

ORIGINAL ARTICLE

Partisan Bias of Perceived Incivility and its Political Consequences: Evidence from Survey Experiments in Hong Kong

Hai Liang¹ & Xinzhi Zhang²

¹School of Journalism and Communication, The Chinese University of Hong Kong

²Department of Journalism, Hong Kong Baptist University

Exposure to presumably uncivil content is neither a necessary nor a sufficient condition for perceptions of incivility and thus could lead to differential political consequences. To examine the emergence and consequences of perceived incivility in disagreement comments, the present study reports on two population-based online survey experiments in Hong Kong ($N_1 = 1,207$, $N_2 = 611$). The results indicate that individuals perceive a higher degree of incivility in disagreement comments directed to in-group members than in those directed to out-group members, regardless of content features. This bias perception is greater when respondents can easily identify the incivility in a comment. Furthermore, exposure to disagreement comments can only influence willingness to participate and affective polarization indirectly via perceived incivility, and such effects are conditional on whether respondents can easily identify the incivility in a comment.

Keywords: Incivility, Perception Bias, Political Disagreement, Affective Polarization, Survey Experiment

doi: 10.1093/joc/jqab008

Uncivil political comments are common on social media (e.g., [Antoci, Delfino, Paglieri, Panebianco, & Sabatini, 2016](#); [Coe, Kenski, & Rains, 2014](#)). Political incivility has been cited as a major problem in online democracy, but researchers have defined and operationalized political incivility in different ways ([Muddiman, 2017](#); [Papacharissi, 2004](#)). Complicating the issue even further, given the vast array of individual and social characteristics that members of any society have, the general public may perceive and interpret ostensibly uncivil messages differently ([Kenski, Coe, & Rains, 2017](#); [Muddiman, 2017](#)). Political incivility is not purely objective, and the boundary between civility and incivility is negotiable, depending on

¹Corresponding author: Hai Liang; e-mail: rainfireliang@gmail.com.

individual characteristics (e.g., gender identity) and social contexts (e.g., swearing between friends or between strangers) (Braunstein, 2018). Variations in definitions and perceptions of political incivility have made it difficult for empirical studies to reach a consensus on the consequences of incivility (see Lu & Vogt, 2020).

Instead of assuming an objective definition of incivility and thus a clear division between civility and incivility, the present study treats the boundary between civility and incivility as a continuous variable that changes across individuals and social contexts. For example, studies have found that individuals with different political identifications may perceive political incivility differently (Muddiman, 2017; Mutz, 2015). Furthermore, based on the theories of social identity and motivated reasoning, the present study hypothesizes that disagreement messages targeted at in-group partisans will be perceived as more uncivil than disagreement directed to out-group partisans, even for messages without presumably uncivil content features. Broadly speaking, the present study demonstrates how incivility perceptions arise as a product of the interaction between individual characteristics (partisanship) and social contexts (messages targeted at in- or out-group partisans).

Furthermore, exposure to presumably uncivil content does not necessarily produce predictable political consequences, because individuals may perceive different levels of incivility in the same content. Previous studies have defined incivility based on presumably uncivil content features like swearing that might be different from perceptions of incivility. The objective approach to political incivility using content analysis or randomized experiments assumes that the same uncivil content can produce perceptions that are at least similar, if not identical. If that assumption is not true, any subsequent consequences would be conditional on the variability of perceived incivility. Using political participation and affective polarization as examples, the present study demonstrates the differential effects of exposure to uncivil content (messages including presumably uncivil content features) and uncivil perceptions. In addition, the present study extends previous theoretical frameworks—mostly with empirical references in the USA—to a non-Western Asian context. Our effort explicates the applicability and generalizability of the related theoretical frameworks on the perceptions and political consequences of online incivility.

Uncivil content versus uncivil perceptions

One central problem in incivility research is that “nobody really agrees on what incivility is” (Chen, Muddiman, Wilner, Pariser, & Stroud, 2019, p. 1). Political incivility has been defined in various ways related to its inappropriateness and undemocratic nature (Muddiman, 2017; Papacharissi, 2004), but precisely which behaviors should be considered inappropriate or undemocratic is socially contingent (Braunstein, 2018). Perceptions of incivility depend on factors such as social position, context of the speech, partisanship, and ideology (Massaro & Stryker, 2012). Nevertheless, based on an extensive literature review, Massaro and Stryker (2012) concluded that sufficient consensus exists about what types of speech count as extremely uncivil:

there are “reasonably shared values that make expression indecent or uncivil in certain contexts” (p. 409). Uncivil speech has certain commonly recognized characteristics, such as profanity directed at a person or speech that is ad hominem, vulgar, or disrespectful.

More recently, incivility has been operationalized according to features of its content: vulgarity, name-calling, hate speech, and so on (e.g., Coe et al., 2014; Kenski et al., 2017). Different types of incivility elicit different perceptions of incivility. For example, Stryker, Conway, and Danielson (2016) found that respondents could reach a consensus on perceived incivility in some forms of uncivil content but not others. Kenski et al. (2017) found that name-calling and vulgarity were rated as more uncivil than other forms, such as accusations of lying and pejorative terms. In addition, perceptions of incivility vary across individuals with different socio-demographic traits or psychological predispositions. For example, females perceive higher levels of incivility than males, and agreeable people perceive messages to be more uncivil than people with other personalities (Kenski et al., 2017). Age and a tendency to conflict-avoidance can also increase people’s reactions to incivility (Ben-Porath, 2008; Mutz & Reeves, 2005). Most pertinent to this article, group identity can bias perceptions of incivility. People are likely to perceive their in-group members as less uncivil than out-group members, even when they engage in precisely the same ostensibly uncivil behaviors (Muddiman, 2017; Mutz, 2015).

To advance the feature-based definition of incivility, research should distinguish between uncivil content (messages with presumably uncivil content features) and uncivil perceptions. Although messages with uncivil content features are usually perceived as uncivil by most people, such features are neither a necessary nor a sufficient condition for perceived incivility. Perceptions of incivility can arise in many ways, even from civil content (messages without presumably uncivil content features), as we demonstrate in this study. If individuals perceive incivility so differently from one another, any attempt to estimate the consequences of political incivility that is evaluated by objective criteria will be problematic.

However, distinguishing between uncivil content and uncivil perceptions does not mean that there is no consensus about what expressions should be considered uncivil. The potentially uncivil content characteristics summarized in previous studies (Coe et al., 2014; Kenski et al., 2017; Massaro & Stryker, 2012) are major factors—though not determinants—influencing uncivil perceptions. Based on this, the present study investigates how partisanship could move perceptions of incivility away from presumably uncivil content and examines those perceptions’ differential effects on political participation and affective polarization.

Partisanship and perceived incivility

It is well documented that partisanship can produce perceptions of news bias (Vallone, Ross, & Lepper, 1985). Political incivility toward out-group members can serve as an identity performance strategy (Rains, Kenski, Coe, & Harwood, 2017).

Based on social identity theory (Tajfel, 1978; Tajfel & Turner, 1979), Kim and Hwang (2018) found that political incivility directed to out-group members could increase party identity salience, and thus increase intergroup bias. Following the logic of social identity theory, partisanship—as a group membership—could cause intergroup bias in favor of in-group members. If political incivility is considered unacceptable in general (Stryker et al., 2016), people—in order to maintain a positive image with their in-group members—may perceive disagreement messages posted by in-group partisans as less uncivil than those posted by out-group partisans. By contrast, disagreements directed to in-group members might be perceived as more uncivil. Similar patterns were found by Muddiman (2017): individuals perceived in-group political figures as less uncivil than those from an out-group party when those figures used precisely the same uncivil statements. However, as we argue below, this partisan bias is not restricted to uncivil content.

According to the theory of motivated reasoning (Kunda, 1987, 1990), disagreement or uncivil messages targeted at in-group members in political discussions can elicit a strong perception of hostility, which can lead to directional (defensive) motivations (Borah, 2014; Kinney & Segrin, 1998). Motivated reasoning is one of the most common approaches to biased political information processing in the context of political misperception (Redlawsk, Civettini, & Emmerson, 2010; Taber & Lodge, 2006). Once directional motivations activate, individuals are more inclined to seek out information to reinforces their preferences, counter-argue information that contradicts their preferences, and view pro-attitudinal information as more convincing than counter-attitudinal information (Taber & Lodge, 2006). Furthermore, partisanship has been demonstrated to be among the most common sources of directionally motivated reasoning (Bolsen, Druckman, & Cook, 2015). When a threat to one's social identity is perceived, defensive motivations can be activated (Sinclair, 2012). Taken together, motivated reasoning provides an additional explanation for the ideological perception bias: disagreement or uncivil content activates directionally motivated reasoning, which leads to self-serving biased processing and eventually a higher level of perceived incivility directed to in-group partisans. Therefore, we posit the following:

H1: Individuals will perceive disagreement statements directed to in-group partisans (in-group) as more uncivil than disagreement statements directed to out-group partisans (out-group).

The moderating role of ambiguity

Previous studies have found that variations in perceived incivility are greater in certain forms of uncivil content. Some forms, such as vulgarity and name-calling, are more easily identified, while others are more ambiguous, such as accusations of lying. According to Kenski et al. (2017), the coefficient of variation (standard deviation/mean) for vulgarity is 0.27, while the coefficient for accusations of lying is 0.33.

This implies that the ambiguity of identifying incivility in a message can play an important role in perceptions of incivility.

Ambiguity has long been recognized as an important factor in both content analysis (Schutz, 1952) and social psychology (Wiener, Carpenter, & Carpenter, 1957), which means that content or experimental stimulus can be interpreted in different ways. Ambiguity has been described as perceived uncertainty about uncertain outcomes (Einhorn & Hogarth, 1985). In the current context, it is the individual who is uncertain about whether to classify a message as uncivil. How ambiguity arises in an individual's perception has to do with the intrinsic properties of the content and the previous experiences of that individual (Schutz, 1952). People may have different estimates of how likely a given message will be interpreted as civil or uncivil, and may thus perceive different levels of ambiguity in the same content. Early studies suggest that ambiguity is positively related to conformity in social influence (Wiener et al., 1957) and confirmation bias is more likely to be activated in ambiguous situations (Nickerson, 1998). People are inclined to interpret ambiguous information as something in favor of in-group members. Therefore, individuals will perceive more incivility in messages directed to in-group than to out-group partisans, especially when they are ambiguous about the incivility of those messages.

However, according to motivated reasoning and social identity theories, defensive motivations are activated when identity is threatened by uncivil or disagreement statements directed to in-group partisans. If the incivility of exposed messages is ambiguous to individuals, they may perceive less identity threat and thus be less likely to activate their directional motivations. This is consistent with empirical evidence showing that ambivalent partisans (i.e., those with less salient group identity) are less likely to engage in directionally motivated reasoning (Lavine, Johnston, & Steenbergen, 2012). Furthermore, if a statement's incivility is completely ambiguous to an individual, it is likely that no threat will be perceived. Following this logic, the group difference stated in *H1* would be conditional on the perceived ambiguity of identifying incivility. Given these contradictory predictions, we propose the following:

RQ1: Will the perception bias proposed in *H1* be conditional on the perceived ambiguity of identifying incivility?

Political consequences of incivility

Researchers have examined different impacts of exposure to uncivil content. In general, incivility has been reported to be detrimental to the realization of online democracy (Massaro & Stryker, 2012). For example, incivility can endanger the cohesiveness of online communities (Stivale, 1997), weaken political trust (Mutz & Reeves, 2005), and make political candidates reluctant to use Twitter to engage with ordinary users (Theocharis, Barbera, Fazekas, Popa, & Parnet, 2016). Exposure to uncivil content also increases negative psychological feelings and thus makes people

less open-minded and more defensive (Chen, 2017; Hwang, Kim, & Kim, 2018; Wang & Silva, 2018).

Incivility can also be beneficial to civic and political engagement. For example, swearing can increase voting intentions (Cavazza & Guidetti, 2014), user attention, and others' online approval (Kwon & Cho, 2017). Other studies have found that uncivil negative information may attract recipients' attention and thus enhance their message recall (e.g., Geer & Geer, 2003), increase their knowledge of political candidates (e.g., Niven, 2006), and raise their interest in political campaigns (e.g., Bartels, 2000).

Most studies on the consequences of incivility employ either content analysis or randomized experiments (Ng, Song, Kwon, & Huang, 2020), examining the direct impacts of exposure to uncivil messages on political outcome variables. Although some previous studies have examined certain cognitive and emotional mediators (e.g., Chen, 2017; Wang & Silva, 2018), perceptions of incivility are missing from most models, except for the one conducted by Muddiman, Pond-Cobb, and Matson (2017); in an experiment, they found that partisan incivility (objectively defined) had no direct effect on news engagement. However, exposure to civil news that emphasized respect and reciprocity increased news engagement via reducing perceptions of incivility by out-group partisans.

To generalize these findings, we argue that perceptions of incivility can mediate the impacts of exposure to uncivil or civil disagreement messages on political outcomes. As we hypothesize above, incivility perceptions arise differently from disagreement messages. It is necessary to test the indirect effects of exposure to messages on political outcomes via perceived incivility. On the one hand, if there are no incivility perceptions generated from a given piece of content, we cannot call the observed differences incivility effects. On the other hand, perceived incivility could be produced without exposure to uncivil content, such as mere exposure to partisan disagreement content (i.e., *H1*). In doing so, the present study tests two such political consequences as examples: political participation and affective polarization.

Political participation

Previous studies have used emotions to connect incivility to political participation: exposure to uncivil content can elicit negative emotions and thus activate defensive motivations and increase political participation (e.g., Chen, 2017; Wang & Silva, 2018). However, the direct effect of uncivil disagreement on political participation is negative because it fails to activate defensive motivations (Chen, 2017). Muddiman et al. (2017) report that perceptions of incivility are negatively associated with news engagement due to automatic reactions to news negativity. In summary, previous studies suggest a negative effect of incivility on participation but the indirect effect could be positive through negative emotions.

In addition to negative emotions, partisanship (in-group vs. out-group disagreement) is another source of directionally motivated reasoning (Bolsen et al., 2015).

Disagreement comments directed to in-group members can increase defensive motivations and thus directly increases political participation, even though, according to *H1*, partisanship increases perceived incivility, which could decrease political participation (Muddiman et al., 2017). That means disagreement statements directed to in-group members would indirectly decrease political participation. Therefore, we posit the following:

H2a: Exposure to disagreement statements directed to in-group partisans will, via perceived incivility, indirectly decrease willingness to participate.

Similarly, compared to presumably civil disagreement, exposure to potentially uncivil disagreement may elicit negative emotions or defensive motivations and thus indirectly increase the likelihood of participation. However, this study proposes another influence path by which uncivil disagreement can increase perceptions of incivility and thus indirectly decrease the likelihood of participation. Therefore, we posit the following:

H2b: Exposure to uncivil disagreement content will, via perceived incivility, indirectly decrease willingness to participate.

Furthermore, the indirect effects might be conditional on the perceived ambiguity of identifying incivility. According to *RQ1*, ideological bias perception might depend on the level of perceived ambiguity. Therefore, the indirect effect of exposure to disagreement directed to in-group partisans via perceived incivility would be either stronger or weaker when individuals feel ambiguous about identifying incivility in a message. In addition, the ambiguity of identifying incivility can narrow the difference in perceptions of incivility between uncivil and civil content. If individuals feel extremely ambiguous about identifying incivility in a pair of messages, it is likely that they would consider them to be similar in terms of perceived incivility. In this situation, any indirect effects via perceived incivility will be either null or small. Therefore, we posit the following:

H2c: The indirect effects posited in *H2a* and *H2b* will be conditional on the ambiguity of identifying incivility.

Affective polarization

Affective polarization refers to the degree to which individuals feel more negatively toward opposing political parties than toward their own (Iyengar, Lelkes, Levendusky, Malhotra, & Westwood, 2019), which is an increasingly important alternative measure of political polarization. Incivility as a form of negativity increases physiological arousal and thus may increase negative feelings toward out-group members (Mutz, 2015). This is also consistent with a prediction drawn from the theory of motivated reasoning: incivility can elicit a strong perception of hostility

(Borah, 2014). In this sense, perceived incivility is positively associated with affective polarization generally.

Furthermore, as stated in *H1*, disagreement or uncivil content directed to in-group partisans could increase perceived incivility, whereas the same content directed to out-group could decrease perceived incivility. Taken together, exposure to in-group or uncivil content could increase affective polarization indirectly via perceived incivility. For example, Druckman, Gubitz, Levendusky, and Lloyd (2019) have demonstrated that exposure to presumably uncivil messages directed to out-group members decreases affective polarization, whereas exposure to presumably uncivil messages directed to in-group members increases affective polarization. Their findings can be fully explained within the mediation framework of perceived incivility: targeting in-group members → perceived incivility → affective polarization. Given that both disagreement statements directed to in-group partisans and presumably uncivil messages could generate perceptions of incivility, we propose the following hypotheses:

H3a: Exposure to disagreement statements directed to in-group partisans will indirectly increase affective polarization via perceived incivility.

H3b: Exposure to uncivil content will indirectly increase affective polarization via perceived incivility.

Based on motivated reasoning and social identity theories, exposure to presumably uncivil messages might not be a threat to group identities when incivility is hard to identify in messages. In this situation, ambiguity could decrease perceptions of hostility and thus directly decrease affective polarization. Meanwhile, ambiguity is a condition for generating perceptions of incivility and could thus moderate the indirect effects on affective polarization. Therefore, we posit the following:

H3c: The indirect effects in *H3a* & *H3b* will be conditional on the ambiguity of identifying incivility; the effects will be stronger when ambiguity is lower.

The theoretical framework, research question, and hypotheses presented above are illustrated and summarized in [Figure S1](#) in the Appendix.

Method

Participants

The present study was implemented in Hong Kong, where different political camps (localism, pro-democracy, and pro-establishment) exist and demonstrate severe intolerance of one another. Online political discussion is active, with online incivility increasing in Hong Kong's major discussion spaces since 2014 (Lee, Liang, & Tang, 2019). An online survey experiment was conducted by Dynata, a web survey panel vendor. To represent the Hong Kong population proportionally, Dynata employed a stratified quota sampling based on gender × age to match the Hong Kong census in

terms of Cantonese speakers aged 18 through 65. In total, 1,207 participants completed the survey; the demographic characteristics are in the Appendix (Table S1). Among survey respondents, 587 (48.6%) reported a clear political identification with one of the three camps (i.e., partisans); they participated in a 2 (in-group vs. out-group) \times 2 (uncivil vs civil) between-subject experiment embedded in the population-based survey.

Materials

Materials were adapted from real news stories and online comments in Cantonese. The adaptation of the stories was based on three criteria: objective (factual messages), balanced (opinions from different political camps), and controversial (having different opinions) (e.g., Chen, 2017). Three stories were selected: the future of one country and two systems in Hong Kong, the interpretation of the Basic Law by the Standing Committee of the National People's Congress, and juxtaposed border controls at Hong Kong's West Kowloon railway station (see Appendix for details).

In addition to the news stories, two lists of disagreement comments (civil vs. uncivil disagreement) were created, based on actual online comments in Cantonese. There were 37 candidate comments including civil disagreement without presumably uncivil content features (e.g., "xxx camp has no evidence and thus the claim is not convincing"), accusations of lying (e.g., "xxx camp is lying"), and vulgarity (e.g., "xxx camp f**k your mother"), as defined and similarly operationalized in Coe et al. (2014). Comments containing accusations of lying and vulgarity were deemed presumably uncivil content. Two types of incivility were chosen to ensure sufficient variance of perceived ambiguity in the identification of incivility. To ensure the quality of the materials, a pretest asked 20 Hong Kong undergraduates to rate the perceived civility of the 37 comments on a 7-point scale. Finally, we selected three civil disagreement comments (perceived civility \geq 5.67), three lying accusation comments ($3.56 \leq$ perceived civility \leq 4.33), and three swearing comments (perceived civility \leq 2.17).

Procedures and manipulation

All participants began by answering demographic questions like gender, age, and political identification (localism, pro-democracy, pro-establishment, neutralism, or no preference). Second, after answering a set of irrelevant questions, participants read a news story randomly selected from the three stories. Participants were then told that the issue discussed in the news story had been the subject of intense online debate and that they would see some comments randomly selected from social media websites.

Third, partisans ($n = 587$) identified with either localism, pro-democracy, or pro-establishment were randomly assigned to one of four conditions: (1) in-group and uncivil, (2) in-group and civil, (3) out-group and uncivil, and (4) out-group and civil. Participants in civil content conditions were presented with two comments

randomly selected from the three comments without uncivil features, while participants in uncivil content conditions were presented with one comment containing an accusation of lying and one swearing comment.

In in-group conditions, the disagreement comments were targeted at in-group members from the same political camp. For example, participants who identified with localism received comments modified by inserting that camp (e.g., “the localism camp has no evidence and thus the claim is not convincing,” “the localism camp is lying,” or “localism camp, f**k your mother”). In the out-group conditions, the disagreement comments were targeted at out-group members from the opposing political camp. Given that localism and pro-democracy respondents in Hong Kong are politically similar in many aspects, we considered pro-establishment to be the opposing camp for both localism and pro-democracy supporters. Participants who identified with the pro-establishment camp received either “the localism camp is lying” or “the pro-democracy camp is lying” messages, for example.

Non-partisans ($n = 620$) were presented with two comments randomly chosen from all displayed comments. After the presentation of a comment, all respondents were asked to rate the perceived incivility of that comment, to indicate the perceived ambiguity of the rating, to report their willingness to participate and feelings toward different political camps, and to answer a set of irrelevant questions.

Measures

Perceived incivility was measured using the four semantic differential items created by [Kenski et al. \(2017\)](#). The four items were rated on 7-point scales with the following anchors (reverse coded): uncivil-civil, impolite-polite, unnecessary-necessary, and disrespectful-respectful. The Cronbach’s alpha of the 4 items was 0.94, so the mean of the 4 items was computed to measure the perceived incivility for each comment. Higher scores indicate greater perceived political incivility. Participants were asked to rate the two comments separately. The final perceived incivility for each participant is the average of the two ratings ($M = 4.01$, $SD = 1.53$).

Perceived ambiguity of identifying incivility was measured by asking participants to rate on a 7-point scale to indicate how much ease they felt when rating the 4 incivility items for each comment (reverse coded). The average of the ambiguity scores for the two comments was calculated as the overall ambiguity for each participant ($M = 3.01$, $SD = 1.31$). Higher scores indicate that participants felt more ambiguous.

Willingness to participate was measured by three questions: (1) “Will you participate in an online discussion about this issue?”; (2) “Will you send emails or postal mail to politicians or government officials to express your opinions?”; and (3) “Will you submit opinion pieces to mass media?” Respondents were asked to rate on 7-point scales their likelihood of participating (1: very unlikely, 7: very likely). Willingness to participate was calculated by the average of the three items ($M = 3.54$, $SD = 1.49$, Cronbach’s alpha = 0.91).

Affective polarization. Drawing on Iyengar, Sood, and Lelkes (2012), we measured affective polarization using favorability ratings of in- and out-party members. Specifically, respondents were asked to rate their feelings toward each political camp from 1 to 7, with higher values denoting greater favorability. Pro-establishment ratings were negatively associated with pro-democracy ratings ($-.39, t = -8.96, p < .001$) and localism ratings ($-.53, t = -15.12, p < .001$), while pro-democracy ratings were positively associated with localism ratings ($.58, t = 17.08, p < .001$). The results suggest that pro-democracy and localism are both opposing camps for the pro-establishment camp. Therefore, a difference measure was constructed by subtracting individuals' rating of pro-establishment from their average rating of localism and pro-democracy. The absolute value of this difference is a measure of affective polarization ($M = 3.04, SD = 1.47$).

Results

Partisan bias of perceived incivility

Before formally testing the hypotheses, we present in Figure 1 the descriptive means of perceived incivility across the manipulated groups. First, respondents that were in uncivil conditions perceived consistently higher incivility than those in civil conditions. Second, respondents reported higher incivility regarding comments directed to the in-group camp than to out-group camps, whether those comments were presumably civil or uncivil. Although the formal participants in the present study are partisans, the centralists (those selecting neutralism or no preference) serve as a

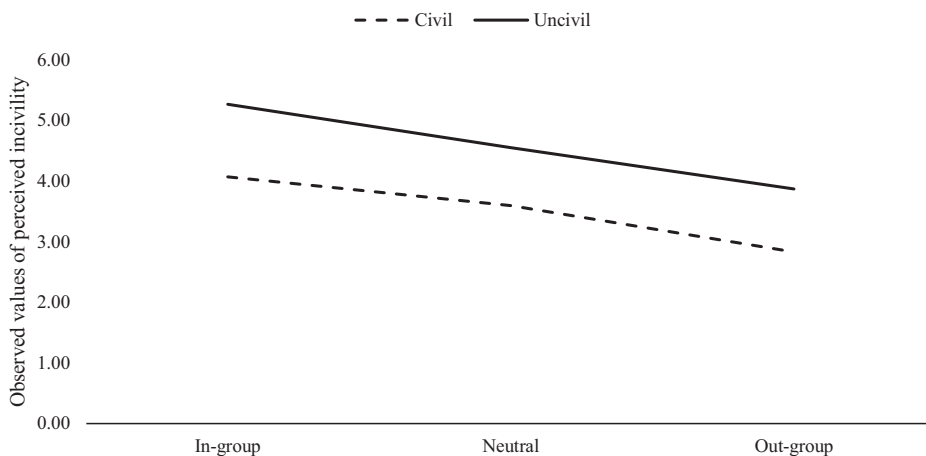


Figure 1 Partisan bias of perceived incivility. In-group: disagreement comments directed to in-group partisans. Out-group: disagreement comments directed to out-group partisans. Neutral: ratings by centralists. Uncivil: disagreement comments including presumably uncivil content features. Civil: disagreement comments excluding presumably uncivil content features.

reference group for objectively evaluating incivility. As Figure 1 shows, the centralists rated incivility consistently higher than participants in the out-group conditions and lower than participants in the in-group conditions, meaning that partisans from all three camps displayed more bias in their perceptions of incivility than non-partisans.

Linear regression models were conducted to examine *H1* and *RQ1*. According to Model I in Table 1, participants in the in-group conditions perceived a higher degree of incivility than those in the out-group conditions ($B = 1.32, p < .001$), regardless of whether those comments were presumably civil or uncivil, *H1* is thus supported. Participants perceived presumably uncivil comments to be more uncivil than civil comments ($B = 1.11, p < .001$). Furthermore, according to Model II in Table 1, the interaction effect between in-group and uncivil is not significant ($B = 0.15, p = .46$), meaning that the ideologically biased perceptions apply to both civil and uncivil disagreement comments in parallel. These findings are consistent with the descriptive statistics presented in Figure 1.

Table 1 OLS Regression Models in Predicting Perceived Incivility

Predictors	Model I <i>Estimates</i> (<i>standard errors</i>)	Model II <i>Estimates</i> (<i>standard errors</i>)	Model III <i>Estimates</i> (<i>standard errors</i>)
Intercept	2.79*** (0.09)	2.83*** (0.10)	1.22*** (0.21)
In-group vs. out-group	1.32*** (0.10)	1.24*** (0.15)	3.46*** (0.25)
Uncivil vs. civil	1.11*** (0.10)	1.04*** (0.21)	1.32** (0.25)
In-group × uncivil		0.15 (0.21)	
Ambiguity			0.55*** (0.07)
In-group × ambiguity			-0.72*** (0.08)
Uncivil × ambiguity			-0.11 (0.08)
Observations	587	587	587
R^2 /adjusted R^2	.32/.32	.32/.32	.45/.42

Note: In-group: disagreement comments directed to in-group members. Out-group: disagreement comments directed to out-group members. Uncivil: disagreement comments including presumably uncivil content features. Civil: disagreement comments excluding presumably uncivil content features. Ambiguity: perceived ambiguity of identifying incivility. OLS: ordinary least squares.

* $p < .05$,

** $p < .01$,

*** $p < .001$.

To answer *RQ1*, perceived ambiguity was included as a moderator in predicting perceived incivility in Model III in [Table 1](#). Perceived ambiguity did not show a significant moderation effect on the perceived incivility gap between uncivil and civil comments ($B = -0.11, p = .14$). The negative coefficient of in-group \times ambiguity ($B = -0.72, p < .001$) suggests that the perceived incivility gap between in- and out-group partisans is conditional on perceived ambiguity. [Figure 2](#) presents the incivility perception gaps between in- and out-group conditions when perceived ambiguity equals 2 (the first quantile: Q1), 3 (Median), and 4 (the third quantile: Q3). Partisan bias is greatest when identifying incivility is easy; however, partisan bias is no longer significant when perceived ambiguity is greater than or equal to 5 ($n = 52$). Therefore, the findings support the conditional hypothesis derived from the theory of motivated reasoning.

Partisan bias of perceived incivility

Linear regression models indicate that the experimental conditions (in- vs. out-group \times uncivil vs. civil) are not significantly related to willingness to participate and affective polarization (see Models I and III in [Table S2](#) in Appendix). Models II and IV in [Table S2](#) indicate that perceived incivility is negatively associated with willingness to participate ($B = -0.18, p = .001$) and positively associated with affective polarization ($B = 0.20, p < .001$). Although the effect sizes are small, [Table S2](#) suggests indirect effects of exposure to in-group incivility on participation and affective polarization via perceived incivility.

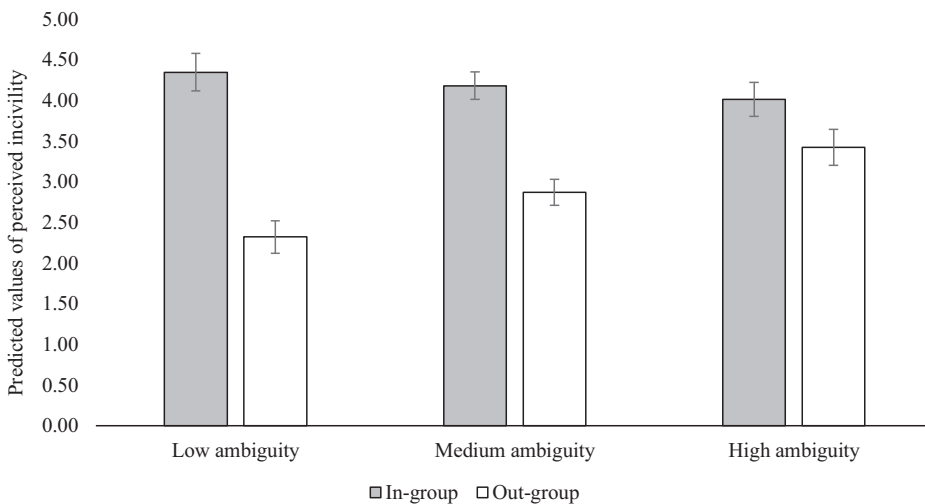


Figure 2 Interaction effect of partisanship by perceived ambiguity on perceived incivility. In-group: disagreement comments directed to in-group members. Out-group: disagreement comments directed to out-group members. Low ambiguity = 2 (Q1), medium ambiguity = 3 (Q2), and high ambiguity = 4 (Q3). Error bars indicate 95% confidence intervals.

Table 2 A Summary of Messages' Direct and Indirect Effects via Perceived Incivility

	Participation Willingness ACME	Participation Willingness ADE	Affective Polarization ACME	Affective Polarization ADE
In-group/Out-group				
Overall	−0.25 ^{***} [−0.40, −0.11]	0.25 [−0.01, 0.51]	0.18 ^{**} [0.03, 0.31]	−0.18 [−0.46, 0.11]
Low ambiguity	−0.43 ^{***} [−0.67, −0.21]	0.41 [*] [0.01, 0.82]	0.41 ^{***} [0.21, 0.63]	−0.52 ^{**} [−0.91, −0.14]
High ambiguity	−0.00 [−0.09, 0.10]	0.18 [−0.17, 0.50]	0.10 [*] [0.02, 0.21]	−0.04 [−0.36, 0.26]
Uncivil/Civil content				
Overall	−0.21 ^{**} [−0.34, −0.09]	0.08 [−0.21, 0.36]	0.15 [*] [0.04, 0.27]	−0.13 [−0.41, 0.16]
Low ambiguity	−0.23 ^{***} [−0.37, −0.11]	−0.02 [−0.33, 0.30]	0.22 ^{***} [0.10, 0.36]	−0.13 [−0.46, 0.19]
High ambiguity	−0.01 [−0.14, 0.14]	0.04 [−0.28, 0.37]	0.15 [*] [0.01, 0.28]	−0.18 [−0.52, 0.15]

Note: The coefficients were estimated based on Imai et al.'s (2010) method. Values in brackets are 95% confidence intervals. Low ambiguity = 2 (Q1), and high ambiguity = 4 (Q3). ACME indicates average causal mediation effects. ADE indicates average direct effects.

* $p < .05$,

** $p < .01$,

*** $p < .001$.

The causal mediation analysis developed by Imai, Keele, and Tingley (2010) was employed to test the indirect effects (see Table 2). Perceived ambiguity was first excluded from the models to obtain the overall indirect effects of targeting in-group members and uncivil content. The average causal mediation effect (ACME) of disagreement directed to in-group partisans on willingness to participate is negative (-0.25 , $p < .001$). The average direct effect (ADE) is not significant. Similarly, the ACME of uncivil comments on willingness to participate is negative (-0.21 , $p = .002$), and the ADE is not statistically significant, meaning that exposure to disagreement directed to in-group partisans or uncivil disagreement comments indirectly decreased, via perceived incivility, willingness to participate. Therefore, *H2a* and *H2b* are supported.

Given the moderating role of perceived ambiguity on perceived incivility, we hypothesize that the indirect effects of targeting in-group members and uncivil content on willingness to participate are also conditional on ambiguity. Moderated mediation models were conducted using Imai et al.'s (2010) method. When identifying

incivility is easy (ambiguity = 2), both the ACME and ADE of the in-group on willingness to participate are significant. When identifying incivility is ambiguous (ambiguity = 4), neither ACME nor ADE is significant. The difference between the ACMEs at different ambiguity levels (2 vs. 4) is -0.43 (95% confidence interval [CI] $[-0.66, -0.20]$, $p < .001$). Similarly, the ACME of uncivil on willingness to participate decreased by 0.23 (95% CI $[0.42, 0.05]$, $p = .02$) from ambiguity = 2 to ambiguity = 4. These findings indicate that the indirect effects of exposure to in-group incivility on willingness to participate exist only when perceived ambiguity is low. Therefore, *H2c* is supported.

In terms of affective polarization, most direct effects are not significant (see the fourth column in Table 2). In general, exposure to disagreement statements directed to in-group partisans decreased affective polarization via perceived incivility at both high and low ambiguity levels (the third column in Table 2). The overall indirect effect is 0.18 ($p = .01$). Exposure to uncivil disagreement comments increased affective polarization via perceived incivility at both ambiguity levels. The overall indirect effect is 0.15 ($p = .012$). Therefore, *H3a* and *H3b* are supported. The difference between the ACMEs of in-group at different ambiguity levels is -0.49 ($p = .046$). The difference for uncivil disagreement is not significant (0.07 , $p = .44$). However, the tests of the moderated mediation effects are heavily influenced by the selection of ambiguity levels. When the ambiguity level is 5, the ACME of in-group is not significant (-0.2 , $p = .53$), and the ACME of uncivil is not significant (0.12 , $p = .17$). Therefore, the results support the conditional hypotheses stated in *H3c*.

Additional study and analysis on ambiguity

This study introduced perceived ambiguity as a moderator to test heterogeneous treatment effects. Perceived ambiguity is measured instead of a manipulated variable. This operationalization is consistent with our non-objective definition of incivility. If incivility cannot be defined objectively, neither can the ambiguity of identifying incivility. However, perceived ambiguity is an endogenous factor, which might be correlated with omitted variables and thus bias the estimation of interaction effects with the treatments. Our data indicate that perceived ambiguity is not associated with partisanship (in- vs. out-group) [$F(1, 585) = 0.05$, $p = .82$] or exposure to uncivil content [$F(1, 585) = 0.11$, $p = .74$]. If the endogenous variable is independent of the treatments (as in this study), OLS estimates of the interaction terms will be consistent and appropriate (Nizalova & Murtazashvili, 2016).

Nevertheless, it might be meaningful to explore the differences between objective and subjective ambiguity. In the survey experiment (Exp.1), the average perceived ambiguity for accusations of lying was 3.07 ($SD = 1.43$), which is higher than the average for swearing comments ($M = 2.92$, $SD = 1.59$), $t_{paired}(308) = 2.1$, $p = .02$. Hence, we may consider exposure to accusations of lying (ambiguous) or swearing (unambiguous) as a natural manipulation of objective ambiguity. Participants in uncivil conditions in Exp.1 were repeatedly exposed to swearing and accusations of

lying. By focusing on these participants ($n = 309$), a random-intercept mixed model was performed to test whether partisan bias is conditional on objective ambiguity (*RQ1*). The results are presented in [Table 3A](#).

In addition, we conducted a second survey experiment (Exp.2, $N = 611$), using a procedure similar to Exp.1. The only difference is that we randomly assigned partisans into either the accusations of lying (ambiguous) or swearing (unambiguous) conditions. The results are presented in [Table 3B](#). The results from Exp.1 and Exp.2 are generally consistent: ambiguous content was perceived less uncivil than unambiguous content, whereas perceived ambiguity was positively associated with perceived incivility; ambiguous content may increase partisan bias, whereas perceived ambiguity decreases partisan bias. Regarding *RQ1*, the manipulation of stimulus ambiguity has either null or an amplification effect on partisan bias. This implies that perceived ambiguity might influence incivility perceptions in a motivated way and that stimulus ambiguity may work in a heuristic way.

In Exp.2, regarding *H2c*, the ACME of exposure to in-group lying accusations is not significant ($-0.01, p = 0.94$), whereas the ACME of exposure to in-group swearing is $-0.39 (p < .001)$. This indicates that stimulus ambiguity, like perceived ambiguity, conditioned the indirect effects of group membership on political participation (difference = $-0.39, p = .008$). However, conditional effects were not found in predicting affective polarization ($p = .852$).

Despite these findings, the reported differences between accusations of lying and swearing may simply be caused by features of uncivil content ([Kenski et al., 2017](#)) other than ambiguity. Ambiguity manipulated by text features is also confounded by the degree of incivility and other content characteristics like familiarity. Even though, our results suggest additional factors that may influence the perceptions of incivility. This question requires additional studies.

Discussion and conclusions

The present study has demonstrated that exposure to uncivil content is neither a necessary nor a sufficient condition to elicit perceptions of incivility. First, simple disagreement comments directed to in-group partisans, without any predefined uncivil content characteristics, can arouse perceptions of incivility, which means that presumably uncivil messages are not a necessary condition for evoking perceived incivility. Second, presumably uncivil messages directed to out-group partisans are perceived as less uncivil than those targeted at in-group members. Indeed, the perceived incivility is very close to how centralists rate civil comments ($p = .64$), which indicates that exposure to uncivil content is not a sufficient condition to make people perceive incivility. In addition, the present study has shown that the perceived ambiguity of identifying incivility is a strong moderator; indeed, it can even eliminate these effects.

Given these findings, the present study suggests a mediation approach to studying incivility consequences via perceived incivility. Our results indicate that

Table 3 Regression Models in Predicting Perceived Incivility by Ambiguity

A (Exp.1)			
Predictors	Model I <i>Estimates</i> (<i>standard errors</i>)	Model II <i>Estimates</i> (<i>standard errors</i>)	Model III <i>Estimates</i> (<i>standard errors</i>)
Intercept	4.30*** (0.11)	4.38*** (0.12)	3.02*** (0.21)
In-group vs. out-group	1.40*** (0.15)	1.23*** (0.17)	3.07*** (0.28)
Lying vs. swearing	−0.85*** (0.07)	−1.01*** (0.10)	1.32** (0.25)
In-group × lying		0.33* (0.14)	
Perceived ambiguity			0.27*** (0.06)
In-group × ambiguity			−0.56*** (0.08)
σ^2/τ_{00}	0.81/1.33	0.80/1.34	1.16/0.94
Observations		309 partisans 618 observations	
Marginal/Conditional R^2	.24/.71	.24/.72	.24/.58
B (Exp.2)			
Predictors	Model I <i>Estimates</i> (<i>standard errors</i>)	Model II <i>Estimates</i> (<i>standard errors</i>)	Model III <i>Estimates</i> (<i>standard errors</i>)
Intercept	4.34*** (0.10)	4.36*** (0.12)	3.05*** (0.19)
In-group vs. out-group	1.70*** (0.12)	1.65*** (0.16)	2.93*** (0.26)
Lying vs. swearing	−0.94*** (0.12)	−0.99*** (0.16)	1.32** (0.25)
In-group × lying		0.11 (0.23)	
Perceived ambiguity			0.29*** (0.06)
In-group × ambiguity			−0.42*** (0.08)
Observations		611	
R^2 /adjusted R^2	.31/.31	.31/.31	.28/.27

Note: Lying: exposure to accusations of lying (ambiguous); Swearing: exposure to swearing comments (unambiguous). Ambiguity: perceived ambiguity.

* $p < .05$,

** $p < .01$,

*** $p < .001$.

intergroup disagreement and potentially uncivil comments did not directly exert significant effects on willingness to participate and affective polarization. Instead, they had indirect effects via perceived incivility (see Table 2). The findings suggest that previous studies that used uncivil content as the treatment and measured incivility effects using the differences of the outcome variables are problematic because the observed differences on the outcome variables may not be caused by perceived incivility but by confounding variables associated with the treatment. The present study paves the way for future studies on incivility effects that can incorporate perceptions of incivility in their theorization and operationalization.

In addition to the main findings, several other points need to be discussed. First, many studies (e.g., Chen, 2017; Wang & Silva, 2018) on the consequences of incivility have been based on the mechanism of defensive motivation, which is characterized by the activation of negative emotions. In our survey, we also asked respondents, after rating incivility, to indicate their emotions related to the displayed comments. Negative emotions were measured by the degree to which the respondents felt sad, angry, or anxious. However, we did not find any significant difference in negative emotions between the conditions. We are not refuting the emotion mediation explanations, but our findings at a minimum suggest that perceptions of incivility could work independently with negative emotions to activate defensive motivations.

However, our finding is consistent with the theory of motivated reasoning in another way. As Flynn, Nyhan, and Reifler (2017) point out, negative affect toward out-group members is only one factor that activates defensive motivations. In the present study, partisan disagreement is another one. Even more importantly, the effects on political participation via negative emotions and perceived incivility are predicted contrarily (increasing vs. decreasing), which might explain why empirical studies have reported contradictory evidence about the impacts of incivility on political participation. As suggested by the theory of motivated reasoning, the present study has found that exposure to disagreement directed to in-group members directly increased willingness to participate (when perceived ambiguity is low) and indirectly decreased it via perceived incivility.

Second, previous studies have noted that group identity can bias perceptions of incivility (Muddiman, 2017; Mutz, 2015) and the impacts of exposure to uncivil messages on affective polarization (Druckman et al., 2019). The present study has extended this line of research in several ways. First, it shows that variations in perceptions of incivility include not only the variability of perceived incivility but also the level of perceived ambiguity of identifying incivility. As we have demonstrated, partisan effects are conditional on the level of perceived ambiguity. Although both confirmation bias and motivated reasoning provide explanations for the role of ambiguity in perceptions of incivility, our findings support the prediction of motivated reasoning: ideological bias is greater when participants perceive less ambiguity in identifying incivility. This implies that incivility is associated with high cognitive arousal, which is consistent with previous findings showing that incivility can increase attention (Kwon & Cho, 2017). Second, the present study shows that group

identity not only biases perceptions of uncivil messages but also of civil disagreement messages. Finally, the study offers an alternative explanation for the impact of intergroup incivility on affective polarization by introducing the mediation of perceived incivility and ambiguity.

Limitations and future research

The study has several limitations that should be addressed in future studies. First, we measured the intensity of perceived incivility as a whole. Perceived incivility could be directed to in-group members or out-group members, which may have differential effects (see [Muddiman et al., 2017](#)). These measurements could introduce another conceptual problem: does perceived incivility indicate the incivility of the message content or the participants involved in the message? Previous studies have measured the perceived incivility of political figures (e.g., [Druckman et al., 2019](#); [Muddiman et al., 2017](#)). In the present study, we have measured the perceived incivility of message content. However, the theoretical and empirical differences between the two types of incivility remain unknown.

Second, the present study drew heavily on theories of social identity and motivated reasoning. However, it did not measure the variables that are directly related to the underlying mechanisms. For example, did the manipulations stimulate a sense of social identity or defensive motivations? Although the theories we employed are widely cited in incivility research, few studies have attempted to validate these basic assumptions. In the present study, we could falsify some competing mechanisms like emotions and confirmation bias, but more empirical studies are needed to understand the full picture by explicitly measuring the defensive goals and salience of, for example, social identity.

Finally, this study proposed a moderated mediation model of perceived incivility. However, the selection of the conditions (group identity and incivility) and outcome variables (participation and affective polarization) was not systematic. Many empirical studies in the field test incivility effects on political deliberation (e.g., open-mindedness) and other forms of participation (e.g., news engagement and voting), areas that were not examined in the present study. More importantly, participants were forced to read the selected disagreement comments in the experiments, which might trigger psychological reactance. And thus the observed effects could vary if they were exposed to the same content in a natural setting. In a free and open online environment, users could also read both disagreement and agreement comments. Researchers are encouraged to systematically test the consequences of incivility with unobtrusive methods under the framework provided in this study.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Please note: Oxford University Press is not responsible for the content or functionality of any [supplementary materials](#) supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Funding

This study was supported by the Direct Grant (4052205) and the Young Researcher Award (4411719) to the first author from the Chinese University of Hong Kong.

Conflict of Interest

We have no conflict of interest to declare.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

References

- Antoci, A., Delfino, A., Paglieri, F., Panebianco, F., & Sabatini, F. (2016). Civility vs. incivility in online social interactions: An evolutionary approach. *PLoS ONE*, *11*(11), doi: 10.1371/journal.pone.0164286
- Bartels, L. M. (2000). Partisanship and voting behavior, 1952-1996. *American Journal of Political Science*, *44*(1), 35–50. doi:10.2307/2669291
- Ben-Porath, E. N. (2008). *Codes of professionalism; Norms of conversation: How political interviews shape public attitudes toward journalists*. Unpublished doctoral dissertation, University of Pennsylvania. Philadelphia, PA.
- Bolsen, T., Druckman, J. N., & Cook, F. L. (2015). Citizens', scientists', and policy advisors' beliefs about global warming. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 271–295. doi:10.1177/0002716214558393
- Borah, P. (2014). Does it matter where you read the news story? Interaction of incivility and news frames in the political blogosphere. *Communication Research*, *41*(6), 809–827. doi: 10.1177/0093650212449353
- Braunstein, R. (2018). Boundary-work and the demarcation of civil from uncivil protest in the United States: Control, legitimacy, and political inequality. *Theory and Society*, *47*(5), 603–633. doi:10.1007/s11186-018-9329-3
- Cavazza, N., & Guidetti, M. (2014). Swearing in political discourse: Why vulgarity works. *Journal of Language and Social Psychology*, *33*(5), 537–547. doi: 10.1177/0261927x14533198
- Chen, G. M. (2017). *Online incivility and public debate: Nasty talk*. Berlin: Springer.

- Chen, G. M., Muddiman, A., Wilner, T., Pariser, E., & Stroud, N. J. (2019). We should not get rid of incivility online. *Social Media + Society*, 5(3), 2056305119862641. doi:10.1177/2056305119862641
- 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. doi:10.1111/jcom.12104
- Druckman, J. N., Gubitza, S. R., Levendusky, M. S., & Lloyd, A. M. (2019). How incivility on partisan media (de)polarizes the electorate. *The Journal of Politics*, 81(1), 291–295. doi:10.1086/699912
- Einhorn, H. J., & Hogarth, R. M. (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, 92(4), 433–461. doi:10.1037/0033-295x.92.4.433
- Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology*, 38, 127–150. doi:10.1111/pops.12394
- Geer, J. G., & Geer, J. H. (2003). Remembering attack ads: An experimental investigation of radio. *Political Behavior*, 25(1), 69–95. doi:10.1023/A:1022904428357
- 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. doi:10.1177/0093650215616861
- Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4), 309–334. doi:10.1037/a0020761
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22(1), 129–146. doi:10.1146/annurev-polisci-051117-073034
- Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, 76(3), 405–431. doi:10.1093/poq/nfs038
- Kenski, K., Coe, K., & Rains, S. A. (2017). Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research*, 0093650217699933. doi:10.1177/0093650217699933
- Kim, Y., & Hwang, H. (2018). When partisans see media coverage as hostile: The effect of uncivil online comments on hostile media effect. *Media Psychology*, 22(6), 845–866. doi:10.1080/15213269.2018.1554492
- Kinney, T., & Segrin, C. (1998). Cognitive moderators of negative reactions to verbal aggression. *Communication Studies*, 49(1), 49–72. doi:10.1080/10510979809368518
- Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, 53(4), 636–647. doi:10.1037/0022-3514.53.4.636
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. doi:10.1037/0033-2909.108.3.480
- Kwon, K. H., & Cho, D. (2017). Swearing effects on citizen-to-citizen commenting online: A large-scale exploration of political versus nonpolitical online news sites. *Social Science Computer Review*, 35(1), 84–102. doi:10.1177/0894439315602664
- Lavine, H. G., Johnston, C. D., & Steenbergen, M. R. (2012). *The ambivalent partisan: How critical loyalty promotes democracy*. New York, NY: Oxford University Press.
- 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, 4940–4959. Retrieved from <https://ijoc.org/index.php/ijoc/article/view/11666/2819>.
- Lu, S. N., & Vogt, O. (2020). *Does online incivility affect engagement? A meta-analysis of research*. Paper presented at the The 69th Annual Conference of International Communication Association (ICA), Gold Coast, Australia.
- Massaro, T. M., & Stryker, R. (2012). Freedom of speech, liberal democracy, and emerging evidence on civility and effective democratic engagement. *Arizona Legal Studies. Discussion Paper No. 12-12*, 54, 375–443.
- Muddiman, A. (2017). Personal and public levels of political incivility. *International Journal of Communication*, 11, 3182–3202. Retrieved from <https://ijoc.org/index.php/ijoc/article/view/6137/2106>.
- Muddiman, A., Pond-Cobb, J., & Matson, J. E. (2017). Negativity bias or backlash: Interaction with civil and uncivil online political news content. *Communication Research*, 0093650216685625. doi:10.1177/0093650216685625
- Mutz, D. C. (2015). *In-your-face politics: The consequences of uncivil media*. Princeton, NJ: Princeton University Press.
- Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, 99(1), 1–15. doi: 10.1017/S0003055405051452
- Ng, Y. L., Song, Y., Kwon, K. H., & Huang, Y. (2020). Toward an integrative model for online incivility research: A review and synthesis of empirical studies on the antecedents and consequences of uncivil discussions online. *Telematics and Informatics*, 47, doi: 10.1016/j.tele.2019.101323
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220. doi:10.1037/1089-2680.2.2.175
- Niven, D. (2006). Throwing your hat out of the ring: Negative recruitment and the gender imbalance in state legislative candidacy. *Politics & Gender*, 2(4), 473–489. doi: 10.1017/S1743923x06060120
- Nizalova, O. Y., & Murtazashvili, I. (2016). Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term. *Journal of Econometric Methods*, 5(1), 71. doi:10.1515/jem-2013-0012
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2), 259–283. doi: 10.1177/1461444804041444
- Rains, S. A., Kenski, K., Coe, K., & Harwood, J. (2017). Incivility and political identity on the Internet: Intergroup factors as predictors of incivility in discussions of news online. *Journal of Computer-Mediated Communication*, 22(4), 163–178. doi:10.1111/jcc4.12191
- Redlawsk, D. P., Civettini, A. J. W., & Emmerson, K. M. (2010). The affective tipping point: Do motivated reasoners ever "get it"? *Political Psychology*, 31(4), 563–593. doi: 10.1111/j.1467-9221.2010.00772.x
- Schutz, W. C. (1952). Reliability, ambiguity and content analysis. *Psychological Review*, 59(2), 119–129. doi:10.1037/h0049742
- Sinclair, B. (2012). *The social citizen: Peer networks and political behavior*. Chicago, IL: University of Chicago Press. [Database].
- Stivale, C. J. (1997). Spam: Heteroglossia and harassment in cyberspace. In D. Porter (Ed.), *Internet culture* (pp. 133–144). New York, NY: Routledge.

- Stryker, R., Conway, B. A., & Danielson, J. T. (2016). What is political incivility? *Communication Monographs*, 83(4), 535–556. doi:10.1080/03637751.2016.1201207
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755–769. doi: 10.1111/j.1540-5907.2006.00214.x
- Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Oxford: Academic Press.
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In M. A. Hogg & D. Abrams (Eds.), *Intergroup relations: Essential readings*. (pp. 33–47). Belmont, CA: Wadsworth.
- Theocharis, Y., Barbera, P., Fazekas, Z., Popa, S. A., & Parnet, O. (2016). A bad workman blames his tweets: The consequences of citizens' uncivil Twitter use when interacting with party candidates. *Journal of Communication*, 66(6), 1007–1031. doi: 10.1111/jcom.12259
- Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut Massacre. *Journal of Personality and Social Psychology*, 49(3), 577–585. doi:10.1037/0022-3514.49.3.577
- Wang, M. Y., & Silva, D. E. (2018). A slap or a jab: An experiment on viewing uncivil political discussions on Facebook. *Computers in Human Behavior*, 81, 73–83. doi: 10.1016/j.chb.2017.11.041
- Wiener, M., Carpenter, J. T., & Carpenter, B. (1957). Some determinants of conformity behavior. *The Journal of Social Psychology*, 45(2), 289–297. doi: 10.1080/00224545.1957.9714311