Social Norms and the Dynamics of Online Incivility

Yotam Shmargad, University of Arizona, yotam@arizona.edu

Kevin Coe, University of Utah, kevin.coe@utah.edu

Kate Kenski, University of Arizona, kkenski@arizona.edu

Steve Rains, University of Arizona, srains@arizona.edu

ABSTRACT

Online discussions are performed in the gaze of fellow users. To increase engagement, platforms typically let these users evaluate the comments made by others through rating systems (e.g., via Likes or Down/Up votes). Understanding how such ratings shape, and are shaped by, features of the underlying discussion is important for our understanding of online behavior. In this study, we focus on an increasingly concerning aspect of online discussions: incivility. We draw on the theory of normative social behavior (TNSB) to analyze a dataset of over 6,000 online newspaper comments. We find that repeated incivility by the same person is more likely when their initial incivility was affirmed by both descriptive norms (incivility in nearby comments) and injunctive norms (Up votes). Repeated incivility receives more Up votes if nearby comments also include incivility but fewer Up votes if they do not, suggesting that injunctive norms are contextual and shaped by descriptive norms. We conclude that online incivility is a dynamic, normative process that is responsive to both positive feedback and proximate incivility.

Keywords: Descriptive and Injunctive Norms, Click Speech, Discussion Context, Incivility, Online Comments, Theory of Normative Social Behavior, Virtual Communities

1. INTRODUCTION

The internet was initially praised for its potential to foster a new era of inclusive and constructive discussion (Rheingold, 1993). More recently, there has been increasing concern that online systems may be contributing to the erosion of democracy (Anderson & Rainie, 2020). Among the concerns is a worry that the internet has led to a deterioration of discussion quality (Rainie et al., 2017), fueled in part by the presence of incivility (Coe et al., 2014), hate speech (Álvarez-Benjumea & Winter, 2018), and disingenuous commenting in the form of "trolling" (Phillips, 2015). These discussion features decrease people's willingness to engage with new information (Kim & Park, 2018) and heighten perceptions of disagreement (Hwang et al., 2014). They serve to silence minority perspectives (Ordoñez & Nekmat, 2019), thereby producing environments in which people are discouraged from expressing their viewpoints (Yun & Park, 2011). Online discussions can thus suffer greatly when incivility runs rampant.

Alongside these concerns is a renewed interest in understanding what drives people to engage in incivility online (Rains et al., 2017). Civility is often characterized as a social norm (Jamieson et al., 2017) and, as such, should be "promoted and sustained through everyday informal positive and negative sanctions" (Massaro and Stryker 2012, p. 439). Yet, Hmielowski et al. (2014) document how online environments, in particular, can socialize people into viewing verbally aggressive behaviors as acceptable and, in some cases, radicalize them into behaving in such a way themselves. Indeed, "the language of disgust and hate play an important role in communication" online, and those "who use the language of hatred and disgust appear to gain a stronger network position" in online discussions (Oegema et al., 2010, p. 7). Although there is a well-established norm of civility in face-to-face discussions (Mutz & Reeves, 2005), incivility appears to be alive and well online.

As Jamieson et al. (2017) point out, uncivil discussions are more emotionally arousing than their civil counterparts (Mutz, 2007). A prominent feature of online environments that may promote incivility is thus the rating systems that they employ (e.g., Likes or Down/Up Votes), which are more likely to reward incivility than civility (Rains et al., 2017). These rating systems impose on users what Cialdini and colleagues (1990) call an *injunctive norm*, or feedback about what people view as appropriate behavior. Cialdini et al. (1990) contrast the injunctive norm with the *descriptive norm*, which refers to the ways in which people actually behave. In their theory of normative social behavior (TNSB), Rimal and Real (2005) outline how descriptive and injunctive norms interact to influence social behaviors. In this study, we apply the TNSB model to examine how incivility manifests in online discussions by analyzing over 6,000 comments posted on the website of an online newspaper.

The remainder of the paper is organized as follows: In the next section, we discuss online incivility and how the TNSB model can help make sense of such incivility as a function of social norms. In Section 3, we describe our dataset of newspaper comments, including how comments were coded for incivility, and outline our empirical strategy for investigating the relationships between descriptive norms, injunctive norms, and incivility by discussion members. We present the results of our analyses in Section 4 and visualize the main effects of interest. In Section 5 we discuss the implications of our study, considering its limitations and suggesting directions for future work.

2. INCIVILITY AND SOCIAL NORMS IN ONLINE DISCUSSIONS

Incivility – messages that are unnecessarily disrespectful to the discussion or its participants (Coe et al., 2014) – has grown into a central public concern over the past several

years. In a recent survey, for instance, 69% of the U.S. public viewed incivility as a "major problem" (Weber Shandwick / Powell Tate, 2018). Another poll found that more than nine in ten people felt that "the lack of civility in politics today is a serious problem" (Quinnipiac University, 2018). Scholars have also displayed a heightened interest in incivility in recent years (e.g., Berry & Sobieraj, 2014; Boatright et al., 2019; Kenski et al., 2020; Muddiman, 2017; Rossini, 2020), especially that which occurs across a range of online spaces, such as blogs (e.g., Anderson et al., 2014; Borah, 2014), news site comment sections (e.g., Rowe, 2015), and social media platforms (e.g., Su et al., 2018). We engage this burgeoning literature by investigating the interactive dimensions of incivility in online spaces. That is, beyond the known effects that site policies and structure can have (e.g., Ksiazek, 2015; Rowe, 2015), how does online incivility reflect and shape the dynamic messaging that occurs online?

The extant literature has explored such dynamics in two ways that are especially relevant here. First, research has considered the extent to which encountering (in)civility in an online environment might influence people's subsequent site behavior. Thus far, the evidence is mixed. Gervais (2015), for example, demonstrated experimentally that encountering incivility – especially passionate "histrionic" incivility – in an online message board leads participants to include more incivility in their subsequent posts, though this effect held only when the original message was consistent with participants' political predispositions. Han and Brazeal (2015) also experimentally demonstrated such "modeling" behavior, finding that those who read an online newspaper article and then viewed civil comments about the article were more likely to post civil comments subsequently themselves, compared to those who had read the same article paired with uncivil comments (see also Sydnor, 2019). Other experimental research has found civil comments, but not uncivil ones, to produce such modeling (Han et al., 2018; Molina & Jennings,

2017). Rösner et al. (2016), using a design similar to Han and Brazeal's but focusing only on modeling incivility, found that exposure to a varying number of uncivil comments (ranging from zero to six) had no effect on the degree of incivility in subsequent posts.

The second strand of incivility research particularly relevant here is that which focuses on incivility in relation to "click speech" (Sklan, 2013) – meaning the use of evaluation features on websites, such as those that allow users to Like or Down/Up vote content. Click speech can serve as a simple heuristic that might affect site users' perceptions of content, and it can also be a form of expression reflecting attitudes about certain content features such as incivility (Kim & Park, 2018). Pang et al. (2016), for example, conducted an experiment in which participants read a Facebook post followed by either civil or uncivil comments. Those exposed to civil comments reported stronger intentions to Like the comments than did those exposed to uncivil comments. This is consistent with the work of Coe, Kenski, and Rains (2014), whose study of a newspaper comment section found that uncivil comments received significantly more Down votes than did civil comments. In an experiment that gave participants the ability to evaluate comments on a mock news site via Down/Up votes, however, Naab and Kalch (2017) found that incivility had no effect on negative evaluations.

The present study weaves together these two threads in research on incivility, situating them in the theoretical context of social norms. Behavioral scholars have long argued that people rely on others to understand their world (Sherif, 1936) and respond to social feedback, both positive and negative, when deciding how to act (Skinner, 1938). Social norms – the informal rules that govern our lives (Massaro & Stryker, 2012) – are in many ways necessary for society to function properly (Coleman, 1990). The role of social norms in shaping human behavior has been demonstrated in a wide variety of settings, including public littering (Cialdini et al., 1990),

alcohol consumption (Borsari & Carey, 2003), household energy conservation (Schultz et al., 2007), political participation (Smith & Louis, 2008), handwashing (Lapinski et al., 2014), and contraception use (Sedlander & Rimal, 2019). The literature discussed above has also begun to indicate that social norms, as made evident in online messaging and click speech, might influence behaviors pertaining to incivility. We delve deeper into this possibility in the paragraphs that follow.

Social norms can be categorized according to whether they are located at the collective or perceived level, as well as whether they are descriptive or injunctive (Lapinski & Rimal, 2005). Collective norms "operate at the level of the social system, which could be a social network" (Lapinski & Rimal, 2005, p. 130). On the other hand, perceived norms are found at the level of the individual and reflect one's understanding of the collective norms. Perceived norms can vary across individuals, even in the same context, because they rely on communication processes that also vary across people (Lapinski & Rimal, 2005). While descriptive norms reflect what people typically do, injunctive norms capture how people ought to behave (Cialdini et al., 1990). Descriptive and injunctive norms can both be analyzed at the collective and perceived levels (Lapinski & Rimal, 2005).

Rimal and Real (2003) suggest that descriptive and injunctive norms interact, whereby people are more likely to comply with a behavior if it is both commonplace and encouraged by others. Lapinski and Rimal (2005) contend that, "when people perceive that social sanctions exist for noncompliance, they are more likely to conform if they also perceive that the behavior is widespread among their peers" (p. 133). This interaction effect was formalized by Rimal and Real (2005) in the TNSB. Since then, the interaction of descriptive and injunctive norms has been found to predict compliance with several behaviors, including alcohol use (Lee et al.,

2007), the signing of political petitions (Smith & Louis, 2008), and handwashing among childcare workers (Lapinski et al., 2014).

Because people rely on communication to learn social norms (Lapinski & Rimal, 2005), online discussions are useful for studying norm evolution (McLaughlin & Vitak, 2011). For example, Chung (2019) maintains that online "comment spaces can be especially relevant as a communication channel for social norms" (p. 552). In their study of the discussion website Reddit, Chandrasekharan et al. (2018) explain that "an understanding of community norms is generally gained through experience: observing posts and comments posted on the subreddit, peer feedback in the form of votes or replies to comments, and interactions with [moderators]" (p. 32:2). The first two of these online features capture descriptive and injunctive norms of the online community, respectively, while the third refers to sanctions from website moderators that can also influence commenting behavior (Stroud et al., 2015).

Scholars have found that social norms are especially useful for predicting anti-social commenting, including prejudice (Hsueh et al., 2015), incivility (Rains et al., 2017), and hate speech (Álvarez-Benjumea & Winter, 2018). Álvarez-Benjumea and Winter (2018) manipulated descriptive and injunctive norms in an experiment, finding that the descriptive treatment was more effective in deterring hate speech. Whereas their descriptive treatment involved censoring hateful comments to create an environment with less hate speech, their injunctive treatment included discouraging replies to hateful comments. While such "counter-speaking" measures have been shown to decrease prejudice in some settings (Munger 2016), Álvarez-Benjumea and Winter (2018) found that this led to a backlash, with sanctioned commenters becoming even more likely to post hateful comments after being sanctioned. Such sanctioning thus represents a "hard" injunctive norm that can backfire.

Unlike counter-speaking, a "softer" injunctive norm can be applied through click speech, (e.g., via Likes or Down/Up votes), which also lets users evaluate the comments made by others. Cheng et al. (2014) show that even in the case of click speech, injunctive norms can lead to a backlash effect. They analyzed data from four different online news communities, finding that Down votes encouraged discussion members to post even lower-quality comments. Injunctive norms alone are thus not enough to deter users from engaging in anti-social commenting. What remains to be seen is whether this is due to injunctive norms interacting with descriptive norms, as implied by TNSB. Following the predictions of TNSB, we propose the following hypothesis:

H1: Commenter incivility is more likely when both descriptive and injunctive norms supporting incivility are present.

In general, there is likely to be more acceptance of anti-social behavior in environments that already feature such behavior. For example, Cheng et al. (2017) discovered that people are more likely to engage in trolling once exposed to trolling themselves, a finding consistent with those of Hmielowski et al. (2014). This occurs for positive behaviors as well. Sukumaran et al. (2011), for example, find that a norm of being thoughtful develops when others are thoughtful. As they put it, "exposure to locally situated social behavior that indicates a normative level of thoughtfulness in an online comment space can induce individuals to conform to this standard in their own commenting activity, as well as cause them to judge such norm-consistent behavior likely on the part of others" (Sukumaran et al., 2011, p. 3406). This suggests that injunctive norms that reward incivility are more likely to form when the descriptive norm around incivility is already present, which leads us to offer a second hypothesis:

H2: Injunctive norms are more likely to reward incivility in contexts where the descriptive norm is to be uncivil.

3. DATA AND MODELS

Data for this project were originally collected by Coe et al. (2014). They examined the general prevalence of incivility and did not consider the roles of descriptive and injunctive norms. Coe et al. (2014) retrieved a census of all comments posted by readers in response to online news stories from a daily newspaper during a three-week period in 2011. The newspaper served a metropolitan area in the United States consisting of approximately 1 million residents. A total of 6,444 comments were collected from 310 different articles. The name of the article, content of each comment, identity of the comment author, and the number of Down and Up votes for each comment were recorded. All comments were evaluated by a team of trained research assistants for one of five types of incivility (inter-coder reliability, via Krippendorff's alpha, is reported in parentheses): name-calling ($\alpha = .67$), aspersion ($\alpha = .61$), lying ($\alpha = .73$), vulgarity ($\alpha = .91$), and pejorative for speech ($\alpha = .74$). Down ($\alpha = 1.00$) and Up ($\alpha = 1.00$) votes took the form of a green "Thumbs Up" and red "Thumbs Down" symbol displayed at the bottom of each comment that could be clicked by readers. One or more forms of incivility was present in 22% of comments. More than 95% of comments had at least one Down ($M = 7.00$, $SD = 10.18$) or Up ($M = 14.26$, $SD = 24.90$) vote.

3.1 Measuring Norms

Those who commented on news articles in our setting could gauge the adequacy of their contribution in two ways: community ratings and the discussion context. Community ratings are the cumulative number of Down and Up votes that comments receive. Down and Up votes reflect the injunctive norms in this setting, informing members about what the community deems to be appropriate to say when commenting. The discussion context, on the other hand, captures comments made by other community members on the same news article. The discussion context thus reflects the descriptive norm in this setting, allowing commenters to observe the behavior of other members and to use it as benchmark when making their own contributions.

Our interest is in understanding how descriptive and injunctive norms shape incivility. We focus on one key indicator of the descriptive norm: the absence or presence of incivility in the comment that immediately follows one's own. We thus assume that people are attentive to the very next comment that is posted after theirs, and that they use this information when deciding whether to be civil or uncivil in subsequent comments. While people may also learn about the community's descriptive norms from the comment that was posted before theirs, this would not be the case if they read the news article, formulated a comment, and then posted it without paying attention to the existing comments on the article. We thus assume that people are most interested in the types of responses that their comments yielded.

Formally, for each comment *t* posted by commenter *i*, we define its *proximate comment* to be the very next comment by another member on the same news article. We are interested in how people use this contextual information, and so only include proximate comments that are made by a different member *j*. We thus exclude from our analysis comments that were followed by a comment from the same member (just under 12% of the comments). For example, suppose commenter *i* posted three comments in a row. We would then only include their third comment,

and the proximate comment would be defined to be the first comment by another commenter *j* after commenter *i*'s third comment.

Formalizing the descriptive norm in this way allows us to map the social network of community members, with network links defined between each commenter *i* and their proximate commenters *j*. We depict this social network of community members in Figure 1. The size of each node captures a member's activity, or the number of comments they posted across all articles, while a node's color captures the proportion of a member's comments that were uncivil (with darker shades indicating higher levels of incivility). Following Lapinski and Rimal (2005), a social network view can capture this particular community's *collective* descriptive norms with respect to incivility. Figure 1 reveals several descriptive aspects of this community. First, this online community is civil at its core. Members that are uncivil in most of their comments are typically found in the periphery, rather than the center, of the network. Second, the network does not break down into separate article-specific clusters, suggesting a fair amount of cross-article commenting and thus a single collective rather than many disconnected discussions.

The network diagram in Figure 1 also suggests that there is variation in incivility *within* members. Many nodes are light or dark gray, implying that many members are uncivil only some of the time. Our motivation in this study is to explain why members choose to be civil in some cases and uncivil in others, and the extent that descriptive and injunctive norms shape these decisions. Given that there is variation in member-level incivility, we also set out to understand when online communities reward and penalize repeated incivility and, in particular, how this varies with incivility in other members' comments (i.e., the descriptive norm).

[Insert Figure 1]

3.2 Empirical Strategy

Before describing our statistical models that relate descriptive norms, injunctive norms, and incivility, we summarize our data preparation and empirical strategy in Figure 2. Our first model investigates how descriptive and injunctive norms interact to shape repeated incivility. This model uses incivility of the initial comment in a chain (commenter *i*), the number of Down and Up votes this comment received, and the presence of incivility in the proximate comment (by commenter *j*) to predict subsequent incivility by commenter *i*. Our second model captures how injunctive norms are shaped by descriptive norms, using incivility in the first and final comment in the chain (by commenter *i*), and in the proximate comment (by commenter *j*), to predict the number of Down and Up votes that the final comment by commenter *i* receives. Next, we discuss our statistical models in more detail.

[Insert Figure 2]

As Figure 2 makes clear, the dependent variable in our first model is the presence of incivility in the final comment, which is a dichotomous variable. The dependent variable in our second model is the number of Down or Up votes that the final comment receives, which are both non-negative integers (i.e., count data). We use multilevel modeling employing logit and negative binomial link functions to test Models 1 and 2, respectively. Multilevel models make it possible to address two sources of non-independence in our data: some articles may draw in more incivility than other comments, while some community members may be more inclined to be uncivil than others. Accordingly, we use a multilevel specification with random effects for

article and community member (using the MEGLM command in Stata). We also show that our results are robust to linear specifications with fixed effects for article and member (using the REGHDFE command in Stata).

Our specification for Model 1 is as follows (with article and commenter effects left out for brevity of exposition):

$$\begin{aligned}\mathrm{Inc}_{it+1} = {} & \alpha + \beta \mathrm{Inc}_{it} + A_1\left(\mathrm{Inc}_{jt} + \mathrm{Down}_{it} + \mathrm{Up}_{it}\right) + A_2 \mathrm{Inc}_{jt} \times \left(\mathrm{Down}_{it} + \mathrm{Up}_{it}\right) \\ & + B_1 \mathrm{Inc}_{it} \times \left(\mathrm{Inc}_{jt} + \mathrm{Down}_{it} + \mathrm{Up}_{it}\right) + B_2 \mathrm{Inc}_{it} \times \mathrm{Inc}_{jt} \times \left(\mathrm{Down}_{it} + \mathrm{Up}_{it}\right) + \Gamma X_{it} + e_{it}\end{aligned}$$

The dependent variable Inc_{it+1} is an indicator of whether the final comment in a chain is uncivil. The right-hand side is a three-way interaction between 1) incivility in the initial comment (Inc_{it}), 2) incivility in the proximate comment (Inc_{jt}), and 3) the number of Down and Up votes that the initial comment receives (Down_{it} and Up_{it}, respectively). For the sake of clarity, we present the effects for initially *civil* chains with α, A_1, and A_2, and the incremental effects for initially *uncivil* chains with β, B_1, and B_2. We also include two control variables, which are found in X_{it}: 1) the number of comments in between the initial and final comment in the chain (Gap), and 2) the position of the initial comment in the thread of comments on a given article (Rank).

In Model 2, we use separate specifications for the number of Down and Up votes that the final comment in a chain receives:

$$\begin{aligned}\mathrm{Down}_{it+1} = {} & \alpha + \beta \mathrm{Inc}_{it} + A_1\left(\mathrm{Inc}_{jt} + \mathrm{Inc}_{it+1}\right) + A_2 \mathrm{Inc}_{jt} \times \mathrm{Inc}_{it+1} \\ & + B_1 \mathrm{Inc}_{it} \times \left(\mathrm{Inc}_{jt} + \mathrm{Inc}_{it+1}\right) + B_2 \mathrm{Inc}_{it} \times \mathrm{Inc}_{jt} \times \mathrm{Inc}_{it+1} + \Gamma X_{it} + e_{it}\end{aligned}$$

$$Up_{it+1} = \alpha + \beta Inc_{it} + A_1(Inc_{jt} + Inc_{it+1}) + A_2 Inc_{jt} \times Inc_{it+1}$$
$$+ B_1 Inc_{it} \times (Inc_{jt} + Inc_{it+1}) + B_2 Inc_{it} \times Inc_{jt} \times Inc_{it+1} + \Gamma X_{it} + e_{it}$$

The right-hand side in this case includes a three-way interaction between 1) incivility in the initial comment (Inc_{it}), 2) incivility in the proximate comment (Inc_{jt}), and 3) incivility in the final comment (Inc_{it+1}). Again, we present the effects for initially civil chains with α, A_1, and A_2, while the incremental effects for initially uncivil chains are found in the parameters β, B_1, and B_2. In addition to the Gap and Rank variables, we also include as a control variable the number of Up votes on the final comment (Up_{it+1}) when modeling its Down votes, and the number of Down votes on the final comment ($Down_{it+1}$) when modeling its Up Votes.

4. RESULTS

Model 1 captures how descriptive norms (i.e., incivility in proximate comments) and injunctive norms (i.e., Down and Up votes) jointly shape repeated incivility by community members. Before discussing the results of our full specification, we first describe the results of separate models for chains in which the initial comment was civil versus uncivil. We report the results of these separate specifications in the two leftmost columns of Table 1. A comparison of the number of observations across these specifications shows that about 80% of chains start out with a civil comment. The only statistically significant predictor of incivility in initially civil chains is the negative constant term, suggesting that members who are initially civil tend to remain civil in subsequent comments. For these chains, neither injunctive norms nor descriptive norms influence whether members stray from their initial civility.

In comparison, for chains that begin with an uncivil comment, the constant term is not statistically significant, and the initial commenter's repeated incivility is predicted by the number of Up votes they receive on their initial comment and the presence of incivility in the proximate comment. In particular, the effect of proximate incivility is negative and statistically significant, implying that members tend to become more civil when a descriptive incivility norm is present but the injunctive incivility norm is absent. However, the effect of proximate incivility tends to zero as the number of Up votes on the initial comment increases and becomes positive at about 20-25 Up votes. After this point, members tend to remain uncivil – the injunctive and descriptive incivility norms interacting to promote repeated incivility. Notably, we do not find a significant effect of Down votes, suggesting that penalties do not influence members' decisions.

In the two rightmost columns of Table 1, we report the results of the full three-way interaction model under logit and linear specifications, respectively. Under both specifications, we find a statistically significant difference in how Up votes and proximate comments shape subsequent incivility in initially civil versus uncivil chains. The direction of these effects is the same as in the specifications that separately model initially civil and uncivil chains: proximate incivility decreases the initial commenter's tendency to remain uncivil in subsequent comments, but the effect reverses as the initial (uncivil) comment receives more Up votes. We thus find strong support for our first hypothesis.

[Insert Table 1]

In Figure 3, we depict the marginal effects of proximate incivility under the linear specification (the rightmost column of Table 1). Estimates from the fixed portion of the logit

specification show similar patterns and are available from the first author upon request. The top panels of Figure 3 show that the effect of proximate incivility does not vary with the number of Down votes the initial comment receives, or with whether the member's initial comment was civil (top left panel) or uncivil (top right panel). Simply put, receiving Down votes on an initial comment does not shape how commenters react when encountering incivility in the discussion context – there is neither a tendency to be more or less civil.

[Insert Figure 3]

The bottom two panels of Figure 3 depict how the effect of proximate incivility varies with the number of Up votes that members received on their initial comment. In the bottom right panel, we see that instances of repeated incivility are most common when 1) the commenter's initial incivility is rewarded with Up votes and 2) their initial incivility was met with incivility in the proximate comment. Repeated incivility thus occurs when two sources of normative support exist, one descriptive (i.e., incivility in the proximate comment) and the other injunctive (i.e., Up votes). Indeed, when there is lack of injunctive support for their initial incivility, incivility in the proximate comment discourages commenters from remaining uncivil. This finding is consistent with the work of Lapinski et al. (2014), who find that descriptive norms around handwashing can backfire when injunctive norms are weak.

Our next set of results speak to which comment chains receive higher levels of negative and positive community ratings. We first report the predictors of negative ratings (i.e., Down votes), again starting with separate specifications for initially civil and uncivil chains. The results of these separate specifications can be found in the two leftmost columns of Table 2. Across both

initially civil and uncivil chains, we find that incivility in the final comment is penalized with more Down votes. Moreover, this effect does not vary by whether the discussion context is civil or uncivil, as is evident from the lack of statistically significant interactions with incivility in the proximate comment. When it comes to Down votes, injunctive norms thus appear to penalize incivility regardless of the descriptive norms.

In the two rightmost columns of Table 2, we report estimates from our full model under negative binomial and linear specifications, respectively. Under both specifications, incivility in the final comment is again associated with more Down votes for that comment, and the effect does not significantly vary by whether the initial or the proximate comment is uncivil. We thus do not find support for our second hypothesis when injunctive norms are captured with Down votes. Estimates of the control variables show several additional predictors of Down votes. The final comment receives more Down votes when it is closer to the initial comment in the chain, when it is earlier in the thread of comments on a given article, and when the comment receives more Up votes.

[Insert Table 2]

Before plotting the marginal effects associated with these estimates, we discuss predictors of positive community ratings (i.e., Up votes). We report these results in Table 3, with the two leftmost columns showing results of separate models for initially civil and uncivil chains and the two rightmost columns showing results from the full model. Here, we find that the number of Up votes on the final comment depends on the presence of incivility in both the member's initial comment and in the proximate comment. In particular, repeated incivility by the same member is

penalized with fewer Up votes in a civil discussion context, but rewarded with more Up votes in an uncivil discussion context. These results do not vary by whether we separately model initially civil and uncivil chains, or by whether the full model takes a negative binomial or a linear form. We thus find strong support for our second hypothesis when injunctive norms are captured with Up votes. Results of the control variables show similar patterns as Down votes, with the final comment receiving more Up votes when it is closer to the initial comment in the chain, earlier in the thread of comments on a given article, and when the it also receives more Down votes.

[Insert Table 3]

We conclude this section by depicting the marginal effects from the linear specifications reported in Tables 2 and 3. Estimates from the fixed portion of a negative binomial specification show similar patterns and are available from the first author. In Figure 4, we depict estimates from the models of Down and Up votes in the top and bottom panels, respectively. The top panels show that the final comment in a chain receives more Down votes when it is uncivil. This effect is directionally more pronounced in civil contexts and when the member's initial comment was uncivil, though in neither case did these differences reach statistical significance. Members thus appear to administer Down votes based on incivility in the comment itself, rather than in response to incivility in the discussion context.

[Insert Figure 4]

The bottom panels of Figure 4 depict marginal effects of incivility on Up votes in the chain's final comment. Here, we see that these effects vary with a member's own initial incivility and incivility in the discussion context. When a member's initial comment is uncivil (bottom right panel), they are penalized for remaining uncivil in a civil discussion context but rewarded for this behavior in an uncivil context. These results suggest that, when community members are administering Up votes, they are sensitive to the descriptive norms around incivility. When a member remains uncivil despite others in the chain being civil, this is seen as inappropriate and fewer Up votes are rewarded than when the member switches to civility. However, in an uncivil discussion context, repeated incivility receives more Up votes, suggesting that injunctive norms are sensitive to the descriptive norm. Repeated incivility is rewarded when it fits the context.

5. DISCUSSION

This study is among the first to examine the effects of descriptive and injunctive norms on incivility in a real-world setting, applying the TNSB (Rimal & Real, 2005) to investigate a concerning aspect of online discussions (Rainie et al., 2017). Focusing on incivility within a three-week census of comments posted to an online newspaper, our analysis reveals the importance of social norms in encouraging and discouraging incivility. In particular, our findings underscore the significant role of the discussion context (Cheng et al., 2017) in determining how incivility unfolds in online comments. Specifically, the comment that immediately follows one's own can influence the likelihood of repeated incivility. Online commenters, it appears, are not simply firing incivility blindly into discussions. Instead, they are sensitive to the behaviors of other members in the discussion community. Along with the known effects that site policies and structure can have in explaining perceived and actual incivility in online communities (e.g.,

Ksiazek, 2015; Rowe, 2015; Van Duyn & Muddiman, 2020), discussion context is another important factor.

Importantly, the discussion context influences incivility through its interaction with an indicator of injunctive norms: Up votes. Our findings reveal that, when a community member's uncivil comment is met with incivility, they tend to become more civil in subsequent comments. A descriptive norm around incivility is thus not enough to keep members uncivil. However, when a member's initial comment also receives many Up votes, they tend to remain uncivil. Descriptive and injunctive norms are thus both needed for repeated incivility to manifest in online discussions. We capture descriptive and injunctive norms using the discussion context and community ratings, respectively. Our empirical strategy thus allows us to map these well-studied social science constructs onto the digital traces that have more recently become available due to the proliferation of online discussion forums (Lazer, 2015).

In addition to applying the TNSB (Rimal & Real, 2005) to incivility, our mapping also allows us to build on this work to better understand how injunctive norms evolve in relation to existing descriptive norms. In a civil discussion context, communities respond negatively to repeated incivility and penalize uncivil comments with more Down votes and fewer Up votes. In an uncivil discussion context, however, repeated incivility tends to be rewarded with Up votes, suggesting that a commenter's fit with the overall discussion context can influence community ratings. All of this confirms the importance of "click speech" as not only a simple heuristic that might influence perceptions, but as a meaningful form of online expression (Kim & Park, 2018; Sklan, 2013). Taking this idea further, our findings illustrate that incivility in online discussions will likely be misunderstood if researchers do not account for the role of click speech.

Situating these findings within the normative considerations that animate much research on incivility (see Boatright et al., 2019; Massaro & Stryker, 2012) can provide additional insights into their importance. Notably, civility – not incivility – was the dominant feature of this online community. Indeed, incivility was not central to this discussion network, and roughly 80% of the discussion chains started out with a civil comment. Moreover, when chains started out civil, they tended to remain civil. Any discussion of normative concerns about incivility should bear that reality in mind: it is often the exception, not the rule. Still, we see in our analysis the potential for incivility to propagate in ways that might harm the deliberative aims of a discussion. This occurs especially when members of the community enact two deliberate behaviors: other commenters respond uncivilly and other members of the community reward commenters' incivility with Up votes. This kind of ramping up is somewhat akin to a schoolyard scuffle, which grows more vicious when the person originally provoked responds with equal force and a crowd of observers roars their approval. Of course, these dynamics are also easily avoided. Commenting civilly and not supporting incivility with Up votes are likely to discourage continued incivility.

This study has several limitations that future research might seek to address. First, we focused on just a single online community. Although Coe et al. (2014) make a case for this site's structural consistency with many other online newspaper sites, there is the possibility that other sites might have different community norms and structures, thus generating different incivility dynamics. For example, this website did not allow "threading," which lets members respond to specific comments. While previous work has shown that threading does not change the overall rate of incivility (Budak et al., 2017), it could limit the scope of descriptive norms in important ways. Future research that compares several different sites and structures would be particularly valuable, especially given the contextual nature of the effects we uncover.

Additionally, our design operationalized the discussion context as the most proximate subsequent comment. We viewed this as the most valid test, working from the assumptions that 1) greater proximity signals the likelihood of greater attention and 2) commenters are likely to be interested in the reactions that their comments garner. Still, given that we opted for a real-world setting rather than an experimental one, there is necessarily some uncertainty around these assumptions. In particular, there could be interesting variation across commenters in where their attention lies both before and after commenting – and thus the set of other comments that form an individual commenter's perceived descriptive norms. Future research could explore additional means of operationalizing the discussion context in order to explore individual-level variation in perceived descriptive norms.

Finally, our study focused on just one element of online discussion: incivility. Consistent with public and scholarly interest, we view this as an important element. Nevertheless, future research could consider whether other features of discussion follow the same normative patterns observed here. Just as norms around incivility may vary from one community to another, so may the extent that attention gets paid towards incivility rather than other discussion features. For example, given recent interest in the spread of misinformation, the norms surrounding comment veracity may also play an important role in commenting behavior. We hope that our mapping of normative constructs onto digital traces serves as a guide for researchers interested in applying theories of normative behavior to better understand online discussions.

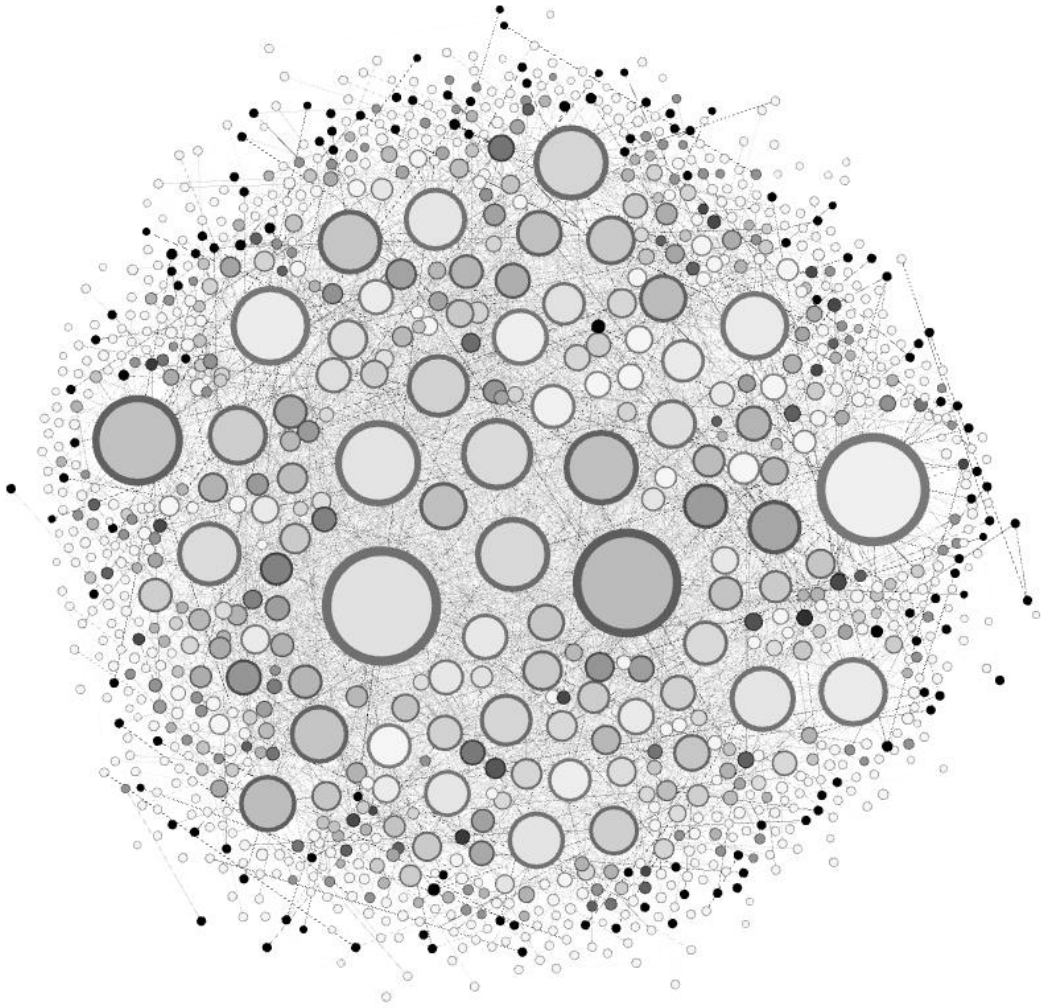

Figure 1: Social Network of Commenters by Activity (size; larger = more comments) and Civility (color; darker = more uncivil)

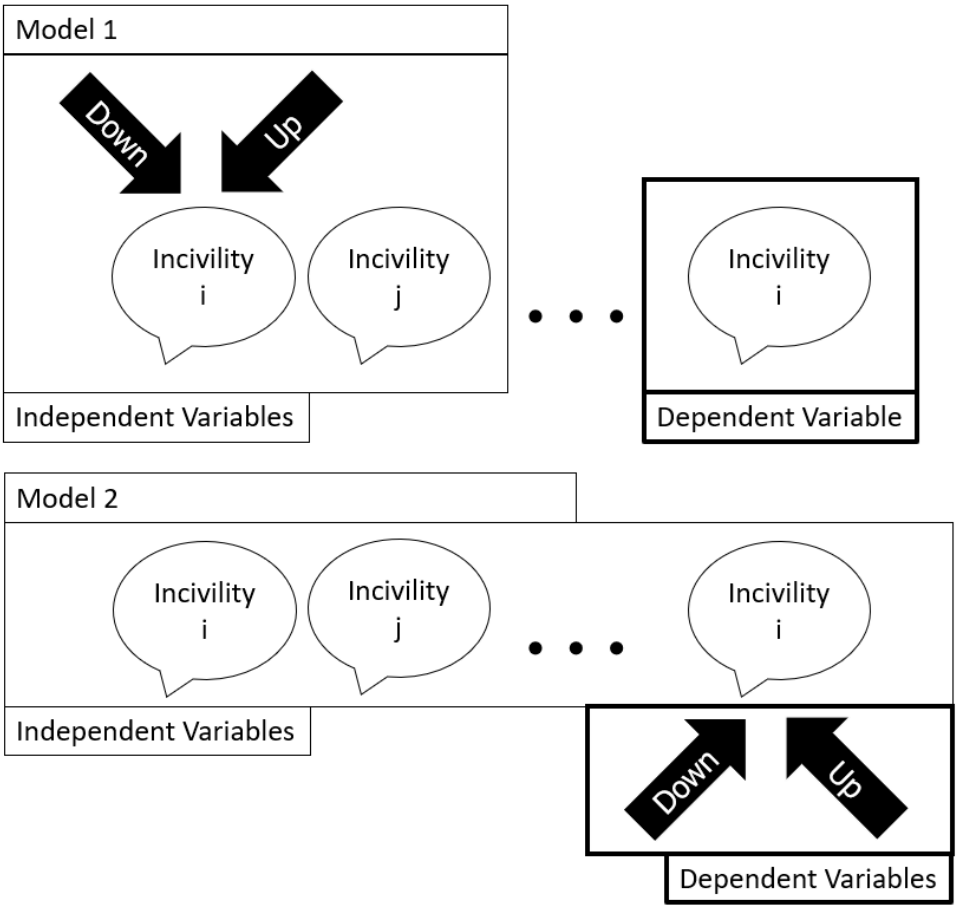

Figure 2: Summary of Data Preparation and Empirical Strategy

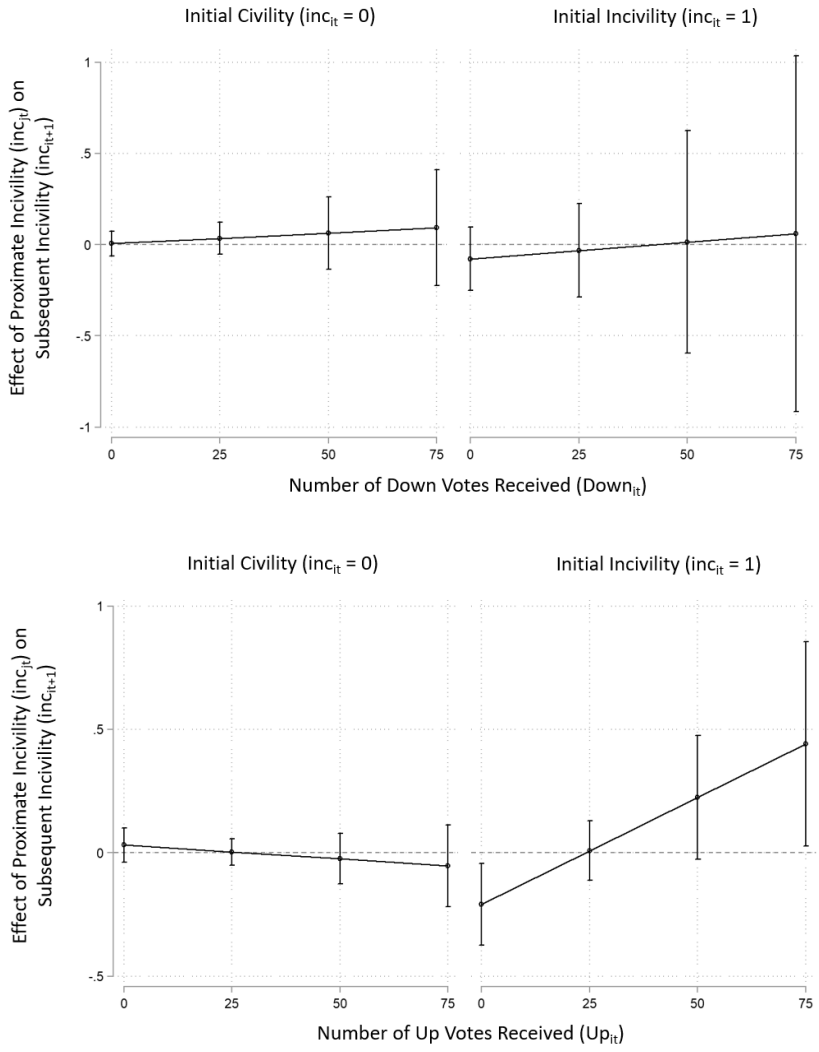

Figure 3: The Interactive Effect of Descriptive and Injunctive Norms on Incivility

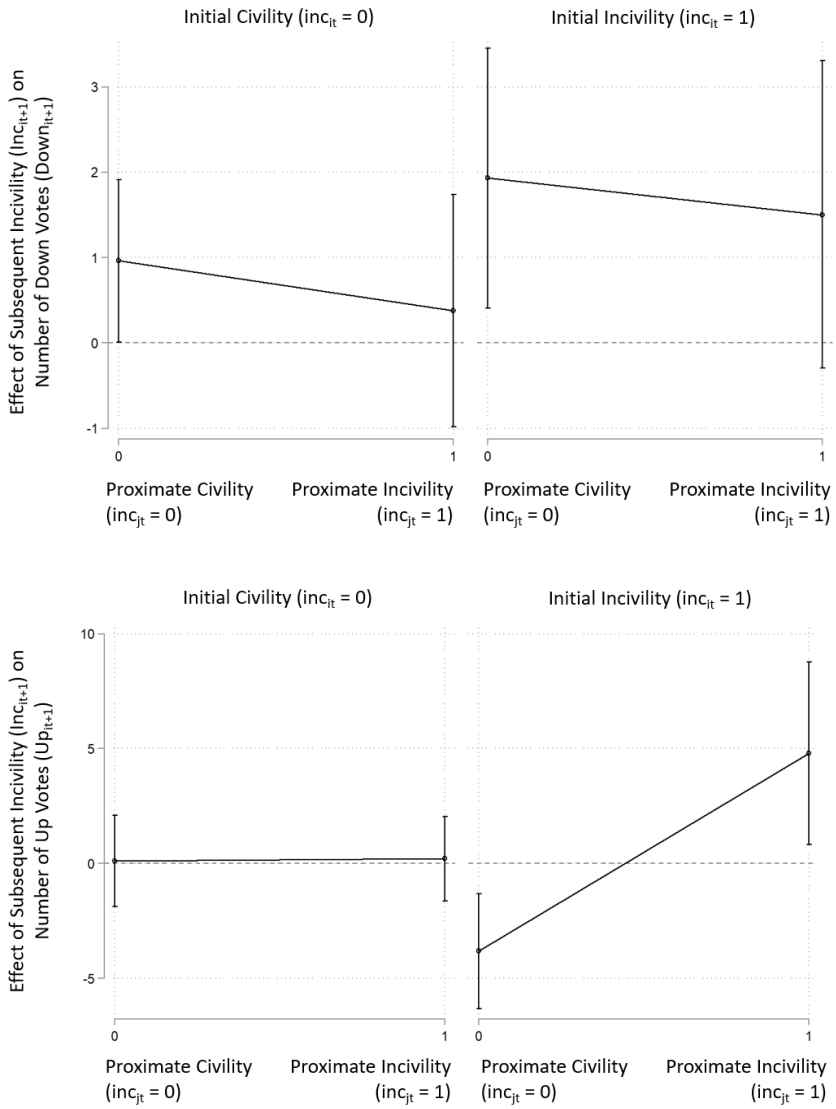

Figure 4: The Role of Descriptive Norms in Shaping Injunctive Norms around Incivility

Table 1 Models of Subsequent Incivility

Model 1 DV: Inc_{it+1}	By Initial Civility (Logit)				Three-way Interaction			
	Initial Civility ($Inc_{it} = 0$)		Initial Incivility ($Inc_{it} = 1$)		Logit		Linear	
	Coef.	RSE	Coef.	RSE	Coef.	RSE	Coef.	RSE
Proximate Civility								
Constant	-2.28**	.218	-.568	.359	-2.03	.165	.158**	.023
$Down_{it}$	.011	.009	.010	.026	.012	.008	.001	.002
Up_{it}	.002	.008	-.018	.013	.002	.007	.000	.001
Inc_{it}					1.02**	.270	.014	.051
$Inc_{it} \times Down_{it}$					-.007	.032	.000	.004
$Inc_{it} \times Up_{it}$					-.015	.018	-.002	.002
Proximate Incivility								
Inc_{jt}	.468	.263	-1.20**	.436	.433	.270	-.022	.041
$Inc_{jt} \times Down_{it}$	.008	.015	-.006	.038	.005	.014	.001	.002
$Inc_{jt} \times Up_{it}$	.003	.008	.054**	.014	-.020	.017	-.001	.000
$Inc_{it} \times Inc_{jt}$					-1.74**	.471	-.249**	.095
$Inc_{it} \times Inc_{jt} \times Down_{it}$					-.013	.045	.001	.008
$Inc_{it} \times Inc_{jt} \times Up_{it}$					.076**	.023	.010*	.004
Control Variables								
Gap	.002	.003	-.006	.034	.001	.002	.000	.000
Rank	-.024	.013	-.056	.051	-.045*	.022	-.001	.001
Article Effects	Random		Random		Random		Fixed	
Commenter Effects	Random		Random		Random		Fixed	
Wald χ^2	26		29		188			
Adjusted R^2							.131	
Observations	1,485		337		1,822		1,822	

Notes: * $p < 0.05$; ** $p < 0.01$. RSE = Robust Standard Errors.

Table 2 Models of Down Votes

Model 2a DV: $Down_{it+1}$	By Initial Civility (Neg. Binomial)				Three-way Interaction			
	Initial Civility ($Inc_{it} = 0$)		Initial Incivility ($Inc_{it} = 1$)		Neg. Binomial		Linear	
	Coef.	RSE	Coef.	RSE	Coef.	RSE	Coef.	RSE
Proximate Civility								
Constant	.974**	.136	.937**	.242	.889**	.128	5.68**	.339
Inc_{it+1}	.123*	.054	.188*	.083	.135*	.054	.965*	.485
Inc_{it}					.012	.057	.147	.390
$Inc_{it} \times Inc_{it+1}$					.060	.114	.968	.900
Proximate Incivility								
Inc_{jt}	.026	.051	-.097	.072	.017	.048	.138	.405
$Inc_{jt} \times Inc_{it+1}$	-.053	.114	.054	.168	-.073	.101	-.586	.841
$Inc_{it} \times Inc_{jt}$					-.072	.104	-.257	.750
$Inc_{it} \times Inc_{jt} \times Inc_{it+1}$					.144	.268	.158	1.41
Control Variables								
Gap	-.005**	.001	-.005*	.002	-.005**	.001	-.022**	.004
Rank	-.056**	.008	-.058**	.019	-.058**	.009	-.163**	.017
Up Votes	.020**	.003	.030**	.009	.021**	.003	.176**	.021
Article Effects	Random		Random		Random		Fixed	
Commenter Effects	Random		Random		Random		Fixed	
Wald χ^2	156		404		296			
Adjusted R^2							.627	
Observations	1,488		337		1,825		1,825	

Notes: * $p < 0.05$; ** $p < 0.01$. RSE = Robust Standard Errors.

Table 3 Models of Up Votes

Model 2b DV: Up_{it+1}	By Initial Civility (Neg. Binomial)				Three-way Interaction			
	Initial Civility ($Inc_{it} = 0$)		Initial Incivility ($Inc_{it} = 1$)		Neg. Binomial		Linear	
	Coef.	RSE	Coef.	RSE	Coef.	RSE	Coef.	RSE
Proximate Civility								
Constant	1.94**	.087	1.96**	.132	1.88**	.088	9.21**	.642
Inc_{it+1}	0.01	.051	-.316**	.089	.009	.052	.116	1.01
Inc_{it}					.058	.046	.912	.881
$Inc_{it} \times Inc_{it+1}$					-.189*	.090	-3.93*	1.58
Proximate Incivility								
Inc_{jt}	.039	.039	-.183*	.082	.027	.041	.368	.628
$Inc_{jt} \times Inc_{it+1}$	-.026	.092	.438*	.188	-.030	.096	.096	1.37
$Inc_{it} \times Inc_{jt}$					-.176*	.079	-2.76	1.46
$Inc_{it} \times Inc_{jt} \times Inc_{it+1}$					.448*	.208	8.51**	2.68
Control Variables								
Gap	-.004**	.000	-.004**	.001	-.004**	.000	-.053**	.008
Rank	-.032**	.010	-.063**	.013	-.033**	.011	-.257**	.028
Down Votes	.040**	.006	.072**	.015	.045**	.007	.572**	.073
Article Effects	Random		Random		Random		Fixed	
Commenter Effects	Random		Random		Random		Fixed	
Wald χ^2	694		56		650			
Adjusted R^2							.481	
Observations	1,488		337		1,825		1,825	

Notes: * $p < 0.05$; ** $p < 0.01$. RSE = Robust Standard Errors.

AUTHOR INFORMATION

Yotam Shmargad is an Assistant Professor in the School of Government & Public Policy at the University of Arizona, yotam@arizona.edu

Kevin Coe is a Professor in the Department of Communication at the University of Utah, kevin.coe@utah.edu

Kate Kenski is a Professor in the Department of Communication at the University of Arizona, kkenski@arizona.edu

Steve Rains is a Professor in the Department of Communication at the University of Arizona, srains@arizona.edu

DATA AVAILABILITY

Data used for analysis will be made available as a repository on the first author's Github page. The authors require that any paper using the data cite the present manuscript.

SOFTWARE INFORMATION

Gephi 0.9.2 was used to produce the social network visualization. Stata MP 13.1 was used for model estimation. A .do file replicating all of the analyses in the paper is available in an online appendix.

REFERENCES

Álvarez-Benjumea, A.& Winter, F. (2018). Normative Change and Culture of Hate: An Experiment in Online Environments. *European Sociological Review*, 34(3), 223-237.

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

Anderson, J. & Rainie, L. (2020). Many Tech Experts Say Digital Disruption Will Hurt Democracy. *Pew Research Center*. https://www.pewresearch.org/internet/2020/02/21/many-tech-experts-say-digital-disruption-will-hurt-democracy/. Accessed July 29, 2020.

Berry, J. M., & Sobieraj, S. (2014). *The outrage industry: Political opinion media and the new incivility*. New York, NY: Oxford University Press.

Boatright, R. G., Shaffer, T. J., Sobieraj, S., & Young, D. G. (Eds.). (2019). *A Crisis of Civility? Political Discourse and its Discontents*. New York, NY: Routledge.

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.

Borsari, B. & Carey, K. B. (2003). Descriptive and Injunctive Norms in College Drinking: A Meta-Analytic Integration. *Journal of Studies on Alcohol*, 64, 331-341.

Budak, C., Garret, R. K., Resnick, P., Kamin, J. (2017). Threading is Sticky: How Threaded Conversations Promote Comment System User Retention. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing.* Portland, OR.

Chandrasekharan, E., Samory, M., Jhaver, S., Charvat, H., Bruckman, A., Lampe, C., Eisenstein, J., & Gilbert, E. (2018). The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing.* New York, NY.

Cheng, J., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2014). How Community Feedback Shapes User Behavior. *Proceedings of the International AAAI Conference on Weblogs and Social Media.* Ann Arbor, MI.

Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions. *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing.* Portland, OR.

Chung, J. E. (2019). Peer Influence of Online Comments in Newspapers: Applying Social Norms and the Social Identification Model of Deindividuation Effects (SIDE). *Social Science Computer Review*, 37(4), 551-567.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places. *Journal of Personality and Social Psychology*, 58(6), 1015-1026.

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments. *Journal of Communication, 64*, 658-679.

Coleman, J. S. (1990). *Foundation of Social Theory*. Cambridge, MA: Belknap Press.

Gervais, B. T. (2015). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics, 12*(2), 167-

Han, S., & Brazeal, L. M. (2015) Playing nice: Modeling civility in online political discussions. *Communication Research Reports, 32*, 20-28.

Han, S., Brazeal, L. M., & Pennington, N. (2018). Is civility contagious? Examining the impact of modeling in online political discussions. *Social Media+ Society*, Online First.

Hmielowski, J. D., Hutchens, M. J., & Cicchirillo, V. J. (2014). Living in an Age of Online Incivility: Examining the Conditional Indirect Effects of Online Discussion on Political Flaming. *Information, Communication & Society, 17*, 1196-1211.

Hsueh, M., Yogeeswaran, K., & Malinen, S. (2015). "Leave Your Comment Below": Can Biased Online Comments Influence Our Own Prejudicial Attitudes and Behaviors? *Human Communication Research,* 41, 557-576.

Hwang, H., Kim, Y., & Huh, C. U. (2014). Seeing is Believing: Effects of Uncivil Online Debate on Political Polarization and Expectations of Deliberation. *Journal of Broadcasting & Electronic Media*, 58(4), 621-633.

Jamieson, K. H., Volinsky, A., Weitz, I., & Kenski, K. (2017). The Political Uses and Abuses of Civility and Incivility. In K. Kenski & K. H. Jamieson (Eds.), *The Oxford Handbook of Political Communication* (pp. 203-217). New York: Oxford University Press.

Kenski, K., Coe, K., & Rains, S. A. (2020). Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research*, 47(6), 795-814.

Kim, J. W., & Park, S. (2018). How Perceptions of Incivility and Social Endorsement in Online Comments (Dis) Encourage Engagements. *Behaviour & Information Technology*, Online first.

Ksiazek, T. B. (2015). Civil Interactivity: How News Organizations' Commenting Policies Explain Civility and Hostility in User Comments. *Journal of Broadcasting & Electronic Media*, *59*, 556-573.

Lapinski, M. K. & Rimal, R. N. (2005), An Explication of Social Norms. *Communication Theory*, 15, 127–147.

Lapinski, M. K., Anderson, J., Shugart, A., & Todd, E. (2014). Social Influence in Child Care Centers: A Test of the Theory of Normative Social Behavior. *Health Communication*, 29(3), 219-232.

Lazer, D. (2015). Issues of Construct Validity and Reliability in Massive, Passive Data Collections. *The City Papers: An Essay Collection from The Decent City Initiative*.

Lee, C. M., Geisner, M. I., Lewis, M. A., Neighbors, C., & Larimer, M. E. (2007). Social Motives and the Interaction Between Descriptive and Injunctive Norms in College Student Drinking. *Journal of Studies on Alcohol and Drugs*, 68(5), 714-721.

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.

McLaughlin, C. & Vitak, J. (2011). Norm Evolution and Violation on Facebook. *New Media & Society*, 14(2), 299-315.

Molina, R. G., & Jennings, F. G. (2017). The role of civility and metacommunication in Facebook discussions. *Communication Studies*, 69, 42–66.

Muddiman, A. (2017). Personal and public levels of political incivility. *International Journal of Communication, 11*, 3182-3202.

Munger, K. (2016). Tweetment Effects on the Tweeted: Experimentally Reducing Racist Harassment. *Political Behavior*, 39, 629-649.

Mutz, D. C. (2007). Effects of "In-Your-Face" Television Discourse on Perceptions of a Legitimate Opposition. *American Political Science Review*, 101(4), 621-635.

Mutz, D. C. & Reeves, B. (2005). The New Videomalaise: Effects of Televised Incivility on Political Trust. *American Political Science Review*, 99(1), 1-15.

Naab, T. K., & Kalch, A. (2017). Replying, evaluating, flagging: How users engage with uncivil and impolite comments on news sites. Paper presented at the International Communication Association annual meeting, San Diego, CA.

Oegema, D., Wang, S., & Kleinnijenhuis, J. (2010). Dynamics of Online Discussions about Politics: A Function of Structural Network Properties, Mass Media Attention or Emotional Utterances? *Proceedings of the Web Science Conference.* Raleigh, NC.

Phillips, W. (2015). *This Is Why We Can't Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture. Cambridge*, MA: MIT Press.

Ordoñez, M. A. M. & Nekmat. E. (2019). "Tipping Point" in the SoS? Minority-Supportive Opinion Climate Proportion and Perceived Hostility in Uncivil Online Discussion. *New Media & Society*, 21(11-12), 2483-2504.

Pang, N., Ho, S. S., Zhang, A. M., Ko, J. S., Low, W. X., & Tan, K. S. (2016). Can spiral of silence and civility predict click speech on Facebook? *Computers in Human Behavior*, *64*, 898-905.

Quinnipiac University (2018, July 3). Harsh words for U.S. family separation policy, Quinnipiac University national poll finds. Retrieved from https://poll.qu.edu/national/release-detail?ReleaseID=2554

Rainie, L., Anderson, J., & Albright, J. (2017). The Future of Free Speech, Trolls, Anonymity and Fake News Online. *Pew Research Center*. https://www.pewresearch.org/internet/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online/. Accessed July 29, 2020.

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, 163-178.

Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, MA: Addison-Wesley.

Rimal, R. N. & Real, K. (2003). Understanding the Influence of Perceived Norms on Behaviors. *Communication Theory*, 13(2), 184-203.

Rimal, R. N. & Real, K. (2005). How Behaviors are Influenced by Perceived Norms: A Test of the Theory of Normative Social Behavior. *Communication Research,* 32, 389-414.

Rösner, L., Winter, S., & Krämer, N. C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior*, *58*, 461-470.

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

Rowe, I. (2015). Civility 2.0: A Comparative Analysis of Incivility in Online Political Discussion. *Information, Communication & Society*, *18*, 121-138.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, 18(5), 429-434.

Sedlander, E. & Rimal, R. N. (2019). Beyond Individual-Level Theorizing in Social Norms Research: How Collective Norms and Media Access Affect Adolescents' Use of Contraception. *Journal of Adolescent Health*, 64, S31-S36.

Sherif, M. (1936). *The Psychology of Social Norms*. New York, NY: Harper & Brothers.

Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York, NY: Appleton-Century.

Sklan, A. D. (2013). @Socialmedia: Speech with a click of a button? #socialsharingbuttons. *Cardozo Arts & Entertainment Law Journal*, *32*, 377-410.

Smith, J. R. & Louis, W. R. (2008). Do As We Say and As We Do: The Interplay of Descriptive and Injunctive Group Norms in the Attitude–Behaviour Relationship. *British Journal of Social Psychology*, 47, 647-666.

Stroud, N. J., Scacco, J. M., Muddiman, A., & Curry, A. L. (2015). Changing Deliberative Norms on News Organizations' Facebook Sites. *Journal of Computer-Mediated Communication*, 20, 188-203.

Su, L. Y. F., Xenos, M. A., Rose, K. M., Wirz, C., Scheufele, D. A., & Brossard, D. (2018). Uncivil and personal? Comparing patterns of incivility in comments on the Facebook pages of news outlets. *New Media & Society*, Online First.

Sukumaran, A., Vezich, S., McHugh, M., & Nass, C. (2011). Normative Influences on Thoughtful Online Participation. *Proceedings of the International Conference on Human Factors in Computing Systems*. Vancouver, Canada.

Sydnor, E. (2019). *Disrespectful democracy: The psychology of political incivility*. New York, NY: Columbia University Press.

Van Duyn, E., & Muddiman, A. (2020). Predicting perceptions of incivility across 20 news comment sections. *Journalism*. https://doi.org/10.1177%2F1464884920907779

Weber S. and Tate, P. (2018). Civility in America 2018: Civility at work and in our public squares. Retrieved from https://www.webershandwick.com/news/civility/

Yun, G. W. & Park, S. (2011). Selective Posting: Willingness to Post a Message Online. *Journal of Computer-Mediated Communication*, 16, 201-227.