

Rossini, P. (*Forthcoming*). Beyond Incivility: Understanding Patterns of Uncivil and Intolerant Discourse in Online Political Talk. *Communication Research*.

**Beyond Incivility:
Understanding Patterns of Uncivil and Intolerant Discourse in Online
Political Talk**

Patricia Rossini
University of Liverpool

Abstract

This paper takes up the popular argument that much online discussion is toxic and hence harmful to democracy and argues that the pervasiveness of incivility is not incompatible with democratically relevant political talk. Instead of focusing on the tone of political talk, scholars interested in understanding the extent to which digital platforms threaten democratic values should focus on expressions of intolerance. I demonstrate the validity of this conceptual model by investigating the discursive and contextual features associated with incivility and intolerance online in the context of public comments in two different platforms—news websites and Facebook. Results show that incivility and intolerance occur in meaningfully different discussion settings. While incivility is associated with features that reveal meaningful discursive engagement, such as justified opinion expression and engagement with disagreement, intolerance is likely to occur in homogeneous discussions about minorities and civil society—exactly when it can hurt democracy the most.

Keywords: Political Talk, Incivility, Political Intolerance, Social Media, Online Discussion.

Introduction

Informal political discussion is a vital component of everyday life in democratic societies, one that shapes citizens' views of the world and yields meaning to matters of public concern (Conover & Searing, 2005; Mansbridge, 1999). With the ubiquitous presence of the internet in citizens' daily lives, informal political talk increasingly takes place in synchronous and asynchronous chats, forums, news websites, and on social media. In this context, concerns with the pervasiveness of incivility have led scholars to question the internet's democratic potential to foster political discussion (Rowe, 2015; Santana, 2014).

Online incivility has been operationalized in different ways, but most approaches have failed to distinguish behaviors that convey impoliteness and interpersonal disrespect from expressions that are threatening, harmful, or violate democratic norms. While some studies have focused on dimensions of incivility that fall under the scope of politeness (e.g. Coe et al., 2014), overlooking harmful behaviors, others have conflated these behaviors with racism, sexism, xenophobia and homophobia, treating incivility as a continuum (e.g. Chen, 2017). As such, little is known about the extent to which online discourse represents is inherently toxic, or whether it is simply characterized by rudeness and profanities—which may coexist with democratically desirable characteristics of discourse (Chen, 2017; Coe et al., 2014).

This study aims to address this gap by proposing an alternative approach. I argue that we need to distinguish antinormative discourse that is detrimental to democracy, which I define as intolerant, from incivility, to better understand of the nature of online discourse. *Uncivil discourse*, in which people express themselves with foul language or harsh tone, is as a rhetorical act that can serve different strategic goals in a political discussion (Benson, 2011; Herbst, 2010) and is not, in itself, incompatible with political talk. I argue that *intolerant discourse*, where

individuals or groups are attacked in ways that violate moral respect and threaten democratic pluralism (Honneth, 1996), poses a more serious threat to democracy. The latter might be expressed in ways that may not always be perceived as uncivil, and yet, is arguably more consequential and detrimental to a pluralist democracy. While these definitions are aligned with an approach to incivility as a function of the tone and the features of discourse—which is not the only way in which civility can be conceptualized (c.f. Laden, 2018)—, it aims at addressing an important empirical problem: disentangling expressions that, albeit potentially offensive or disrespectful are not inherent threats to democratic norms from those that have clearer detrimental consequences and hence undermine the value of political talk. The validity of this model is analyzed in the context of public comments in a wide range of political stories in two platforms—the Facebook page of a news organization, and the comments section of news websites. By comparing distinct platforms that are heavily used for political talk, this research takes into account how platform affordances may affect expressions of incivility and intolerance.

Results demonstrate that uncivil and intolerant discourse can be meaningfully distinguished in online political talk, and are associated with different types of discussion—while incivility occurs in comments with justified opinion expression and heterogeneous debates, intolerance is most likely to surface in homogeneous threads—indicating that many people feel more comfortable expressing extreme opinions when there is a lack of diverse expressions, which may happen either because opposing voices are not a part of the discussion, or because they are silenced. After examining discursive and contextual features associated with antinormative expressions, I argue that incivility is not necessarily incompatible with political talk online, as it does not impede participants to meaningfully engage with opposing views and

justify their perspectives. Intolerant expressions, however, are voiced precisely in conditions that make them more threatening to democratic conversations, as they tend to remain uncontested.

This study helps advance theory and research by providing a better conceptualization of interaction norms in online discussions, while taking into account how contextual discursive features and platform affordances may affect the expression of uncivil and intolerant opinions. In doing so, it argues that focusing on the function of these expressions is more important than their form to understand whether or not they can be described as harmful or toxic. Instead of looking for elusive deliberative values in online discussion, this study acknowledges the democratic value of heated or rude political exchanges, in line with scholars who have investigated strategic uses of uncivil discourse by political and media elites (Herbst, 2010; Mutz, 2015). Additionally, as most research has focused in the U.S. context, this study provides further insight about both the volume and the tone of online incivility beyond English-speaking Western democracies.

Political Discussion, Disagreement, and Democracy

Conversation is at the core of democratic citizenship, described as the capacity of listening, understanding and mediating affection, identities and individualities to build communities and negotiate conflict (Barber, 2003)—necessary conditions for a strong democracy. Although this perspective is disputed by Schudson (1997), for whom democratically relevant conversation should be oriented to problem-solving—echoing the deliberative model (Habermas, 1996)—, research suggests that informal political talk, regardless of its quality or characteristics, has intrinsic benefits for its participants (Amsalem & Nir, 2019). Political talk is an important predecessor of more sophisticated forms of political engagement, and helps participants learn about others' views, refine their own arguments, build collective identities, and

become better informed about political issues (Conover & Searing, 2005; Eveland & Hively, 2009; Moy & Gastil, 2006).

Talking about politics is particularly valued in the context of disagreement, as conversation among homogeneous groups can lead to polarization and intolerance (Mutz, 2006)—a concern echoed by early studies of online discussions (Sunstein, 2009). Exposure to heterogeneous political views helps citizens learn about others' perspectives, may promote respect towards opposing views (Mutz, 2006), increase political knowledge (Eveland & Hively, 2009), and lead to more openness to political difference (Moy & Gastil, 2006).

Online platforms such as social media may create opportunities for users to engage in cross-cutting political talk insofar as they facilitate access to political difference (Anspach, 2017). Social media users are more likely to perceive disagreement online compared to face-to-face, and to perceive disagreement on social media (Barnidge, 2017). In line with research examining face-to-face discussions, researchers have found a positive association between talking about politics online and other forms of political engagement, both on and offline (Gil de Zúñiga et al., 2014; Valenzuela et al., 2012; Valeriani & Vaccari, 2017).

Many of the benefits of political talk are associated with frequency, rather than quality or sophistication, and have been assessed through self-reported measures, challenging the view that only certain types of political discussion are democratically relevant. Nevertheless, studies examining the content of online political talk are often framed by the criteria of deliberation¹ (Freelon, 2010)—which, unsurprisingly, are not met in most online debates. Instead of looking for elusive normative ideals, this study aligns with the perspective that the search for an “online

¹ Deliberation is a form of communication oriented towards problem-solving through respectful exchange of arguments between heterogeneous individuals. It is characterized by a set of normative criteria, such as inclusion, reasoned justification, reflexivity, respect, equality and autonomy (Habermas, 1996).

public sphere” that fulfills the criteria of ideal discourse is unrealistic, and that doing so prevents scholars from understanding how different forms of expression can contribute to democratic political talk (Freelon, 2010). It is worth noting the emphasis on Habermas’ original notion of a bourgeois public sphere overlooks recent developments in deliberative theory, which has been focusing on other forms of expression, such as humor, emotions, or story-telling, as relevant forms of expression to advance deliberative goals (Parkinson & Mansbridge, 2012). While the empirical and systemic turn in deliberative theory has focused on renegotiating normative principles to understand deliberation in the wild, scholarship in online political talk continues to prioritize demanding criteria to analyze discussions.

Uncivil Discourse and Online Political Talk: A Call for Nuance

Scholars have investigated the rise (and fall) of uncivil discourse by politicians, the media, and the public, focusing on the volume and tone of discourse, its potential effects, and consequences (Herbst, 2010; Mutz, 2016; Sobieraj & Berry, 2011). Incivility is conceptually hard to define, as behaviors that may be perceived as uncivil in certain situations might be considered in other contexts (Herbst, 2010; Jamieson et al., 2017), and researchers have offered different definitions. Approaches rooted in deliberative theories approach incivility as the lack of respect for a participant in a discussion, or for their arguments, and lack of cooperation (Habermas, 1996). Empirical research tends to operationalize incivility as the tone of discourse, for instance, as involving “gratuitous asides that show a lack of respect and/or frustration with the opposition” (Mutz & Reeves, 2005, p. 5), or as “features of discussion that convey an unnecessary disrespectful tone towards the discussion forum, its participants, or its topics” (Coe et al., 2014, p. 660). Most studies investigating online discussions follow the impoliteness approach, which can be explained by the focus on content analysis. Nevertheless, as Bejan

(2017, p. 8) convincingly argues, incivility is seen as intolerable and inappropriate in a way that impoliteness is not and has been deemed as a problematic feature of online discourse.

Incivility has been measured in various ways, including behaviors that are rude, but not inherently harmful, such as messages in “all caps”, using profanities, or name-calling, while some approaches also include expressions that can be seen as detrimental to democratic discourse, such as discriminatory or hateful speech (Chen, 2017; Coe et al., 2014; Papacharissi, 2004; Santana, 2014). This conceptual fuzziness has prevented scholars from understanding the rhetorical role of uncivil discourse, and from accepting that the norms that determine what is acceptable or not are flexible and context dependent (Herbst, 2010). As a result, incivility online has been described as a sign of “low quality conversations” (Santana, 2014), “nasty” or “toxic” (Anderson et al., 2014, 2018). The view that incivility is inherently bad for online discussions has led news organizations to adopt several measures to prevent it, such as content moderation, stricter rules, identity-enforcement, as well as shutting down the comments section (Huang, 2016; Meltzer, 2015). By approaching incivility as a problematic feature by default, scholarship on online political talk fails to recognize its potential value as a legitimate form of political engagement that can yield benefits to its participants, and to society.

Findings regarding the effects of incivility online are inconclusive to support the claim that incivility is inherently harmful for political talk. Research suggests that it may influence polarization (Anderson et al., 2014; Borah, 2014), undermine trust, and affect perceptions of quality or credibility of news media (Meltzer, 2015; Prochazka et al., 2018; Thorson et al., 2010). There is evidence that exposure to uncivil posts may trigger incivility by those on the same side (Gervais, 2014), and that civil comments are seen as more persuasive than uncivil ones (Chen & Ng, 2016). Experimental research has observed a “third-person effect”: participants

assume that others will be angry or upset because of uncivil comments, but do not report being angry or upset themselves (Chen & Ng, 2017). However, most of these studies focused on specific, and often polarized, topics.

While there is evidence that exposure to incivility online may influence how people interpret information, form opinions, or participate in discussions, the effects associated with incivility are not inherently negative. Incivility can raise attention, awareness, and recall of arguments (Mutz, 2015), increase interest in politics (Brooks & Geer, 2007), and boost engagement and participation in online comments (Borah, 2014; Coe et al., 2014). Particularly on social media, uncivil discussions can be seen as entertaining (Sydnor, 2018). However, while incivility has the potential to mobilize and engage citizens who enjoy heated debates, it has the opposite effect in conflict-avoidant individuals (Sydnor, 2019). Incivility may also be used strategically to foster a sense of community and belonging among those on the same side when used to disqualify a common “enemy” (Berry & Sobieraj, 2016)—which explains its use by media pundits and politicians (Herbst, 2010). There is some evidence that incivility may become normalized for those who discuss politics online more often (Hmielowski et al., 2014). There is empirical support to the argument that participants in online discussions may have distinct expectations of civility than those who are not in these environments, reinforcing the argument that incivility is a communicative practice, “situational and contextual” (Benson, 2011, p.22).

Research examining citizens’ perceptions of uncivil behavior provides further support to this perspective. For instance, people evaluate behaviors related to personal conduct or character more uncivil than attacks to an opponent’s policy positions (Stryker et al., 2016), and citizens are less offended by uncivil discourse related to political opinions than by expressions that are offensive or demeaning towards other people (Muddiman, 2017). Name-calling and vulgarity are

seen as highly uncivil, while messages containing aspersions or pejorative tone are not (Kenski et al., 2017). Taken together, these studies signal that the perception of incivility depends not only on individual characteristics, but also on the type and target of uncivil expressions.

Scholars focused on digital media have argued that civility cannot be reduced to interpersonal politeness because the "definition ignores the democratic merit of robust and heated discussion" (Papacharissi, 2004). Papacharissi (2004) argued for a distinction between politeness and incivility, with the latter referring to threats to democratic values. Although I agree with the perspective that online political talk should not be dismissed due to "impoliteness", I argue that intolerance is a more robust concept than incivility to describe democratically threatening behaviors which, as Papacharissi (2004) has argued, should be disentangled from heated and rude expressions online.

Approaching civility as a communicative practice (Benson, 2011), I define incivility as a context-dependent feature of discourse that may convey a rude or disrespectful tone, but is not necessarily incompatible with political talk (Coe et al., 2014; Herbst, 2010; Sydnor, 2018). Following definitions offered by Coe et al. (2014), Brooks and Geer (2007), Mutz (2016), and Sydnor (2018), I operationalize uncivil discourse online as a matter of tone, not substance, of discourse. In other words, what makes a comment uncivil is a particular feature, such as the use of a vulgar word, name-calling, or potentially offensive language that, if removed, would make the same comment "civil" without changing its substance. By this definition, discourse that threatens democratic pluralism and values is beyond the scope of incivility, as it cannot be reduced to an inflammatory, emotional or heated tone. These types of discourse should instead be characterized as political intolerance, which focuses on substance (Gibson, 2007; Hurwitz & Mondak, 2002). Intolerant expressions are less susceptible to interpretation, as they offend,

derogate, silence, or undermine particular groups due to personal, social, sexual, ethnical, religious or cultural characteristics, attack individual rights, and may incite or promote violence and harm (Gibson, 2010). Political intolerance threatens moral respect—a fundamental condition for individuals to be recognized as free and equal in a pluralist democracy (Habermas, 1996; Honneth, 1996). From this perspective, hate speech, which has been largely investigated in online contexts, is a manifestation of intolerance: expressions of hatred towards groups and their members, as well as discourse that aims to humiliate, abuse, or deeply insult them (c.f. Davidson et al., 2017). I approach hate speech as a subtype of intolerance because its definition is too narrow in scope to address intolerant expressions that occur in relatively public digital spaces. Research focused on hateful speech online has often proposed other categories to classify abusive speech that does not express “hatred”, such as ‘derogatory’, ‘offensive’, ‘insults’, highlighting the limitation of the definition of hate speech to address problematic or abusive forms of online discourse (Nobata et al., 2016). Intolerance, on the other hand, includes a broader scope of discriminatory or exclusionary speech which may not be expressed as hatred, but still represent a violation of moral respect and democratic pluralism. Examples would include limiting people’s rights, undermining or silencing their participation in the public sphere (e.g. people on welfare should not be allowed to vote), as well as speech that is discriminatory or derogatory. Research on perceptions of incivility has identified some examples these behaviors (e.g., using racial slurs, threatening or encouraging harm) as “extreme” and “very uncivil” (Stryker et al., 2016), providing further evidence that intolerant expressions are consistently perceived as unacceptable in political discourse.

Distinguishing incivility from intolerance in online political talk is important for a set of reasons. First, accepting that incivility can be used to justify or emphasize political opinions,

instead of representing an inherent violation of discussion norms, means accepting that these norms are temporary and changeable. Second, standards of political civility can serve as a mechanism to silence particular forms of expression and to limit the types of discourse that are accepted in the public sphere (Benson, 2011). Accounts of protests and social movements as uncivil are examples of how those who hold power use civility to silence and delegitimize minoritarian voices (Bates, 2019; Elliot, 2019; Zerilli, 2014). Thus, calls for more online “civility” should be taken with a grain of salt. Third, this approach acknowledges that incivility is not incompatible with political talk online. While digital platforms may facilitate incivility (Hmielowski et al., 2014; Papacharissi, 2004), these expressions do not necessarily prevent political conversations to take place—for some, it may actually encourage participation (Sydnor, 2018). The same cannot be said for intolerant discourse, as it signals moral disrespect and profound disregard towards individuals or groups, and as such is incompatible with, and has the potential to damage, values of democratic pluralism that are necessary in free and equal societies (Gibson, 2007; Hurwitz & Mondak, 2002).

Platform Affordances and Online Incivility

Online incivility is associated with a discussion around the affordances of digital discussion spaces. Different platforms may allow or constrain antinormative discourse because of intrinsic affordances, such as anonymity and moderation, as well as the type of social connection they allow, if any (Ellison & boyd, 2013; Rossini & Stromer-Galley, 2019). In spite of an acknowledgement that conversational environments shape and change discussion norms, most studies on online incivility have focused on single platforms, with a few exceptions (see Halpern & Gibbs, 2013; Rowe, 2015).

The platforms chosen for this study— news websites and Facebook—are characterized by distinct affordances that may affect the quality and substance of political talk. Specifically, they differ in level of identification, moderation, and community features. Anonymity is considered one of the most relevant affordances to explain online incivility (Coe et al., 2014; Hmielowski et al., 2014)—in the absence of social cues and constraints inherent to face-to-face interaction, users may feel disconnected from their identities and are prone to express their opinions harshly without worrying about the consequences (Suler, 2004) . Studies of online deliberation have found that discussions in which participants have profiles or use real names tend to foster sincerity, rationality and higher quality of justification (Friess & Eilders, 2015).

Concerns around incivility have driven news websites to move away from complete anonymity. Most sites now require users to register, which still ensures some anonymity for participants, who can adopt pseudonyms to protect their identities from other discussants (Huang, 2016; Meltzer, 2015; Stroud et al., 2014). On Facebook, users are identified by real names and have profiles with pictures and visible social connections—features that may constrain self-expression insofar as online identities become closely connected to “real life” (Ellison & boyd, 2013). Studies have found that incivility is less frequent on Facebook (Halpern & Gibbs, 2013; Rowe, 2015) than in anonymous spaces (e.g. news sites, YouTube, blogs).

Online platforms may use moderators to enforce discussion norms and community values. In some platforms (e.g. Reddit), moderators are active discussants, who participate in threads and engage with others while incentivizing community values. News websites typically adopt “invisible” moderation—conducted by humans, and/or using automatic filters, lists of prohibited words, and algorithms—to systematically monitor and remove messages that violate rules for participation (Friess & Eilders, 2015; Stroud et al., 2014; Wise et al., 2006). On social

media, moderation practices are more complicated: platforms have rules about types of content that are subjected to moderation and rely on a set of practices (user reports, algorithms) to determine what needs to be reviewed and removed (Gillespie, 2018).

If the affordances of computer-mediated conversations facilitate incivility—making discussions uncivil “by design”—, a nuanced approach is needed to understand incivility as a communicative practice and to examine its role in online political talk. This study contributes to the growing literature in online by focusing on the contextual and discursive features that may help explain the conditions in which incivility and intolerance happens. While prior research has investigated the relationship between incivility and reason-giving (Coe et al., 2014), and incivility in different platforms (Halpern & Gibbs, 2013; Maia & Rezende, 2016; Rowe, 2015), research on more “extreme forms of incivility” has either included behaviors such as violent threats, harassment, and discrimination under a broad concept of incivility that also includes less extreme forms (e.g. Chen, 2017), or narrowly focused on hate speech in a single platform (Burnap & Williams, 2015; ElSherief et al., 2018). Thus, investigating expressions of incivility and intolerance in different platforms is needed for a better understanding of the conditions in which these types of discourse emerge online.

Starting from the premise that incivility is a rhetorical asset, based on prior findings (Coe et al., 2014; Rowe, 2015), I hypothesize that uncivil discourse is associated with the presence of disagreement because the exposure to contrasting views might elicit emotional and heated responses from those who participate in these discussions (Gervais, 2014; Sydnor, 2019).

H1.a) There is a positive relationship between disagreement and uncivil messages.

Prior research on extreme opinion expression has suggested that these views thrive in online spaces that are ideologically homogeneous and extreme (Wojcieszak, 2011). Considering

that the lack of exposure to diversity tends to lead to more intolerant and extreme views, I hypothesize that expressions of intolerance will be less likely to occur in discussions where there is disagreement.

H1.b) There is a negative relationship between disagreement and intolerant messages.

Scholars have found an association between incivility and anonymity, providing support to hypothesize different identification norms may have a significant impact in the tone of discussion (Halpern & Gibbs, 2013; Maia & Rezende, 2016). Specifically, the visibility and connectedness of Facebook may create social constraints to incivility, while users who comment on news sources preserve a high level of anonymity (Halpern & Gibbs, 2013; Rowe, 2015).

H2) Comments will be more uncivil on news websites than on Facebook.

Facebook pages of news organizations and news websites also differ in moderation, as there are no available tools for systematic moderation on Facebook by publishers. Comments on news websites are typically moderated as they affect the image of the outlet, its credibility, and may have legal implications (Huang, 2016; Thorson et al., 2010). That is the case for Portal UOL: there are strict moderation rules on the website, but comments on Facebook are not systematically moderated due to the lack of tools to handle them at scale². Given these constraints, I expect that intolerant messages will be consistently moderated on news websites because they signal a more extreme sentiment towards other people and/or groups, violence, or hate, which are likely to violate moderation rules. While these behaviors may also violate Facebook's "community standards", I expect news websites to moderate intolerance more systematically, while Facebook relies on user reports and imperfect machine-learning approaches (Gillespie, 2018). Thus,

² The editor responsible for online engagement at Portal UOL was interviewed as part of this project and discussed in detail the moderation practices adopted by the on their own websites as compared with their Facebook pages.

H3) Comments will be less intolerant on news websites than on Facebook.

Justification is a key characteristic of persuasion, as arguments that are backed by reasons are more likely to resonate than arguments that are not justified (Conover & Searing, 2005; Habermas, 1996; Mansbridge, 1999). Justification is a paramount requirement for deliberative discussions, as participants need to explain the reasons behind their opinions to be understood (Moy & Gastil, 2006; Stromer-Galley, 2007; Stromer-Galley et al., 2015). Many of the benefits associated with political discussion refer to argumentation—improving one’s own argumentative skills, understanding the arguments behind diverse viewpoints, and developing a better understanding of the problems that affect society (Eveland & Hively, 2009; Moy & Gastil, 2006; Scheufele et al., 2006). Considering uncivil discourse as a rhetorical asset mobilized to express opinions (Herbst, 2010)—one that may help opinions stand out (Mutz, 2016)—, I hypothesize that incivility is positively associated with opinion justification.

H4) Incivility is positively associated with justified opinion expression.

To the extent that expressions of intolerance undermine individuals and groups and is a mechanism of social exclusion, it is an open question whether it they are associated with justified opinion expression, as individuals may not be eager to provide defensible reasons when they discriminate against groups, nor feel like they have to if they believe others will agree. Thus:

RQ1) What is the relationship between justified opinion expression and intolerance?

Finally, most studies of online incivility have focused on a single issue or on institutional politics. One notable exception is Coe et al. (2014), who found incivility to be associated with “hard news” and sports more frequently than with topics such as health, entertainment and lifestyle. While these findings hint that topic is an important factor in determining incivility, this aspect of political talk has not been sufficiently explored, particularly beyond the U.S. context.

The following questions investigate the relationship between news topics and uncivil or intolerant discourse:

RQ2) Is there a significant relationship between the topic of the news stories and expressions of uncivil or intolerant discourse?

RQ3) Are the topics associated with uncivil discourse different than the ones associated with intolerant discourse?

Internet Use and Political Context in Brazil

Brazil is the fifth most populous country in the world with a population of more than 207 million. In 2017, 75% of the population had internet at home (IBGE, 2018). Brazilians are heavy social media users and represent Facebook's third largest market, with over 130 million users³.

Brazil is a presidential democracy with a multi-party system that has been dealing with ongoing political crises. Popular uprisings in 2013 and 2014 were triggered by dissatisfaction and mistrust in the political system, leading to a polarized presidential election in 2014 and a narrow victory of the incumbent, Dilma Rousseff. The polarized climate continued in 2015, with several public demonstrations, ultimately leading to an impeachment in 2016. These protests, along with corruption scandals, placed the political sphere on the spot, contributed to sink the President's approval rates, and to fuel public dissatisfaction. This was the political context when the data for this paper was collected.

Methods

This study compares online comments made on the Facebook page of a popular news portal, to comments left in the original sources of the stories shared on Facebook, allowing for a

³ We Are Social, & DataReportal, & Hootsuite. (July 18, 2019). Leading countries based on number of Facebook users as of July 2019 (in millions). In Statista. Retrieved October 16, 2019, from <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/>

direct comparison of comments to the same stories in two different platforms. Stories were selected from the Facebook page of the largest online news portal in Brazil, Portal UOL⁴. This page was selected because the portal hosts several media outlets, such as Folha de São Paulo—the main national newspaper—, regional newspapers, entertainment websites, and opinion blogs, which grants a variety of topics, stories and perspectives in the sample. At the time of data collection, UOL had the largest Facebook audience among Brazilian news pages, with over 6 million likes. While most studies have focused on a single news website, this approach allows for the include a more diverse set of sources in the sample, preserving the ability to compare comments to the same stories in two different platforms. Portal UOL shares similarities with other content portals, such as AOL, in the United States: UOL started as an internet provider in 1996, and, since its early days, has focused on providing diverse content from multiple sources to its users, along with web services such as email, discussion forums, and chatrooms. Today, UOL has its own newsroom, but continues to share stories from regional and national news outlets. As such, Portal UOL is not aligned to a particular media group or political ideology.

The sample of news stories was build using constructed week sampling to ensure that the variability of the media cycle is properly represented, with each day being randomly selected within the timeframe of analysis (Hester & Dougall, 2007). Following Hester and Dougall's (2007) recommendation for online news, two constructed weeks were sampled to represent a six-month period of news coverage (February to July 2015).

To compare comments to stories shared on Facebook with news sources, I followed links to each news outlet shared by Portal UOL included in the sample—including *hard* news,

⁴ Portal UOL is the largest news portal in Brazil. In April 2016, it was ranked by Alexa.com as the fifth most accessed site in the country—the first news website on the list— and is currently the 8th most accessed portal (2019).

entertainment news outlets, and political blogs⁵—and collected all public comments. While it is not possible to make inferences about the demographics of users in each platform⁶, this approach aimed at keeping the topics of discussion constant.

Portal UOL posted a total of 1,609 stories during the two constructed weeks.

DiscoverText was used to collect Facebook posts and to sample posts that shared stories about politics, based on their headlines. Stories were included in the sample if they mentioned politicians, government-related issues and policies, corruption, any of the three branches of power, as well as topics related to civic society, such as organized protests, stories about racism or discrimination, and stories about social struggles and inequalities. The assessment was made by the author, and a post-hoc intercoder agreement test was conducted on two samples⁷ of the data to ensure the reliability of the sampling criteria (Krippendorff's α : 0.79, pairwise agreement 90.4%).

After excluding nine cases in which the same story was posted on Facebook twice, this procedure identified 218 political stories. This initial set was filtered based on number of comments, excluding stories without any comments or with just one comment either on Facebook or on the source (N=29), posts with broken links to a source (N=2), and posts from sources that did not have a comments section (N=28) or used Facebook's plugin for comments (N=2). The final sample of stories (N = 157) had a total of 55,053 comments, with 70% of them being on Facebook (n = 38,594). Finally, comments were sampled using a random stratified

⁵ Because these blogs are "opinion blogs" written by journalists, and follow the moderation practices of the hosting websites, they were aggregated with other news sources.

⁶ UOL and Folha de São Paulo did not disclose demographics data.

⁷ I used two samples of 167 posts each (~20% of the initial constructed week sample of news stories) to calculate reliability. One sample was balanced to include 50% of the posts originally classified as political, and the second sample was random, with 19% of the posts originally classified as political. Krippendorff's α was calculated for each sample. Balanced sample: $\alpha = 0.75$ (pairwise agreement: 87.5%); Random sample: $\alpha = 0.80$ (pairwise agreement: 93.4%).

approach⁸, considering the proportion of comments in each platform and the number of comments on each thread (e.g. threads with 1000, 100 or 10 comments were proportionally represented). The content analysis was conducted on 12,330 comments, and all news stories were coded by theme. Because we were interested in discussion dynamics instead of random comments, we sampled comments as threads, selecting consecutive comments using a random starting point starting from a random comment. Comments were collected using Facepager, an open-source tool that interacts with Facebook's Graph API (Junger & Keyling, 2012/2019).

Content Analysis

I developed a systematic content analysis (Neuendorf, 2002) coding scheme to capture a set of dimensions of online political talk and incivility, inspired by prior research (Coe et al., 2014; Stromer-Galley, 2007). The data was coded by two independent coders. Inter-coder agreement was measured using approximately 5 percent of the sample (n = 636). All categories were reliable (Table 1).

[TABLE 1 ABOUT HERE]

The codebook (online appendix) operates with two units of analysis: news stories and messages. News stories were coded by topic: formal politics; civil society (NGOs, social movements); celebrities; minorities; public policy; international affairs (see codebook for details). To code a news story for topic, coders followed the link to the source to read the entire story and were instructed to read each story before coding its comments to be able to understand the context of the discussion. The subcategory for celebrities captured stories in which celebrities were engaged politically (e.g. cases of racism, advocacy for LGBT rights).

⁸ Confidence interval: 99%; Margin of error: 1%.

Messages were coded in the following categories: target of interaction; disagreement; opinion expression; incivility; and intolerance. Following ethical guidelines from the Association of Internet Researchers (Markham & Buchanan, 2012), examples of intolerant and uncivil messages are not provided, as individuals could be identified and exposed.

Target of interaction. This category aims at identifying comments that are direct replies to other participants in the discussion — referring directly to someone’s name or alias, responding directly to a previous speaker or arguments made previously, or using the ‘reply’ feature on Facebook.

Incivility. Following prior research (Coe et al, 2014; Mutz, 2015; Sydnor, 2018), incivility is conceptualized as a violation of discussion and social norms, as signaled by the tone (insults, vulgarity) of a message or by the style of its delivery (e.g. shouting). Messages were classified as uncivil following some of the subtypes identified by Coe et al. (2014), adapted for the purposes of this study: use of profane or vulgar language, personal attacks, lying and aspersions (combined), and attacks towards arguments or perspectives.

Intolerance. Intolerant messages are those that express a harmful—e.g. in the form of violent threats or harassment—or discriminatory intent towards people or groups based on personal characteristics, preferences, social status, and beliefs (Gibson, 1992, 2007). It also includes expressions that deny individual liberties and participation in the public sphere. Intolerant messages were classified in the following subcategories: xenophobia, racism, hate speech, violent threats, homophobia, religious intolerance, offensive stereotyping, and attacks towards gender, sexual preferences or economic status. Incivility and intolerance are not mutually exclusive categories: a message can be civil and intolerant or uncivil and tolerant.

Disagreement. Messages were coded as disagreement when they 1) diverged from the general tone of the discussion (considering the previous message(s))⁹, indicating heterogeneous opinions in the thread, or 2) explicitly diverged from another commenter in form of either name tagging or reply.

Opinion Expression. Following Stromer-Galley (2017), opinion expression is defined as judgements made by the speaker about a person, an event, a fact, a problem, or an issue.

Messages were coded in the following subcategories: 1) no opinion expression, referring to messages that did not disclose a personal opinion; 2) opinion expression, operationalized as any remark that revealed a commenter's take on a topic, without providing further explanation; and 3) justified opinion expression, which refers to opinions that are accompanied by attempts to explain or elaborate the speaker's position. Justified opinions were not judged in terms of quality of the reasons offered. For the purposes of statistical analysis, this variable was recoded as a binary (justified opinion expression = 1).

Results

Similarly to the US and the UK (Coe et al., 2014; Rowe, 2015; Santana, 2014), this study finds that uncivil discourse occurs frequently in online political talk in Brazil: 37.8% of all comments were uncivil. When bivariate variables are considered, there is a significant difference between platforms, with incivility being more frequent on Facebook, $X^2(1) = 22.06677$, $p < 0.0001$. Intolerance occurs substantially less than incivility, in only 7.8% of all messages in both platforms. There is a significant difference between platforms, with intolerance being more

⁹ Because coders analyzed sequences of comments in each story, they coded for disagreement when a comment explicitly diverged with previous messages or when participants directly disagreed from others by replying to them. For example, if two comments criticized a given political party and another commenter followed up defending the party, this message was coded as disagreement. While it cannot be assumed that users read previous comments, the strategy focuses on the presence of heterogeneous opinions.

frequent on Facebook, $X_2(1) = 99.93801$, $p < 0.0001$. Considering that these variables were coded independently, it is relevant to describe the extent to which they co-occur. Among uncivil comments, just 10.8% were also classified as intolerant, while 48% of the comments classified as intolerant were civil and 52%, uncivil.

The first hypothesis posed that incivility in public comments is predicted by the presence disagreement (H1a). By contrast, H1b posed that intolerance would have a negative relationship with disagreement. To test both hypotheses, two logistic regression models were used to examine the relationship between incivility or intolerance (as binaries) and contextual discussion features, such as disagreement, opinion expression, replies (e.g. direct responses), and the platform (Facebook vs. news sites). The topics of the news story (formal politics, civil society, minorities, policy-related topics, celebrities, international affairs — reference category being stories that did not fit under the main categories) and the character count for each comment were added to the model as control variables¹⁰. Log odds were transformed to odd ratios to facilitate interpretation, and confidence intervals are reported in addition to p-values. Results are in Table 2.

[TABLE 2 ABOUT HERE]

The model indicates that uncivil discourse can be predicted by the presence of disagreement, supporting H1a. Specifically, incivility is about 1.3 times more likely to occur when there is disagreement, and about 39% less likely to occur when users were replying to others in a discussion. However, the relationship between disagreement and intolerance is not significant, rejecting H1b.

¹⁰ Models with fewer independent variables were tested. This model was selected based on the Akaike Information Criterion (AIC) of 15,614, compared to 16,161 for a model without news topics. For intolerance, the model's AIC was 6,177, which indicates a slightly better fit than a model without news topics (AIC = 6244.3)

The second hypothesis predicted a relationship between the platform of the comment and uncivil discourse, suggesting that news stories would be more likely to drive uncivil comments than stories posted on Facebook. While differences were significant at the bivariate level, the differences between platforms are not significant in the multivariate analysis.

Starting from the premise that intolerant discourse is systematically moderated on news websites, the third hypothesis predicted that comments would be less likely to be intolerant on news sources than on Facebook. The model on Table 2 confirms this hypothesis, revealing that comments on Facebook were about 43% more likely to contain intolerant expressions than those on news websites.

The third hypothesis was tested with a logistic regression model using justified opinion as the dependent variable, with topic of news and platform included as control variables¹¹. The results in Table 3 confirm H3, showing that people were about 61% more likely to use incivility when justifying opinions. Justified opinion expression is also strongly associated with disagreement, but 37% less likely to occur when people were replying to others. The differences in platforms were also salient, with justified comments being 80% less likely to appear on Facebook. Differences by topic were significant: comments on stories about the formal political sphere, international affairs or civil society were less likely to present justified opinion expression, while comments on stories about policy-related issues were likely to be justified.

[TABLE 3 ABOUT HERE]

Table 3 reveals a negative relationship between intolerance and justified opinions, the focus of the first research question. Justified opinions were 21% less likely to be provided when comments were intolerant. The second research question inquired about the relationship between

¹¹ The full model was selected based on AIC, which indicated a better fit for the full model (AIC = 13,012) when compared to a model with no topics (AIC = 13,349)

intolerance, incivility, and the topic of the news story, and the third question asked if there were differences in the topics associated with each type of discourse. Topic has a significant impact in both types of discourse, with important differences between those associated with uncivil discourse and the ones that relate to intolerance. Discussions about formal politics, organized civil society, policy-related topics (e.g. education, violence) and international affairs were negatively associated with incivility. Conversely, intolerance was positively associated stories about minorities or civil society and activism, which were not significantly associated with uncivil behavior. Intolerance was also substantially more likely to occur in comments about international affairs and policy-related issues, which were negatively and significantly associated with incivility. In sum, for most topics, the associations with intolerance were a mirror image of those with incivility, which confirms that the two concepts are substantively different and surface in considerably different discursive contexts.

Discussion

The multivariate analysis reveals that uncivil and intolerant behavior occur in substantially different conversational contexts and are associated with distinct features of discourse. The coefficients for the two regression models were plotted in figure 1 for visualization of the distinct effects associated with incivility and intolerance.

[FIGURE 1 HERE]

Uncivil discourse online is associated with justified political expression and more likely to occur in heterogeneous discussions. Many of the benefits associated with discussing politics stem from being exposed to and engaging with cross-cutting perspectives — yet, disagreement tends to be avoided in face-to-face (Mutz, 2006; Walsh, 2004). Although incivility is positively associated with disagreement, the analysis suggests that incivility is not necessarily used to

directly attack others in a discussion— interpersonal exchanges are, in fact, significantly less likely to be uncivil. To the extent that online discussions might have greater levels of disagreement than offline conversations—and that the social constraints to face-to-face disagreement are absent—, heated heterogeneous debates in which participants exchange contrasting views should not be considered democratically dysfunctional just because they are uncivil in tone.

Moreover, uncivil discourse may help citizens express their views and stand out among the crowd — as incivility helps raise attention, improve recall, and may be seen as “entertaining” on social media (Mutz, 2016; Sydnor, 2018). The strong and positive association between incivility and justified opinion expression suggests that uncivil discourse, then, is not empty offensive shouting, and appears along attempts to elaborate opinions—a desirable characteristic of democratically relevant political talk (Conover & Searing, 2005; Mansbridge, 1999). While this study did not seek to evaluate deliberative standards, justified opinion expression is an important condition for political talk to yield positive democratic outcomes—as it is necessary for participants in a discussion to be made aware of, and to understand, diverse perspectives. Taken together, the findings suggest that online political talk, albeit sometimes uncivil, allows participants to meaningfully engage with heterogeneous opinions and present arguments to back their own positions — two characteristics that are associated with intrinsic benefits of political talk (Eveland & Hively, 2009; Scheufele et al., 2006). In other words, the fact that conversations are uncivil in tone does not mean they are not relevant in substance.

Comparing different venues for online discussion, this study shows that there are meaningful differences in the types of discourse facilitated on Facebook when compared to news websites. The finding that platform was not a significant predictor of incivility is counter-

intuitive, as prior research had suggested that these behaviors would be constrained by social media affordances—such as the semi-public nature of users’ profiles and connections (Halpern & Gibbs, 2013; Rowe, 2015). The fact that incivility prevails both in the privately public and setting of Facebook and the publicly anonymous context of news comments provides some indication that uncivil expressions might be considered acceptable by those who write them and engage in these debates, echoing prior research (Hmielowski et al., 2014). Moreover, the prevalence of incivility in the moderated setting of news websites suggest that these expressions are not universally perceived as toxic by moderators, nor violate their norms, unlike expressions of intolerance. It is relevant to note that comments are less likely to be uncivil when participants engage with others, providing further support for its strategic use in elaborating opinions instead of offending others.

Differently than incivility, intolerance is more likely to be expressed on Facebook and is not affected by disagreement. These results can be interpreted in two ways. First, Facebook is a less controlled environment than news websites, as page administrators have limited capabilities to moderate in large-scale¹², while news websites tend to be systematically moderated, which may reduce the volume of intolerant comments. Secondly, if people perceive that their opinions will be shared by others, they might be more willing to make intolerant comments. Considering that Facebook users are connected to people they know, it is possible users’ perceptions of the opinion environment around them as homogeneous may encourage intolerant expressions (Liu & Fahmy, 2011). Third, aligned with the argument that intolerant expressions are less susceptible to

¹² Interviews with editors and moderators from the main sources of news (Portal UOL and Folha de São Paulo) revealed different moderation approaches in the news websites and on Facebook—the former is systematically moderated, while the latter is not. The alleged reason not to moderate Facebook is that the platform does not provide tools to moderate in large scale other than the use of filters with lists of words to hide comments on a Facebook page. Facebook may moderate content based on reports of abuse, and by using automated approaches, but page owners and administrators do not have oversight over that process.

interpretation than incivility, these results suggest that moderation systematically suppresses expressions of racism, hate speech, violence, and the like. The reliance on user reports for content moderation might also explain why intolerance prevails on Facebook: as this study has demonstrated, this type of discourse surfaces in homogeneous debates, in which it can be assumed participants would be less likely to report it.

This study makes an important contribution by including several types of news stories instead of a single issue. While there was no positive association between topics and incivility, intolerance was associated with topics covering a broader scope of news beyond formal politics — which suggests that prior studies focused on single issues or strictly hard news might have overlooked those types of expression. Specifically, intolerant discourse is more likely to surface in stories about minority groups, e.g. LGBTQ, women, blacks, and those in social or economic disadvantage, as well as policy-related topics, international affairs, and stories related to organized civil society. What these findings suggest is that intolerance is more likely to occur precisely when and where it can hurt democracy the most, by disparaging minorities and targeting civil society groups.

By disentangling expressions widely perceived as democratically harmful from behaviors that signal harsh, vulgar or rude tone, this study suggests that the types of incivility that are more common in online discussions do not represent threats to democratic values and, as such, should not be enough to disqualify online political talk as democratically relevant. While some level of incivility might come with the territory when people discuss politics online, most discussions do not cross the boundaries of intolerant discourse and should not be treated as inherently toxic. To the contrary, some degree of incivility in informal online discussions may be acceptable to the

extent that incivility is a rhetorical asset that people mobilize to express and, in particular, to defend their views when faced with disagreement—without necessarily attacking others.

These findings have limitations. While this study extends the current literature beyond English-speaking countries and finds comparable amounts of incivility to prior work in the U.S. and the U.K., the different ways in which incivility has been measured make it challenging to compare results. Future studies need to adopt a comparative approach to investigate online incivility in different contexts and countries. Second, the content analysis did not account for the use of memes, images, and emojis, which are now embedded in political talk. Finally, the claim that incivility may serve as a rhetorical asset in online discussions is based on the characteristics of messages, not on users' perceptions. As such, I advance the hypothesis that intolerant communication is more harmful from a theoretical point of view, considering how intolerance threatens liberal democratic values, but this study did not test the effects of these behaviors may have on citizens. Future experimental projects focused on perceptions and effects of antinormative discourse online are needed to examine how both uncivil and intolerant discourse may affect those who are exposed to or targeted by it.

Conclusion

This study challenges the notion that uncivil discourse online is inherently threatening to democracy, a perspective that disregards the complexities of interpersonal communication in the digital age and does not recognize that interaction norms are flexible, and affected by the context and the nature of relationships (Benson, 2011; Herbst, 2010). Instead, I argue that it is important to disentangle expressions of incivility, which refer to tone, from expressions of intolerance, which refer to substance, to understand the conditions in which online speech is dangerous and democratically harmful from the ones it is not. Unlike incivility, which is contextual, not

inherently offensive, and subject to interpretation (Kenski et al., 2017; Stryker et al., 2016), intolerant discourse that threatens, harasses or silences people or groups, particularly based on personal or cultural characteristics undermine the positive outcomes of political talk.

Online discussions characterized by incivility are not simply shouting matches where participants offend one another. To the contrary, our results suggest that online incivility does not prevent participants from engaging in the types of political discussion deemed as important and beneficial for democracy — that is, heterogeneous conversations in which people are exposed to different views. It might be the case that those who participate in these discussions are not offended by incivility, but enticed by it: as Sydnor (2019) argues, individuals who enjoy heated debates are entertained by incivility, may become more engaged because of it, and are themselves more likely to use it. In addition to the role of platform affordances in reducing social constraints, the pervasiveness of uncivil discourse online might be a reflection of the types of people who opt to engage in these debates, who are more likely to experience positive emotions when faced with incivility and are not necessarily offended by it (Sydnor, 2019). This is not to say that incivility is necessarily positive—just that it might be more useful to treat it as a feature, and not a “bug”, to better understand its function in online political talk. Conversely, intolerant discourse online is associated with discursive and contextual features that may potentialize its harm. Intolerance is voiced in homogeneous conversations, in the absence of justified opinions. As such, it can potentially lead to more extreme opinions, contributing to silencing, segregating, and excluding minorities.

The contributions of this paper can be summarized as follows. First, it shows that incivility and intolerance can be meaningfully distinguished and are characterized by opposing features, providing empirical evidence to the need to disentangle extreme and harmful

expressions from rude discourse to better understand online political talk. While incivility surfaces in conditions that are associated to the intrinsic benefits of political talk, intolerance appears in discussions that can amplify its harmful consequences. Second, this study has demonstrated that topics can influence expressions of incivility and intolerance, the latter being more likely to be associated with news stories that feature minorities, activists and civic organizations—which is alarming insofar as it demonstrates that intolerance occurs precisely when and where it threatens democratic values the most. Third, this study contributes to expand the literature in online incivility by examining one of the most relevant economies in the Global South, contributing to the global debate on content moderation and platform governance. Recognizing that uncivil discourse is not, in itself, incompatible with political talk online is an important step to shift the contemporary discussion around content moderation from tone to substance. Considering the size of the Brazilian social media market, the results of this study may help shape future social media content moderation policies by shifting the focus from the tone of discourse to detecting and preventing more harmful types of expression.

Finally, this research makes an important contribution for advancing theory and research in incivility and online political talk by differentiating behaviors that are inherently harmful and threatening to democracy from those that are not only widely present in online conversations but are also not necessarily problematic. To fully examine the extent to which political discussions online are democratically relevant, researchers should not readily assume that incivility is problematic in itself and need instead to focus on the conditions in which it is used to understand its communicative functions. Those worried about harmful online conversations, particularly in the realm of content moderation practices and government regulation, currently being discussed globally, should focus on substance instead of tone, examining intolerant discourse, in order to

have a more productive discussion about the situations in which digital platforms should be required to take action in order to prevent discussions that may have harmful consequences for both participants and bystanders.

References

- Amsalem, E., & Nir, L. (2019). Does Interpersonal Discussion Increase Political Knowledge? A Meta-Analysis: *Communication Research*. <https://doi.org/10.1177/0093650219866357>
- 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: Crude comments and concern. *Journal of Computer-Mediated Communication*, *19*(3), 373–387. <https://doi.org/10.1111/jcc4.12009>
- Anderson, A. A., Yeo, S. K., Brossard, D., Scheufele, D. A., & Xenos, M. A. (2018). Toxic Talk: How Online Incivility Can Undermine Perceptions of Media. *International Journal of Public Opinion Research*, *30*(1), 156–168. <https://doi.org/10.1093/ijpor/edw022>
- Anspach, N. M. (2017). The New Personal Influence: How Our Facebook Friends Influence the News We Read. *Political Communication*, *34*(4), 1–17. <https://doi.org/10.1080/10584609.2017.1316329>
- Barnidge, M. (2017). Exposure to Political Disagreement in Social Media Versus Face-to-Face and Anonymous Online Settings. *Political Communication*, *34*(2), 302–321. <https://doi.org/10.1080/10584609.2016.1235639>
- Bates, K. G. (2019, March 14). *When Civility Is Used As A Cudgel Against People Of Color*. NPR.Org. <https://www.npr.org/sections/codeswitch/2019/03/14/700897826/when-civility-is-used-as-a-cudgel-against-people-of-color>

- Benson, T. W. (2011). The Rhetoric of Civility: Power, Authenticity, and Democracy. *Journal of Contemporary Rhetoric*, 1(1).
- Berry, J. M., & Sobieraj, S. (2016). *The Outrage Industry: Political Opinion Media and the New Incivility* (Reprint edition). Oxford University Press.
- 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.
<https://doi.org/10.1177/0093650212449353>
- Brooks, D. J., & Geer, J. G. (2007). Beyond negativity: The effects of incivility on the electorate. *American Journal of Political Science*, 51(1), 1–16.
- Burnap, P., & Williams, M. L. (2015). Cyber Hate Speech on Twitter: An Application of Machine Classification and Statistical Modeling for Policy and Decision Making. *Policy & Internet*, 7(2), 223–242. <https://doi.org/10.1002/poi3.85>
- Chen, G. M. (2017). *Online Incivility and Public Debate: Nasty Talk* (1st ed. 2017 edition). Palgrave Macmillan.
- Chen, G. M., & Ng, Y. M. M. (2016). Third-person perception of online comments: Civil ones persuade you more than me. *Computers in Human Behavior*, 55(Part B), 736–742.
<https://doi.org/10.1016/j.chb.2015.10.014>
- Chen, G. M., & Ng, Y. M. M. (2017). Nasty online comments anger you more than me, but nice ones make me as happy as you. *Computers in Human Behavior*, 71(Supplement C), 181–188. <https://doi.org/10.1016/j.chb.2017.02.010>
- 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.
<https://doi.org/10.1111/jcom.12104>

- Conover, P. J., & Searing, D. D. (2005). Studying ‘Everyday Political Talk’ in the Deliberative System. *Acta Politica*, 40(3), 269–283. <https://doi.org/10.1057/palgrave.ap.5500113>
- Elliot, D. (2019, March 20). “*Hear Me By Any Means Necessary*”: Charlottesville Is Forced To Redefine Civility. NPR.Org. <https://www.npr.org/2019/03/20/704902802/hear-me-by-any-means-necessary-charlottesville-is-forced-to-redefine-civility>
- Ellison, N. B., & boyd, danah. (2013). Sociality Through Social Network Sites. In W. H. Dutton (Ed.), *The Oxford Handbook of Internet Studies* (1st ed.). Oxford University Press.
- ElSherief, M., Kulkarni, V., Nguyen, D., Wang, W. Y., & Belding, E. (2018). Hate Lingo: A Target-based Linguistic Analysis of Hate Speech in Social Media. *ArXiv:1804.04257 [Cs]*. <http://arxiv.org/abs/1804.04257>
- Eveland, W. P., & Hively, M. H. (2009). Political Discussion Frequency, Network Size, and “Heterogeneity” of Discussion as Predictors of Political Knowledge and Participation. *Journal of Communication*, 59(2), 205–224. <https://doi.org/10.1111/j.1460-2466.2009.01412.x>
- Freelon, D. G. (2010). Analyzing online political discussion using three models of democratic communication. *New Media & Society*, 12(7), 1172–1190. <https://doi.org/10.1177/1461444809357927>
- Friess, D., & Eilders, C. (2015). A Systematic Review of Online Deliberation Research. *Policy & Internet*, 7(3), 319–339.
- Gervais, B. T. (2014). Incivility Online: Affective and Behavioral Reactions to Uncivil Political Posts in a Web-based Experiment. *Journal of Information Technology & Politics*, 1–19. <https://doi.org/10.1080/19331681.2014.997416>

- Gibson, J. L. (1992). The Political Consequences of Intolerance: Cultural Conformity and Political Freedom. *American Political Science Review*, 86(02), 338–356.
<https://doi.org/10.2307/1964224>
- Gibson, J. L. (2007). Political intolerance in the context of democratic theory. In R. Dalton & H.-D. Klingemann, *The Oxford Handbook of Political Behavior*. Oxford University Press.
- Gil de Zúñiga, H., Molyneux, L., & Zheng, P. (2014). Social Media, Political Expression, and Political Participation: Panel Analysis of Lagged and Concurrent Relationships. *Journal of Communication*, 64(4), 612–634. <https://doi.org/10.1111/jcom.12103>
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. Yale University Press.
- Habermas, J. (1996). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. The MIT Press.
- Halpern, D., & Gibbs, J. (2013). Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, 29(3), 1159–1168. <https://doi.org/10.1016/j.chb.2012.10.008>
- Herbst, S. (2010). *Rude democracy: Civility and incivility in American politics*. Temple University Press.
- Hester, J. B., & Dougall, E. (2007). The Efficiency of Constructed Week Sampling for Content Analysis of Online News. *Journalism & Mass Communication Quarterly*, 84(4), 811–824. <https://doi.org/10.1177/107769900708400410>
- 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(10), 1196–1211.
<https://doi.org/10.1080/1369118X.2014.899609>
- Honneth, A. (1996). *The struggle for recognition: The moral grammar of social conflicts*. MIT Press.
- Huang, C. L. (2016). *Do Comments Matter? Global Online Commenting Study 2016*. The World Association of Newspapers and News Publishers (WAN-IFRA). <http://www.wan-ifra.org/reports/2016/10/06/the-2016-global-report-on-online-commenting>
- Hurwitz, J., & Mondak, J. J. (2002). Democratic Principles, Discrimination and Political Intolerance. *British Journal of Political Science*, 32(01).
<https://doi.org/10.1017/S0007123402000042>
- IBGE. (2018). *Pesquisa Nacional por Amostra de Domicílios: Acesso à internet e à televisão e posse de telefone móvel celular para uso pessoal: 2017*. IBGE.
- Jamieson, K. H., Volinsky, A., Weitz, I., & Kenski, K. (2017). The Political Uses and Abuses of Civility and Incivility. In *The Oxford handbook of political communication*. Oxford University Press.
- Junger, J., & Keyling, T. (2018). *Facepager was made for fetching public available data from Facebook, Twitter and other JSON-based APIs* [Python]. Source code and releases available at <https://github.com/strohne/Facepager/>. (Original work published 2012)
- Kenski, K., Coe, K., & Rains, S. A. (2017). Perceptions of Uncivil Discourse Online: An Examination of Types and Predictors. *Communication Research*, 0093650217699933.
- Liu, X., & Fahmy, S. (2011). Exploring the spiral of silence in the virtual world: Individuals' willingness to express personal opinions in online versus offline settings. *Journal of Media and Communication Studies*, 3(2), 45–57.

- Mansbridge, J. (1999). Everyday Talk in the Deliberative System. In S. Macedo (Ed.), *Deliberative Politics: Essays on Democracy and Disagreement* (pp. 1–211). Oxford University Press.
- Markham, A., & Buchanan, E. (2012). *Recommendations from the AoIR Ethics Working Committee (Version 2.0)* (p. 19). Association of Internet Researchers.
- Meltzer, K. (2015). Journalistic Concern about Uncivil Political Talk in Digital News Media: Responsibility, Credibility, and Academic Influence. *The International Journal of Press/Politics*, 20(1), 85–107. <https://doi.org/10.1177/1940161214558748>
- Moy, P., & Gastil, J. (2006). Predicting Deliberative Conversation: The Impact of Discussion Networks, Media Use, and Political Cognitions. *Political Communication*, 23(4), 443–460. <https://doi.org/10.1080/10584600600977003>
- Muddiman, A. (2017). Personal and Public Levels of Political Incivility. *International Journal of Communication*, 11(0), 21.
- Mutz, D. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press.
- Mutz, D. C. (2016). *In-Your-Face Politics: The Consequences of Uncivil Media*. Princeton University Press.
- Mutz, D., & Reeves, B. (2005). The New Videomalaise: Effects of Televised Incivility on Political Trust. *American Political Science Review*, 99(01). <https://doi.org/10.1017/S0003055405051452>
- Neuendorf, K. A. (2002). *The Content Analysis Guidebook* (1st edition). SAGE Publications, Inc.

- Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). Abusive Language Detection in Online User Content. *Proceedings of the 25th International Conference on World Wide Web - WWW '16*, 145–153. <https://doi.org/10.1145/2872427.2883062>
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2), 259–283. <https://doi.org/10.1177/1461444804041444>
- Parkinson, J., & Mansbridge, J. J. (2012). *Deliberative systems: Deliberative democracy at the large scale*.
- Prochazka, F., Weber, P., & Schweiger, W. (2018). Effects of Civility and Reasoning in User Comments on Perceived Journalistic Quality. *Journalism Studies*, 19(1), 62–78. <https://doi.org/10.1080/1461670X.2016.1161497>
- Rossini, P., & Stromer-Galley, J. (2019). Citizen Deliberation Online. In E. Suhay, B. Grofman, & A. Tresel (Eds.), *Oxford Handbook of Electoral Persuasion*. Oxford University Press.
- Rowe, I. (2015). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, Communication & Society*, 18(2), 121–138. <https://doi.org/10.1080/1369118X.2014.940365>
- Santana, A. D. (2014). Virtuous or Vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice*, 8(1), 18–33. <https://doi.org/10.1080/17512786.2013.813194>
- Scheufele, D. A., Hardy, B. W., Brossard, D., Waismel-Manor, I. S., & Nisbet, E. (2006). Democracy Based on Difference: Examining the Links Between Structural Heterogeneity, Heterogeneity of Discussion Networks, and Democratic Citizenship.

- Journal of Communication*, 56(4), 728–753. <https://doi.org/10.1111/j.1460-2466.2006.00317.x>
- Schudson, M. (1997). Why conversation is not the soul of democracy. *Critical Studies in Mass Communication*, 14(4), 297–309. <https://doi.org/10.1080/15295039709367020>
- Sobieraj, S., & Berry, J. M. (2011). From Incivility to Outrage: Political Discourse in Blogs, Talk Radio, and Cable News. *Political Communication*, 28(1), 19–41. <https://doi.org/10.1080/10584609.2010.542360>
- Stromer-Galley, J. (2007). Measuring deliberation's content: A coding scheme. *Journal of Public Deliberation*, 3(1), 12.
- Stromer-Galley, J., Bryant, L., & Bimber, B. (2015). Context and Medium Matter: Expressing Disagreements Online and Face-to-Face in Political Deliberations. *Journal of Public Deliberation*, 11(1), 1.
- Stroud, N. J., Scacco, J. M., Muddiman, A., & Curry, A. L. (2014). Changing Deliberative Norms on News Organizations' Facebook Sites. *Journal of Computer-Mediated Communication*, 20, 188–203. <https://doi.org/10.1111/jcc4.12104>
- Stryker, R., Conway, B. A., & Danielson, J. T. (2016). What is political incivility? *Communication Monographs*, 83(4), 535–556. <https://doi.org/10.1080/03637751.2016.1201207>
- Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior*, 7(3), 321–326. <https://doi.org/10.1089/1094931041291295>
- Sunstein, C. R. (2009). *Going to extremes: How like minds unite and divide*. Oxford University Press.

- Sydnor, E. (2018). Platforms for Incivility: Examining Perceptions Across Different Media Formats. *Political Communication*, 35(1), 97–116.
<https://doi.org/10.1080/10584609.2017.1355857>
- Sydnor, E. (2019). *Disrespectful Democracy: The Psychology of Political Incivility*. Columbia University Press.
- Thorson, K., Vraga, E., & Ekdale, B. (2010). Credibility in Context: How Uncivil Online Commentary Affects News Credibility. *Mass Communication and Society*, 13(3), 289–313. <https://doi.org/10.1080/15205430903225571>
- Valenzuela, S., Kim, Y., & Gil de Zúñiga, H. (2012). Social Networks that Matter: Exploring the Role of Political Discussion for Online Political Participation. *International Journal of Public Opinion Research*, 24(2), 163–184. <https://doi.org/10.1093/ijpor/edr037>
- Valeriani, A., & Vaccari, C. (2017). Political talk on mobile instant messaging services: A comparative analysis of Germany, Italy, and the UK. *Information, Communication & Society*, 0(0), 1–17. <https://doi.org/10.1080/1369118X.2017.1350730>
- Wise, K., Hamman, B., & Thorson, K. (2006). Moderation, Response Rate, and Message Interactivity: Features of Online Communities and Their Effects on Intent to Participate. *Journal of Computer-Mediated Communication*, 12(1), 24–41.
<https://doi.org/10.1111/j.1083-6101.2006.00313.x>
- Wojcieszak, M. (2011). Pulling Toward or Pulling Away: Deliberation, Disagreement, and Opinion Extremity in Political Participation*. *Social Science Quarterly*, 92(1), 207–225.
- Zerilli, L. M. G. (2014). Against Civility: A Feminist Perspective. In A. Sarat (Ed.), *Civility, Legality, and Justice in America* (pp. 107–131). Cambridge University Press; Cambridge Core. <https://doi.org/10.1017/CBO9781107479852.005>

Table 1. Pairwise-agreement and Krippendorff's Alpha

Category	Facebook α	News α	Combined α
Topic (of the story)	0.69	0.79	0.72
Disagreement	0.82	0.90	0.86
Incivility	0.79	0.88	0.83
Target of incivility	0.85	0.88	0.86
Intolerance	0.90	0.85	0.88
Target of intolerance	0.82	0.82	0.82
Opinion expression	0.74	0.91	0.83
Reply	0.78	0.96	0.87
	N=413	N=223	N=636

Note: Intercoder agreement was measured for each platform, and for the total sample.

Table 2. Logistic regression predicting uncivil and intolerant discourse

	Uncivil Comment		Intolerant Comment	
	O.R. ($e\beta$)	C.I. [2.5%, 97.5%]	O.R. ($e\beta$)	C.I. [2.5%, 97.5%]
Constant	0.64 ***	[0.49, 0.82]	0.05 ***	[0.03, 0.08]
Disagreement	2.29 ***	[1.98, 2.65]	0.94	[0.72, 1.24]
Reply	0.61 ***	[0.54, 0.69]	0.83	[0.66, 1.05]
Facebook (vs. news)	1.02	[0.93, 1.13]	1.42 ***	[1.16, 1.74]
Story topic				
Politics	0.77 *	[0.61, 0.98]	0.45 **	[0.28, 0.73]
Minorities	1.15	[0.90, 1.47]	2.78 ***	[1.73, 4.49]
Policy	0.31 ***	[0.24, 0.40]	1.71 *	[1.05, 2.80]
International	0.40 ***	[0.28, 0.57]	3.96 ***	[2.29, 6.83]
Civ. Society	0.53 ***	[0.37, 0.76]	1.95 *	[1.07, 3.57]
Celebrities	0.87	[0.58, 1.31]	1.59	[0.77, 3.27]
N. Characters	1.00 ***	[1.00, 1.00]	1.00 ***	[1.00, 1.00]
Pseudo R2	0.08		0.12	

*Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.*

Table 3. Logistic regression predicting justified opinion expression.

	O.R. (e^{β})	C.I. (2.5%)	C.I. (97.5%)
(Intercept)	1.14	0.86	1.50
Intolerance	0.80 **	0.68	0.94
Incivility	1.62 ***	1.480	1.76
Reply	0.64 ***	0.562	0.74
Disagreement	3.32 ***	2.848	3.88
Facebook (vs. news site)	0.20 ***	0.185	0.23
Story topic			
Politics	0.51 ***	0.39	0.66
Minorities	1.12	0.84	1.48
Policy	1.54 **	1.15	2.05
Civ. Society	0.66 *	0.44	0.99
Celebrities	0.68	0.43	1.08
International	0.40 ***	0.26	0.62
Pseudo R2	0.17		

Note: p = 0 '***' 0.001 '**' 0.01 '*' 0.05

Figure 1. Visualization of Regression Coefficients