 1957). Sunstein (2001) has argued that after the emergence of the Internet, individuals can more easily find information which confirms their personal views. This is because there are many options available online. Empirical studies have found that some people tend to select information sources, on the Internet, according to their political attitudes (for example, Garrett, 2009; Stroud, 2008). The same tendency has been observed on Twitter (Vaccari et al., 2016).

This causes individuals' opinions to become more extreme. Sunstein (2001) argues as follows: 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 pre-existing opinions. Empirical research has also found that partisan-selective exposure leads to more extreme attitudes (Stroud, 2010). Similarly, network heterogeneity on social media is negatively correlated with opinion polarization (i.e., network homogeneity on social media is positively correlated with opinion polarization) (J. Lee & Choi, 2020).

However, some evidence has shown that exposure to non-like-minded information leads to extreme opinions. When individuals are exposed to a balanced set of like-minded and non-like-minded arguments, they reinforce their pre-existing attitudes (Taber & Lodge, 2006). This is because they skeptically process counterattitudinal information, while uncritically accepting information that supports their pre-existing opinions (Taber & Lodge, 2006). Similarly, experimental research has shown that conservative participants who were randomly assigned to follow a liberal Twitter bot became more conservative (Bail et al., 2018). These findings suggest that exposing individuals to non-like-minded information might not solve the problem of polarization caused by selective exposure.

As mentioned earlier, there are growing concerns about the relationship between exposure to (non-)like-minded information and political attitudes. However, to my knowledge, there are few studies that have investigated the relationship between exposure to (non-)like-minded information and the use of political incivility.

# Previous Studies

## Like-Minded Information and Incivility

A few studies have previously investigated the relationship between exposure to like-minded information and the use of incivility. F. L. Lee et al. (2019) have found that an increase in the level of cyberbalkanization (the state in which contents are frequently shared within communities, 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 pre-existing beliefs, which in turn leads to more extreme expressions (F. L. Lee et al., 2019).

Although Lee and colleagues' work has made a significant contribution to this area of research, there is one problem: the unit of analysis is not the individual. This may lead to ecological fallacy problems. In other words, even though the degree of cyberbalkanization is correlated with the level of incivility at the collective level, this is not necessarily true at the individual level. Individual-level analysis is essential to understand political incivility because the phenomenon of incivility is caused by individuals' decision-making through their cognitive processes.

Another study found that content shared within communities is more uncivil than content shared across communities on Facebook (Chan et al., 2019). This finding suggests a relationship between like-minded ties and the use of incivility. However, this does not directly answer the question of the present study, because the finding does not indicate a difference in the degree of incivility between individuals who are exposed to like-minded and non-like-minded information. Instead, it merely indicates that content that is more moderate in expression is more likely to be accepted by a wider audience.

## Non-Like-Minded Information and Incivility

A few prior studies have suggested that there is a relationship between exposure to non-like-minded information and the use of incivility. Although these studies do not directly answer the question of the present study, it is worth paying attention to them because they suggest a mechanism that is contradictory to the findings introduced in the previous section.

Hopp & Vargo (2019) have shown that individuals with high levels of bonded social capital are less likely to use political incivility on Facebook. Hopp & Vargo (2019) have argued that this is because 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 opinions. Another study found that low levels of partisan polarity (high levels of partisan conflict) or high levels of racial heterogeneity in districts are positively correlated with the use of incivility on Twitter (Vargo & Hopp, 2017). Furthermore, Rossini (2020) investigated the comment section of news websites and a Facebook page in Brazil. The results show that high levels of disagreement are associated with the use of incivility in the comment sections.

These studies have suggested that individuals are more likely to be uncivil in a non-like-minded environment because there are more conflicts of opinion and, thus, more opportunities for aggression. This mechanism conflicts with Lee and colleagues' mechanism, which was introduced in the previous section. Lee and colleagues mentioned that both of these competing mechanisms are theoretically possible (F. L. Lee et al., 2019). However, although these studies are thought-provoking, they do not directly answer my questions. This is because they have not directly focused on the way individuals are exposed to information on the Internet, where they have the freedom to choose their information sources.

Another study investigated political incivility on five platforms in Brazil, and the results showed that there are more uncivil comments on heterogeneous platforms than on homogeneous ones (Maia & Rezende, 2016). This is an important finding, suggesting that exposure to non-like-minded information leads to the use of incivility. However, this finding is not sufficient to answer the question because the unit of analysis is a platform instead of an individual. This raises two problems. First, ecological fallacy problems may occur. Second, endogeneity may exist. It is possible that platforms created for the purpose of active discussion are more likely to attract people with diverse opinions, and discussions are more likely to be intense and uncivil. These problems are solved by analyzing individuals on a single platform.

## Previous Studies' Limitations

As discussed above, few previous studies have investigated the impact of exposure to like-minded and non-like-minded information on the use of incivility on the Internet. There are three main problems with the previous studies. First, previous studies have suggested two conflicting mechanisms. Second, the analysis at the individual level was insufficient. Third, few studies have directly focused on the ways in which individuals are exposed to political information on the Internet, where they have the freedom to choose their information sources from a wide range of options. The present study aims to overcome these issues and provide new findings.

# Theories and Hypotheses

## Like-Minded-Information and Polarization

This section describes the mechanism by which exposure to like-minded information leads to the use of incivility. As mentioned earlier, individuals reinforce their original opinions through selective exposure to like-minded information (Sunstein, 2001; Stroud, 2010). The more extreme an individual's views become, the greater the distance between their views and those of the out-group, and, thus, the more likely they are to perceive the out-group as a threat. As a result, they attack the out-group to protect the in-group from the threat (Böhm et al., 2016). An empirical study has shown that individuals with extreme opinions are more likely to use uncivil expressions (Suhay et al., 2015). If these mechanisms are correct, it is expected that individuals are more likely to post uncivil comments when exposed to like-minded information.

**H1: Internet users are more likely to use political incivility when they are exposed to like-minded information.**

## Non-Like-Minded Information and Backlash

This section explains the mechanism by which exposure to non-like-minded information leads to incivility. The key point here is that individuals who are exposed to non-like-minded information have more opportunities to experience conflicts with out-groups. When individuals perceive a threat from out-groups, they attack the out-groups to protect the in-groups (Böhm et al., 2016). Based on this finding, the more opportunities people have to be exposed to non-like-minded information posted by out-group members, the more frequently they perceive out-group threats, and, thus, the more likely people are to post uncivil comments. If such a mechanism is correct, then I can predict that exposure to non-like-minded information leads to the use of incivility.

**H2: Internet users are more likely to use political incivility when they are exposed to non-like-minded information.**

These two hypotheses are conflicting, but both are theoretically plausible. Therefore, the present study adopts both of them as hypotheses and clarifies which one is correct (or that neither is correct) through empirical analysis.

# Method

## Data

I collected political tweets in Japan over a period of eight weeks.[3] The collection was conducted using Twitter API. Specifically, I collected Japanese tweets containing the name of a political party or its leader, randomly extracting 500 tweets per day (that is, I collected a total of 28,000 tweets). Only extracted tweets were used for the analysis. This is because Twitter API limits the number of requests that can be made per hour, and it was, thus, not possible to obtain the necessary information described below for all the collected tweets. For each tweet, I collected the followee list of the user who posted the tweet, his/her Twitter bio text, his/her followees' Twitter bio texts, and old tweets posted by him/her within the first 30 days of opening his or her account. I collected old tweets only for users whose total number of tweets was less than 3,200 at the time of data collection. This is because I could collect only the latest 3,200 tweets of any given user due to the limitations of the Twitter API. In addition, I collected a list of parliament members' Twitter accounts to conduct the indexing described later (Kokkaigiin Ichiran List, n.d.-a,-b). All data collection and indexing were performed using Python.

## Dependent Variable

The dependent variable was the level of incivility of the tweet. I used Latent Semantic Scaling (LSS) (Watanabe, 2020), a semi-supervised learning method, to index this variable. The LSS evaluates the polarity of a text using the cosine similarity between the vector of seed words that represent each

[3]I collected data over eight weeks, from November 2, 2020, to December 27, 2020.

polarity and the vector of words in the text. According to Watanabe, if we select the appropriate seed words for a given purpose, the LSS can evaluate a text on the desired dimension. Thus, by selecting typical uncivil and civil words as seed words, LSS was applied to evaluate the level of incivility of tweets.

A word vector is a vector that represents the meaning of a word in approximately 200 dimensions. This vector is calculated using a technique called word embedding. There are several methods for word embedding. According to Watanabe, who uses one of them, Latent Semantic Analysis, it is possible to use other methods as well. Therefore, I used Word2Vec (Mikolov et al., 2013).

The procedure for evaluating the level of incivility of a tweet is as follows: First, I obtained a pre-trained Word2Vec model, hottoSNS-w2v (Matsuno et al., 2019), created by Hotto Link Inc. Matsuno et al. trained the Word2Vec model using a Japanese corpus built from the text of blogs, Twitter, Japanese Wikipedia, and other web pages.

Next, I preprocessed the text. Using Mecab (Kudo, n.d.) and mecab-ipadic-NEologd (NEologd, n.d.), I conducted a morphological analysis of the tweets. Words without substantive meanings were excluded. I excluded parts of speech other than nouns, verbs, adjectives, and adjectival verbs. I also excluded numbers, URLs, and words included in SlothLib (SlothLib, n.d.) and Marimo (Watanabe, n.d.), which are the famous Japanese stop word lists. Additionally, I excluded the names of political parties and party leaders, as these words are not necessary for assessing the level of incivility.

I then prepared seven uncivil and civil seed words each (see Table 1). The level of the incivility of a word was obtained by summing the cosine similarity of the vector of the word and the seven uncivil seed words, summing the cosine similarity of the vector of the word and the seven civil seed words multiplied by -1, and dividing the sum by 14. The level of incivility of a tweet was obtained by calculating the level for all the words in the tweet using the aforementioned method and dividing the sum by the number of words (for details, see Watanabe (2020)).

$$\begin{aligned} incivlity\ of\ a\ word &= \tfrac{1}{14}\{\sum_{i=1}^{7} cos\_sim(uncivil\ seed\ word_i, word) \\ &\quad + \sum_{i=1}^{7} cos\_sim(civil\ seed\ word_i, word) \times (-1)\} \\ incivility\ of\ a\ tweet &= \tfrac{1}{n} \sum_{i=1}^{n} incivility\ of\ word_i \end{aligned}$$

Table 1: Seed Words

| Polarity | Seed Words |
|---|---|
| Univil | バカ, アホ, 無能, 死ね, キモい, 糞, 売国奴 |
| Civil | 聡明, 優秀, 有能, 応援, ありがとう, 素敵, 誠実 |

## Independent Variable (Index 1)

The independent variable is the degree of like-mindedness of the information to which the individuals, who posted the tweets, were exposed on Twitter. I measured this by focusing on who they were following. When individuals use Twitter, their Twitter home timelines show the tweets posted by other users that they follow. Thus, if they follow only others with the same political views as themselves, their home timelines show only tweets that express the same political views as themselves. However, if they follow others with diverse political opinions, their home timelines show a variety of political opinions. Therefore, to know the degree of like-mindedness of the information that they were exposed to on Twitter, I focused on the degree of like-mindedness of their Twitter followees.

Specifically, I used the semantic similarity between their Twitter bios and their followees' ones as a measure of like-mindedness. When both Twitter bios are semantically similar, they follow others who are similar to themselves, and they are, thus, exposed to like-minded tweets when looking at their Twitter home timelines.

For indexing, I used Word2Vec. As mentioned above, by using this, we can obtain the vectors that represent the meaning of words. In addition, by averaging the vectors of the words in the text, we can obtain a vector that represents the meaning of the text. Furthermore, by calculating the cosine similarity of the two vectors representing the two texts, we can evaluate their semantic similarity.

The specific procedure for indexing like-mindedness using Word2vec is described below. First, I obtained a pre-trained model, hottoSNS-w2v (Matsuno et al., 2019), and preprocessed the texts. I then found the average of the vectors of words in the Twitter bio of the users who posted the tweets in my data to obtain a vector of their Twitter bio. Next, I averaged the vectors of words in the Twitter bio of their followees to obtain the vector of their Twitter bio. I then calculated the cosine similarity between the vector of Twitter bio of the user who posted the tweet and the vector of Twitter bio of their followees, summed them, and divided it by the number of followees. However, users and followees who did not write their Twitter bios were excluded from the dataset.

## Independent Variable (Index 2)

To increase robustness, I created another measure of like-mindedness. This measure focused on MPs among the followees and was calculated using the following formula: If he/she followed more conservative MPs than liberal ones, I used formula A. If he/she followed more liberal MPs than conservative ones, I used formula B. If he/she followed an equal number of conservative and liberal MPs, I gave him/her a value of 0.5.

$$\text{A: } \textit{like-mindedness (2)} = \frac{\textit{the number of conservative MPs whom he/she follows}}{\textit{the number of MPs whom he/she follows}}$$

$$\text{B: } \textit{like-mindedness (2)} = \frac{\textit{the number of liberal MPs whom he/she follows}}{\textit{the number of MPs whom he/she follows}}$$

It is important to note that I assumed that most people follow more MPs who have the same ideology as them than MPs who do not. If he/she followed only MPs who shared the same ideology as him/her, the indicator was close to 1. In contrast, when he/she followed MPs on both sides in a balanced manner, the indicator was close to 0.5. The criteria for labeling MPs as conservative or liberal are listed in Table 2.

Table 2: The criteria for labeling MPs as conservative or liberal

| Label | Criteria |
|---|---|
| Conservative | MP who belongs to the Liberal Democratic Party, Komeito or Japan Innovation Party |
| Liberal | MP who belongs to Constitutional Democratic Party, Japanese Communist Party, Social Democratic Party or Reiwa Shinsengumi |
| Excluded from the dataset | MP who belongs to Democratic Party For the People or The Party to Protect Citizens from NHK |

Note: Since the Democratic Party For the People's platform states that it aims to be "a citizens-driven, reforming centrist party that encompasses both moderate conservatives and liberals," it is difficult to classify it as either left or right. In addition, it is difficult to place the Party to Protect Citizens from NHK on the left-right scale, because it is a single-issue party and its issue is not ideological.

I excluded, from the dataset, those who did not follow any MPs because it was impossible to calculate. I also excluded those who followed only one legislator from the dataset because it was highly likely that they did so only because the MP was famous.

The differences between indexes 1 and 2 are as follows: The former provides a more accurate measure of general-level like-mindedness. This is because the former uses information from almost all the followees. However, the former has the disadvantage that the like-mindedness it measures may not necessarily be like-mindedness in a political sense, because Twitter bios may contain information that is not related to politics. The latter compensates for the disadvantages of the former. This is because the ideology of MPs can be estimated accurately by their party affiliations. Conversely, the latter does not accurately reflect the actual like-mindedness of their timelines because it ignores the information of non-MP followees. Thus, the first and second indicators reflect different aspects of the concept of like-mindedness.

## Control Variables

To address the possibility of reverse causality, where an originally aggressive person follows a non-like-minded person to attack those who have different opinions from their own, I controlled for the initial level of incivility when they first opened their Twitter accounts. As the initial incivility level indicator, I used the level of incivility of tweets posted during the first 30 days of opening their Twitter account.

I also controlled for the duration of Twitter use. This was to address the possibility that the duration of Twitter use might affect both their tendencies to follow and levels of incivility.

In addition, I controlled for the number of followees and the number of MPs they followed. Since the number of Twitter users who post political tweets is fixed and the number of MPs is also fixed, the more Twitter users they follow, or the more MPs they follow, their like-mindedness score is likely to be lower. In addition, the more people they follow, the more information they are exposed to, and, thus, under the assumptions of hypothesis 1 or 2, their level of incivility is likely to be higher. Therefore, it is possible that the number of followees and the number of MPs they follow are confounding factors.

# Results

Prior to the analysis, if there were multiple tweets posted by the same user in the dataset, the values were averaged and aggregated into one row. This was done to avoid such users' data being overly reflected in the estimates.

After that, I conducted multiple regression analyses using the data and variables described above. In addition to the analyses using a normal dataset, I also conducted regression analyses using only those users whose initial levels of incivility were below average (initially not uncivil dataset). This is because I was interested in discovering whether originally not uncivil people could become uncivil through exposure to like-minded or non-like-minded information on Twitter.

Table 3: Descriptive Statistics

| | Mean | SD | Median | Min | Max |
|---|---|---|---|---|---|
| Like-mindedness (1) | 0.210 | 0.119 | 0.212 | -0.298 | 1.000 |
| Like-mindedness (2) | 0.903 | 0.144 | 1.000 | 0.500 | 1.000 |
| Level of incivility | 0.042 | 0.030 | 0.042 | -0.183 | 0.301 |
| Initial level of incivility | 0.024 | 0.022 | 0.026 | -0.146 | 0.120 |
| Followee count | 1224.324 | 8641.604 | 400.000 | 1.000 | 997812.000 |
| MP followee count | 10.961 | 18.080 | 5.000 | 0.000 | 374.500 |
| Duration of Twitter use | 1914.616 | 1445.109 | 1685.500 | 0.000 | 18577.000 |

The results of the regression analysis using the normal dataset and index 1 are presented in Table 4. These results show that a higher *like-mindedness (1)* statistically significantly reduces the *level of incivility* at the 5% level. That is, the more non-like-minded information individuals are exposed to on Twitter, the more uncivil tweets they post. Therefore, Hypothesis 1 is not supported, while Hypothesis 2 is supported.

As shown in Table 5, the results of the analyses using index 1 and the initially not uncivil dataset show that the lower an individual's *like-mindedness (1)*, the higher their *level of incivility*, which is statistically significant at the 5% level. Again, Hypothesis 2 is supported.

The results of the regression analyses using the normal dataset and index 2 are presented in Table 6. The results provide no evidence that *like-mindedness* has a statistically significant effect on the *level of incivility* at the 5% level. However, the fact that a coefficient is not statistically significant does not prove that there is no effect. Subsequently, regression analyses were conducted using the initially not uncivil dataset. As shown in Table 7, low *like-mindedness* significantly increases the *level of incivility* at the 5% level. These results support Hypothesis 2.

The above results provide multiple pieces of evidence supporting Hypothesis 2, while providing no evidence to support Hypothesis 1. Therefore, I conclude that Hypothesis 2 was supported. I illustrate the magnitude of the coefficient using Model 4 as an example. Since it is difficult to interpret the substantive meaning of the value calculated by cosine similarity by itself, I will evaluate the effect size by comparing it to the standard deviations.

Table 4: Analysis using the Normal Dataset and Index 1

| | *Dependent variable:* | |
|---|---|---|
| | Level of incivility | |
| | (1) | (2) |
| Like-mindedness (1) | −0.012*** | −0.026*** |
| | (0.002) | (0.005) |
| Followee count (log) | 0.00004 | −0.0001 |
| | (0.0002) | (0.0004) |
| Duration of Twitter use | −0.00000*** | −0.00000 |
| | (0.00000) | (0.00000) |
| Initial level of incivility | | 0.318*** |
| | | (0.029) |
| Constant | 0.046*** | 0.040*** |
| | (0.001) | (0.002) |
| Observations | 15,617 | 2,778 |
| $R^2$ | 0.003 | 0.058 |
| Adjusted $R^2$ | 0.003 | 0.056 |
| Residual Std. Error | 0.030 (df = 15613) | 0.032 (df = 2773) |
| F Statistic | 14.617*** (df = 3; 15613) | 42.431*** (df = 4; 2773) |
| *Note:* | | *p<0.1; **p<0.05; ***p<0.01 |

Table 5: Analysis using the Initially not Uncivil Dataset and Index 1

| | *Dependent variable:* | |
|---|---|---|
| | Level of incivility | |
| | (3) | (4) |
| Like-mindedness (1) | −0.040*** | −0.039*** |
| | (0.007) | (0.007) |
| Followee count (log) | −0.001** | −0.001** |
| | (0.001) | (0.001) |
| Duration of Twitter use | −0.00000 | 0.00000 |
| | (0.00000) | (0.00000) |
| Initial level of incivility | | 0.270*** |
| | | (0.062) |
| Constant | 0.050*** | 0.049*** |
| | (0.003) | (0.003) |
| Observations | 1,231 | 1,231 |
| $R^2$ | 0.033 | 0.048 |
| Adjusted $R^2$ | 0.031 | 0.045 |
| Residual Std. Error | 0.033 (df = 1227) | 0.033 (df = 1226) |
| F Statistic | 13.922*** (df = 3; 1227) | 15.378*** (df = 4; 1226) |
| *Note:* | | *p<0.1; **p<0.05; ***p<0.01 |

Table 6: Analysis using the Normal Dataset and Index 2

| | *Dependent variable:* | |
|---|---|---|
| | Level of incivility | |
| | (5) | (6) |
| Like-mindedness (2) | 0.002 | −0.008* |
| | (0.002) | (0.005) |
| Followee count (log) | 0.0004* | 0.001 |
| | (0.0002) | (0.001) |
| MP Followee count | −0.0001*** | −0.0002*** |
| | (0.00002) | (0.00005) |
| Duration of Twitter use | −0.00000** | 0.00000 |
| | (0.00000) | (0.00000) |
| Initial level of incivility | | 0.377*** |
| | | (0.035) |
| Constant | 0.040*** | 0.039*** |
| | (0.002) | (0.006) |
| Observations | 12,070 | 1,843 |
| $R^2$ | 0.005 | 0.075 |
| Adjusted $R^2$ | 0.004 | 0.073 |
| Residual Std. Error | 0.031 (df = 12065) | 0.032 (df = 1837) |
| F Statistic | 14.146*** (df = 4; 12065) | 29.851*** (df = 5; 1837) |
| *Note:* | | *p<0.1; **p<0.05; ***p<0.01 |

Table 7: Analysis using the Initially not Uncivil Dataset and Index 2

| | *Dependent variable:* | |
|---|---|---|
| | Level of incivility | |
| | (7) | (8) |
| Like-mindedness (2) | −0.025*** | −0.023*** |
| | (0.008) | (0.008) |
| Followee count (log) | 0.0004 | 0.0003 |
| | (0.001) | (0.001) |
| MP Followee count | −0.0005*** | −0.0004*** |
| | (0.0001) | (0.0001) |
| Duration of Twitter use | 0.00000** | 0.00000*** |
| | (0.00000) | (0.00000) |
| Initial level of incivility | | 0.251*** |
| | | (0.075) |
| Constant | 0.056*** | 0.053*** |
| | (0.009) | (0.009) |
| Observations | 790 | 790 |
| $R^2$ | 0.072 | 0.085 |
| Adjusted $R^2$ | 0.067 | 0.079 |
| Residual Std. Error | 0.034 (df = 785) | 0.033 (df = 784) |
| F Statistic | 15.114*** (df = 4; 785) | 14.508*** (df = 5; 784) |
| *Note:* | | *p<0.1; **p<0.05; ***p<0.01 |

The coefficient of the effect of *like-mindedness (1)* on the *level of incivility* is -0.039, the standard deviation of *like-mindedness (1)* is 0.119, and the standard deviation of the *level of incivility* is 0.030. Thus, when *like-mindedness* increases by an amount equivalent to one standard deviation, the *level of incivility* decreases by an amount equivalent to approximately 1/7 of the standard deviation. An effect size equivalent to approximately 1/7 of the standard deviation is a substantially meaningful effect size.

# Discussion and Conclusion

The question of whether exposure to (non-)like-minded information on the Internet leads to political incivility has not been adequately studied, with previous studies suggesting conflicting answers. Furthermore, there have been insufficient studies on the individual level. This study conducted an individual-level analysis using data from Twitter, and this resulted in more valid findings. The results of the empirical analysis show that, contrary to the findings of a previous study (F. L. Lee et al., 2019), users who are exposed to non-like-minded information on the Internet have higher levels of incivility. Sunstein (2001) proposed that political web pages should always include a link to a page with opposing views to reduce the opinion polarization caused by the echo chamber phenomenon on the Internet. However, based on the findings of this study, designing a platform that encourages exposure to non-like-minded information may induce aggressive communication and contribute to social fragmentation.

This study makes three important contributions to the literature. First, it provided new insights that contradicted the findings of previous studies. I pointed out two competing possibilities and found empirical evidence that contradicts the findings of the previous study. Second, I conducted the analysis at the individual level. In previous studies, the unit of analysis was not at the individual level. Therefore, they could not eliminate the effect of platform culture and the possibility of ecological fallacy. This study was able to avoid these problems because it analyzed individuals on the same platform. Third, it expands the regional scope of political incivility research. To the best of my knowledge, this is the first study to explore the factors of political incivility in the Japanese language.

This study has some limitations. First, it is unclear whether similar results can be found in other languages and cultures. Second, it is also unclear whether the same causal relationship can be observed on other platforms such as Facebook and Instagram. Third, since this study conducted regression analyses using observational data at one time point, the possibility of endogeneity cannot be completely avoided. Therefore, comparative analysis using datasets from multiple languages, cultures, and platforms, as well as analysis using more sophisticated research designs, need to be considered in future studies.

# References

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., … Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216–9221.

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.

Chan, C.-h., Chow, C. S.-l., & Fu, K.-w. (2019). Echoslamming: how incivility interacts with cyberbalkanization on the social media in Hong Kong. *Asian Journal of Communication*, *29*(4), 307–327.

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, *64*(4), 658–679.

Festinger, L. (1957). *A theory of cognitive dissonance* (Vol. 2). Stanford university press.

Garrett, R. K. (2009). Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of computer-mediated communication*, *14*(2), 265–285.

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*, 21.

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.

Kokkaigiin Ichiran List. (n.d.-a). *Members of the House of Councillors.* (`https://democracy.minibird.jp/councillors.php`)

Kokkaigiin Ichiran List. (n.d.-b). *Members of the House of Representatives.* (`https:democracy.minibird.jp`)

Kudo, T. (n.d.). *mecab.* (`https://github.com/taku910/mecab`)

Lee, F. L., Liang, H., & Tang, G. K. (2019). Online incivility, cyberbalkanization, and the dynamics of opinion polarization during and after a mass protest event. *International Journal of Communication*, *13*, 20.

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.

Maia, R. C., & Rezende, T. A. (2016). Respect and disrespect in deliberation across the networked media environment: examining multiple paths of political talk. *Journal of Computer-Mediated Communication*, *21*(2), 121–139.

Matsuno, S., Mizuki, S., & Sakaki, T. (2019). Nihongo daikibo SNS + Web corpus ni yoru tangobunsanhyogen no model kochiku [Constructing of the word embedding model by Japanese large scale SNS + Web corpus]. *Proceedings of the Annual Conference of JSAI/Jinkochino Gakkai Zenkokutaikai Ronbunshu*, 4Rin113–4Rin113.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *arXiv preprint arXiv:1310.4546*.

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

NEologd. (n.d.). *mecab-ipadic-neologd.* (`https://github.com/neologd/mecab-ipadic-neologd`)

Otto, L. P., Lecheler, S., & Schuck, A. R. (2020). Is context the key? the (non-) differential effects of mediated incivility in three European countries. *Political Communication*, *37*(1), 88–107.

Oz, M., Zheng, P., & Chen, G. M. (2018). Twitter versus Facebook: Comparing incivility, impoliteness, and deliberative attributes. *New media & society*, *20*(9), 3400–3419.

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.

Rossini, P. (2020). Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. *Communication Research*, 0093650220921314.

Rowe, I. (2015). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, communication & society*, *18*(2), 121–138.

SlothLib. (n.d.). *Japanese.txt.* (`http://svn.sourceforge.jp/svnroot/slothlib/CSharp/Version1/SlothLib/NLP/Filter/StopWord/word/Japanese.txt`)

Sobieraj, S., & Berry, J. M. (2011). From incivility to outrage: Political discourse in blogs, talk radio, and cable news. *Political Communication*, *28*(1), 19–41.

Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, *30*(3), 341–366.

Stroud, N. J. (2010). Polarization and partisan selective exposure. *Journal of communication*, *60*(3), 556–576.

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.

Sunstein, C. R. (2001). *Republic. com*. Princeton university press.

Sydnor, E. (2018). Platforms for incivility: Examining perceptions across different media formats. *Political Communication*, *35*(1), 97–116.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American journal of political science*, *50*(3), 755–769.

Vaccari, C., Valeriani, A., Barberá, P., Jost, J. T., Nagler, J., & Tucker, J. A. (2016). Of echo chambers and contrarian clubs: Exposure to political disagreement among German and Italian users of Twitter. *Social media+ society*, *2*(3), 2056305116664221.

Vargo, C. J., & Hopp, T. (2017). Socioeconomic status, social capital, and partisan polarity as predictors of political incivility on twitter: A congressional district-level analysis. *Social Science Computer Review*, *35*(1), 10–32.

Watanabe, K. (n.d.). *stopwords_ja.yml.* (`https://github.com/koheiw/marimo/blob/master/yaml/stopwords_ja.yml`)

Watanabe, K. (2020). Latent Semantic Scaling: A semisupervised text analysis technique for new domains and languages. *Communication Methods and Measures*, 1–22.

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*, 20.