Future directions for online incivility research
Masullo, Gina M.

Erstveröffentlichung / Primary Publication
Sammelwerksbeitrag / collection article

Digital
Communication
Research.de

Abstract: This chapter makes a normative argument that incivility scholars should shift directions in exploring aversive online communication. Specifically, it is vital for scholars to consider various subsets of incivility (e.g., profanity or hate speech), rather than treat incivility as a monolith and to acknowledge that different types are not equally damaging to democracy or interpersonal relations. Furthermore, this chapter calls for more attention to how incivility of all types hurts those from marginalized groups and how and why those with less societal power are more frequent targets of toxicity, as well as how to protect them. It also proposes that the role of online platforms, like Facebook, WeChat, and WhatsApp, be integorrated more fully in regard to incivility and that incivility be studied in concert with other types of problematic speech, such as misinformation and disinformation.

Gina M. Masullo

Future Directions for Online Incivility Research

1 Rethinking incivility research

Ask average people what they think of online comments on news websites or social media, and soon enough you are likely to hear the common adage: "Don't read the comments." The thinking is that online comments are vast cesspools of vitriol. The implication is that the incivility that peppers these comments renders them useless for political discourse. Scholars all too often have adopted a similar approach, highlighting a clear normative assumption that "incivility is bad and should be eliminated" (Chen, Muddimann, et al., 2019, p. 1).

This chapter puts forth a theoretical argument to challenge this notion. While incivility is certainly rife online (e.g., Chen, 2017; Coe et al., 2014), clear evidence exists that various types of incivility are not equally damaging to political discussions (Rossini, 2020) or perceived as such (Stryker et al., 2016). Indeed, when incivility is defined as impoliteness, research suggests it may actually jump-start political engagement at least in the short term (Chen, 2017; Chen & Lu, 2017) even as it foments negative emotions (e.g., Gervais, 2015; Rösner et al., 2016). Thus, understanding the role and influence of incivility in online discussions is decidedly complicated. As the study of incivility has progressed and our understanding has grown, we must shift to asking new questions, considering different outcome

variables, and abandoning old assumptions. I put forth two principles that undergird the approach that I am calling for: Specify what type of incivility you are considering, rather than treating incivility as a monolith, and realize that different types of incivility are not equally damaging to democracy or interpersonal relations (Stryker et al., 2016; Chen, Muddimann, et al., 2019).

Keeping these two principles in mind, I call for an expanded online incivility research agenda with a broader vantage point. Instead of asking—what is the influence of uncivil discourse? – we should ask more specific questions. How do different types of uncivil content, such as pejorative speech versus profanity, differ in their harm? Instead of assuming incivility will have negative effects, we should consider positive outcomes such as a boost in political engagement (e.g., Chen, 2017; Chen & Lu, 2017) and parse how they vary between different subsets of people or across platforms. Rather than focus our inquiry on the more common types of aversive speech, profanity and insults (Chen, 2017; Coe et al., 2014), we should delve more deeply into the effects of the less frequent but more antagonistic types of discourse, such as hateful or intolerant speech (Rossini, 2020). I argue that hate speech fits under the umbrella term of what scholars and the public label incivility (see Chen, 2017, for an overview of this arugment), although certainly some scholars see hate speech as conceptually distinct (see Paasch-Colberg et al., 2021, for an overview; see Sponholz and Frischlich in this volume, for a discussion of the hate speech concept).

Specifically, I urge that incivility research be expanded in three main areas, which I will discuss below. First, more attention should be paid to how incivility of all types hurts those from marginalized groups and how and why those with less societal power are more frequent targets of toxicity, as well as how to protect them. This approach puts more research emphasis on hate speech, arguably the most virulent type of incivility, rather than impoliteness, the least antagonistic type. Second, we should interrogate the role of online platforms, such as Facebook, Twitter, WeChat, and WhatsApp, in managing uncivil attacks. Finally, more research should probe how incivility intersects with and perhaps amplifies other problematic forms of online communication, such as misinformation and disinformation. While misinformation and disinformation are clearly problematic types of communication, I argue they are conceptually distinct from incivility because their most potent effect is in misleading the public, which is not part of incivility. Incivility and false information warrant study together because they are arguably the two major issues that trouble scholars and the public alike about online discussions.

2   Online incivility is not a monolith

First, I will unpack the two principles I have outlined that should be foundational to online incivility research going forward. Then I will examine the three areas for expanded research in greater detail. The first principle I put forth is that scholars should examine specific types of aversive online content, rather than treat incivility as a monolith. There is great debate in the literature over what constitutes uncivil communication although most definitions fall into two main camps (see Bormann & Ziegele in this volume, for an overview of different incivility conceptions). In the impoliteness camp, incivility is defined as profanity, name-calling, and insults (Chen, 2017; Muddiman, 2017; Rossini, 2020). This approach often relies on impoliteness or face theories, which argue that uncivil speech threatens people's constructed sense of self, called face, leading to emotional pain (Brown & Levinson, 1987; Goffman, 1959; Metts & Cupach, 2015). The other approach defines incivility more virulently as threats to democracy; stereotyping; or racists, sexist, xenophobic, and homophobic communication (Papacharissi, 2004), a subcategory that Rossini (2020) calls "intolerant discourse" (p. 2). This type of incivility is rooted in the theory of deliberative democracy, which relies on the normative ideal that discussions across differences should be rational, respectful exchanges that seek to reach consensus and are, therefore, valuable in a democracy (Fishkin, 1991; Gutmann & Thompson, 1996; Jacobs et al., 2009). These ways of considering incivility offer distinct theoretical and operational approaches that should not be conflated. Even more importantly, other types of content that might fit definitions of incivility—such as lying accusations (Kenski et al., 2018), hyperbole and distortion (Gervais, 2015; 2017), and lack of political compromise (Muddiman, 2017) – that have received scant study should receive more attention. Indeed, Stryker et al. (2016) considered 23 types of behavior or speech that might be considered uncivil, including vulgarity, refusing to listen to others, and shouting, and found that slurs and threatening or encouraging harm were perceived as most uncivil. Yet, study after study, including some of my own (e.g., Chen, 2017; Chen & Lu, 2017), focus on forms of rudeness (e.g., Lee et al., 2019; Rösner et al., 2016), the type of incivility that Stryker et al. (2016) found was perceived as less damaging to political discourse. Our efforts should move to a multi-dimensional approach to incivility when possible (e.g., Chen et al., 2020; Oz et al., 2018; Ozler et al., 2020; Ziegele et al., 2020; see also Bormann & Ziegele

in this volume). I would even go so far as to suggest we should stop calling every-thing incivility and, rather, use the specific terms (e.g., "hate speech" or "profani-ty") that more narrowly pertain to what we mean. Our research then should focus on these more extreme types of speech that offer more troubling implications (e.g., Chen, Fadnis, & Whipple, 2019; Murthy & Sharma, 2019; Paasch-Colberg et al., 2021). Of course, from a practical standpoint, we may need to retain the concept of incivility as an umbrella term for all these types of communication and as a theoretical perspective, but, when possible, we should use more specific terminology. For example, studies that examine specific types of speech, such as obsencity or politically motivated hate speech (e.g., Bodrunova et al., 2021), offer more knowledge than those that aim to tackle *incivility* in general.

In the earlier days of incivility research, methodological issues may have led to reductive operational definitions for incivility. For example, experiments re-quire that researchers focus only on few forms of incivility because manipulating several types of incivility would require exhaustively large sample sizes. Techno-logical limitations initially meant that efforts to automate detection of incivility using machine learning could only capture less-nuanced attributes of incivility, such as profanity (e.g., Lee et al., 2019). Human coders were often used to better detect subtle uncivil attributes (Guo et al., 2016). However, this was expensive and time-consuming and limited the amount of content that could be reasonably analyzed (Muddiman et al., 2019). Newer approaches, such as using manually val-idated organic dictionaries (Guo et al., 2016; Muddiman et al., 2019), and machine learning models have shown success in detecting multi-faceted forms of incivil-ity across various domains, such as comments and tweets (Davidson et al., 2020; Ozler et al., 2020), although many still misclassify complex types of incivility (e.g., Stoll et al., 2020). Even when human coders are employed for smaller datasets (e.g., Chen et al., 2020; Oz et al., 2018) or in combination with automated coding (Kenski et al., 2018), efforts should be made to consider and compare different attributes of incivility. Experiments, of course, may still focus on only several types of incivility, and that is fine, as long as these types are identified and some experiments delve into the more virulent types of incivility.

These methodological issues suggest that we need to expand how we explore in-civility, employing both quantitative and qualitiative approaches to provide a fuller understanding of how the public perceives incivility and the effects it has. In-depth interviews with the public or with content moderators can provide insights into

how incivility is perceived and identified, for example, in ways where quantitative coding may fall short. Theoretical work can forge connections between disparate quantitative or qualitative studies that is woefully lacking.

3   All incivility is not equally harmful

The second underlying principle is that different types of online incivility are not equally harmful. When communication is considered dichotomously—either uncivil or civil—there can be a tendency to view one as always good and one as always bad. This is a flawed assumption (Chen, Muddimann, et al., 2019; Rossini, 2020). We need more research that asks questions that tap into more subtle questions: Under what conditions is incivility harmful? What attributes of incivility are more harmful than others? Are there situations where civility might not be beneficial? We need more research that offers insight into how people experience different types of incivility and what it means for public discourse. What effect on political discourse or emotions does profanity have that differ from the effects of threats to democracy or stereotyping? Are there particular subgroups of the population that encounter greater or lesser effects? These questions need to be addressed. Rather than look at incivility as always harmful and civility as always righteous, we need to understand the overlap between good and bad. Deciding what is civil or uncivil, Herbst (2010) argues, is a "strategic tool or weapon in politics" (p. 6), such that those in power can squelch speech they disagree with by labeling it as uncivil. It is temporary and changeable and fluid across contexts, she posits (see Litvinenko in this volume, for a similar argument). If that is true, and I believe it is, clearly incivility is also malleable. Yet, the literature often assumes that online incivility by default is harmful. In addition, we need more research outside the United States and other western democracies. We have limited evidence that people perceive types of incivility differently across cultures and countries (Weber et al., 2020), but more cross-cultural work in this area is direly needed to have a fuller understand of online incivility's effects on discourse and society.

4   Marginalized groups

I have outlined the two main principles that should be kept in focus in incivility research going forward. Now I shift my attention to the three areas that I suggest need greater attention. The first is that we need more research about how online incivility of all types disproportionately targets those from marginalized groups, and, even more importantly, how people from these groups can be more protected online. For example, we know that women and people of color and other marginalized groups are frequent targets of incivility online (Chen, Fadnis, & Whipple, 2019; Chen et al., 2020; Edström, 2016; Ferrier, 2018; Pain & Chen, 2019; Searles et al., 2020; Sobieraj, 2020; Southern & Harmer, 2021), but we know little about what interventions are most effective to safeguard them. Journalists provide a cogent example. We know that female journalists face a barrage of online attacks (Chen et al., 2020; Searles et al., 2020) from threatening messages to the unauthorized release of private information, and that this sometimes compels them to change how they tell stories or even consider leaving the profession (Ferrier, 2018). Politicians offer another notable example. While all politicians run the risk of being attacked online, barbs against people identifying as female are more likely to be gendered or stereotypical assaults on their identity (Southern & Hamer, 2021). But it is not just journalists or politicians. All women who dare to participate in online public discourse, particularly about politics, face the threat of violent resistance: "The abuse targets their identities, pummeling them with rape threats, attacks on their appearance, and presumed sexual behavior, and a cacophony of misogynistic, racist, xenophobic, and homophobic stereotypes and epithets" (Sobieraj, 2020, p. 4). Expanding incivility research to more countries and cultures will help address this in some ways, but we also need more studies that specifically focus on marginalized groups, such as people of color, refugees, or LGBTIA+ people. We need more research into how the digital space can be changed or managed at a structural level to prevent this. What roles should newsrooms, platforms, or government play in solving these problems? Can existing laws be better employed or are new laws necessary? How can newsrooms help their employees feel safe? This is important because understanding how to protect the marginalized online will help achieve a more user-friendly digital space for everyone. The strategies that work for those with less power in society will improve discourse for all.

5   Role of platforms

The second line of research that I argue deserves more focused attention is what role platforms should play in managing uncivil attacks online. Currently, a patchwork of efforts aim to ensure a productive commenting space on social media and news websites. Moderators police content, and these efforts improve discussions (Ksiazek, 2018; Masullo et al., 2020; Stroud et al., 2015), but the task is emotionally exhausting (Riedl et al., 2020). Users flag unseemly content (Naab et al., 2018) or even dialog with commenters in hopes of improving discussions (Ziegele et al., 2020). Platforms and news organizations are performative by telling users in advance through terms of service or online posts what type of behavior and content will be permitted (Gillespie, 2018). Yet, despite all these efforts, calls are frequent that more should be done (see Sobieraj et al., 2020, for related arguments). Are platforms or newsrooms responsible for ensuring a robust democratic discourse can take place on their sites? Is it right or ethical for privately owned companies to take on this role? Should governments regulate platforms as public utilities to ensure they do this task? Will that improve discussions? Does that force platforms into a role they shouldn't have? How would that even work, considering many platforms cross national boundaries? These questions need research-based answers. This is particularly true at this current moment as some social media platforms took the unprecedented step of banning former U.S. President Donald J. Trump because his combative posts culminated with a mob of his supporters violently attempting to prevent Congress from certifying Joseph Biden as the victor over Trump (Denham, 2021). Regardless of how scholars feel about this particular banning, the banning raised urgent questions about internet governance and the role and power of social media platforms in contemporary lives and highlights the need for more study in this area. It leads to important questions, such as: Who really controls speech? What entities should have the power to govern online discourse? What are the ramifications of banning politicians, or others, from engaging online? What are the ethical and legal questions surrounding such bans? All are ripe area for more inquiry.

6   Online incivility and other problematic discourse

The third and final avenue for research on online incivility that I propose is understanding the intersection between incivility and false information online. These two concepts should be studied together because they are arguably the two major issues that scholars and the public raise about online discourse, and we know that in general incivility can undermine the persuasiveness of communication (Chen & Ng, 2016; Jenkins & Dragojevic, 2013). Yet, for the most part incivility and false information are treated as separate research streams. The problem of misinformation, the unintentional spread of false information, and disinformation, the purposeful spread of inaccurate communication (Tandoc et al., 2018), are focal concerns in the public consciousness. We know correcting misperceptions from faulty information, whether purposeful or not, is challenging (Nyhan & Reifler, 2010) but possible (e.g., Bode & Vraga, 2015), but we know less about how incivility may influence how people process false information or corrections of that information. This is important to understand because, given the free-wheeling discourse online, it seems logical the people may come across uncivil misinformation or disinformation or acerbic corrections to this false information. Bode et al. (2020) have shown in an experiment that corrections to misleading information about the safety of raw milk on Facebook are effective regardless of whether the tone of the correction is uncivil, affirmational, or neutral, suggesting tone is not the driving force in whether people embrace or reject a correction to misinformation. But more research in this area is warranted. Kim and Chen (2021) demonstrated that angry emoticons on social media comments that attempt to correct misinformation altered how those messages are perceived. Yet, so many questions remain unanswered. Do people discount misinformation that is uncivil or is it more arousing and, therefore, more powerful? Do people reject or embrace uncivil disinformation that outrages them morally, such as accusations that politicians are not telling the truth? Does how people respond to these messages depend on whether an out-group or in-groups is spreading the falsehood or whether a person realizes the message is not true? Does this vary based on what type of false information is considered? Given the monumental concern that false information and incivility present online, it is vital to understand more about how these two concepts intersect.

7   Going forward

In summary, there is a fruitful path ahead for the study of online incivility. While the topic has received a plethora of study, there are holes in the literature that scholars should fill. We have a firm foundation of incivility research at this point, but research going forward should focus more specifically on differences between various types of incivility, rather than treat it as a monolith. Also, the days of seeing incivility as always normatively bad and civility as always normatively good should be over, I urge. We need to consider different types of harm for different types of incivility, and also leave open the idea that incivility may have benefits even if they are unintended. Research going forward should look to finding solutions, not just illuminating problems. In particular, we need more work on how the digital space can be improved so that it is safer for those from marginalized groups. We need more study into how platforms can and should manage incivility and what ramifications their actions have on the larger society. Finally, we need to consider incivility in concert with other forms of aversive online communication, such as misinformation and disinformation.

These trajectories of research will offer many benefits. First, they will bring a richer, more nuanced understanding of how online incivility operates and its effects. It will help us theorize more about the role of incivility in society, and it will help us solve problems related to acerbic online communication more broadly. I cannot imagine a day anytime soon when online communication will disappear. If anything, we likely will be communicating more and more online than we do today. That means online incivility is with us in the future, so the need to bring incivility research to a new level is particularly important.

Gina M. Masullo is Associate Professor in the School of Journalism and Media and Associate Director of the Center for Media Engagement, both at The University of Texas at Austin, USA. https://orcid.org/0000-0002-4909-2116

References

Bode, L., & Vraga, E. K. (2015). In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication*, *65*(4), 619–638. https://doi.org/10.1111/jcom.12166

Bode, L., Vraga, E. K., & Tully, M. (2020). Do the right thing: Tone may not affect correction of misinformation on social media. *Harvard Kennedy School Misinformation Review.* https://doi.org/10.37016/mr-2020-026

Bodrunova, S. S., Litvinenko, A., Blekanov, I., & Nepiyushchikh, D. (2021). Constructive aggression? Multiple roles of aggressive content in political discourse on Russian YouTube. *Media and Communication*, *9*(1), 181–194. https://doi.org/10.17645/mac.v9i1.3469

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage.* Cambridge University Press.

Chen, G. M. (2017). *Online incivility and public debate: Nasty talk.* Palgrave Macmillan.

Chen, G. M., Fadnis, D., & Whipple, K. (2019). Can we talk about race? Exploring online comments about race-related shootings. *Howard Journal of Communications*, *31*(1), 35–49. https://doi.org/10.1080/10646175.2019.1590256

Chen, G. M., & Lu, S. (2017). Online political discourse: Exploring differences in effects of civil and uncivil disagreement in news website comments. *Journal of Broadcasting & Electronic Media, 61*(1), 108–125. https://doi.org/10.1080/08838151.2016.1273922

Chen, G. M., Muddiman, A., Wilner, T., Pariser, E., & Stroud, N. J. (2019). We should not get rid of incivility online. *Social Media + Society*, *5*(3). https://doi.org/10.1177/2056305119862641

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., Pain, P., Chen, V., Mekelburg, M., & Springer, N. (2020). 'You really have to have a thick skin': A cross-cultural perspective on how online harassment influences female journalists. *Journalism*, *21*(7), 877–895. https://doi.org/10.1177/1464884918768500

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

Davidson, S., Sun, Q., & Wojcieszak, M. (2020). Developing a new classifier for automated identification of incivility in social media. *Proceedings of the Fourth Workshop on Online Abuse and Harms*, 95–101. Association for Computational Linguistics. https://www.aclweb.org/anthology/2020.alw-1.12

Denham, H. (2021, January 14). These are the platforms that have banned Trump and his allies. *The Washington Post.* https://www.washingtonpost.com/technology/2021/01/11/trump-banned-social-media/

Edström, M. (2016). The trolls disappear in the light: Swedish experiences of mediated sexualised hate speech in the aftermath of Behring Breivik. *International Journal for Crime, Justice and Social Democracy*, *5*(2), 96–106. https://doi.org/10.5204/ijcjsd.v5i2.314

Ferrier, M. (2018). Attacks and harassment: The impact on female journalists and their reporting. International Women's Media Foundation. https://www.iwmf.org/attacks-and-harassment/

Fishkin, J. S. (1991). *Democracy and deliberation: New directions for democratic reform.* Yale University 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), 167–185. https://doi.org/10.1080/19331681.2014.997416

Gervais, B. T. (2017). More than mimicry? The role of anger in uncivil reactions to elite political incivility. *International Journal of Public Opinion Research*, *29*(3), 384–405. https://doi.org/10.1093/ijpor/edw010

Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media.* Yale University Press.

Goffman, E. (1959). *Presentation of self in everyday life.* Doubleday.

Guo, L., Vargo, C. J., Pan, Z., Ding, W., & Ishwar, P. (2016). Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling. *Journalism & Mass Communication Quarterly*, *93*(2), 332–359. https://doi.org/10.1177/1077699016639231

Gutmann, A., & Thompson, D. (1996). *Democracy and disagreement.* Harvard University Press.

Herbst, S. (2010). *Rude democracy: Civility and incivility in American politics.* Temple University Press.

Jacobs, L. R., Cook, F. L., & Delli Caprini, M. X. (2009). *Talking together: Public deliberation and political participation in America.* The University of Chicago Press.

Jenkins, M., & Dragojevic, M. (2013). Explaining the process of resistance to persuasion: a politeness-theory based approach. *Communication Research,* *40*(4), 559–590. https://doi.org/10.1177/0093650211420136

Kenski, K., Filer, C. R., & Conway-Silva, B. A. (2018). Lying, Liars, and Lies: Incivility in 2016 presidential candidate and campaign tweets during the invisible primary. *American Behavioral Scientist*, *62*(3), 286–299. https://doi.org/10.1177/0002764217724840

Kim, J. W., & Chen, G. M. (2021). Exploring the influence of comment tone and content in response to misinformation in social media news. *Journalism Practice*, *15*(4), 456–470. https://doi.org/10.1080/17512786.2020.1739550

Ksiazek T. B. (2018). Commenting on the news. *Journalism Studies, 19*(5), 650–673. https://doi.org/10.1080/1461670X.2016.1209977

Lee, F. L. F., Liang, H., & Tang, G. K. Y. (2019). Online incivility, cyberbalkanization, and the dynamics of opinion polarization during and after a mass protest event. *International Journal of Communication*, *13*, 4940–4959.

Masullo, G. M., Riedl, M. J., & Huang, Q. E. (2020). Engagement moderation: What journalists should say to improve online discussions. *Journalism Practice*. Advance online publication. https://doi.org/10.1080/17512786.2020.1808858

Metts, S., & Cupach, W. R. (2015). Face theory: Goffman's dramatist approach to interpersonal interaction. In D. O. Braithwaite & P. Schrodt (Eds.), *Engaging theories in interpersonal communication* (pp. 203–214). Sage.

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

Muddiman, A., McGregor, S. C., & Stroud, N. J. (2019). (Re)Claiming our expertise: Parsing large text corpora with manually validated and organic dictionaries. *Political Communication*, *36*(2), 214–226. https://doi.org/10.1080/10584609.2018.1517843

Murthy, D., & Sharma, S. (2019). Visualizing YouTube's comment space: Online hostility as a networked phenomena. *New Media & Society*, *21*(1), 191–213. https://doi.org/10.1177/1461444818792393

Naab, T. K., Kalch, A., & Meitz, T. G. (2018). Flagging uncivil user comments: Effects of intervention information, type of victim, and response comments on bystander behavior. *New Media & Society*, *20*(2), 777–795. https://doi.org/10.1177/1461444816670923

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*(2), 303–330. https://doi.org/10.1007/s11109-010-9112-2

Oz, M., Zheng, P., & Chen, G. M. (2018). Twitter versus Facebook: Comparing incivility, impoliteness, and deliberative attributes. *New Media & Society, 20*(9), 3400–3419. https://doi.org/10.1177/1461444817749516

Ozler, K. B., Kenski, K., Rains, S., Shmargad, Y., Coe, K., & Bethard, S. (2020). Fine-tuning for multi-domain and multi-label uncivil language detection. *Proceedings of the Fourth Workshop on Online Abuse and Harms*, 28–33. Association for Computational Linguistics. https://www.aclweb.org/anthology/2020.alw-1.4

Paasch-Colberg, S., Strippel, C., Trebbe, J., & Emmer, M. (2021). From insult to hate speech: Mapping offensive language in German user comments on immigration. *Media and Communication*, *9*(1), 171–180. https://doi.org/10.17645/mac.v9i1.3399

Pain, P., & Chen, V. (2019). This reporter is so ugly, how can she appear on TV? *Journalism Practice*, *13*(2), 140–158. https://doi.org/10.1080/17512786.2017.1423236

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

Riedl, M. J., Masullo, G. M., & Whipple, K. N. (2020). The downsides of digital labor: Exploring the toll incivility takes on online comment moderators. *Computers in Human Behavior*, *107*, Article 106262. https://doi.org/10.1016/j.chb.2020.106262

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. https://doi.org/10.1016/j.chb.2016.01.022

Rossini, P. (2020). Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. *Communication Research*. Advance online publication. https://doi.org/10.1177/0093650220921314

Searles, K., Spencer, S., & Duru, A. (2020). Don't read the comments: The effects of abusive comments on perceptions of women authors' credibility. *Information, Communication & Society*, *23*(7), 947–962. https://doi.org/10.1080/1369118X.2018.1534985

Sobieraj, S. (2020). *Credible threat: Attacks against women online and the future of democracy.* Oxford University Press.

Sobieraj, S., Masullo, G. M., Cohen, P. N., Gillespie, T., & Jackson, S. J. (2020). Politicians, social media, and digital publics: Old rights, new terrain. *American Behavioral Scientist*, *64*(11), 1646–1669. https://doi.org/10.1177/0002764220945357

Southern, R., & Harmer, E. (2021). Twitter, incivility and "everyday" gendered othering: An analysis of tweets sent to UK members of parliament. *Social Science Computer Review, 39*(2), 259–275. https://doi.org/10.1177/0894439319865519

Stoll, A., Ziegele, M., & Quiring, O. (2020). Detecting impoliteness and incivility in online discussions: Classification approaches for German user comments. *Computational Communication Research, 2*(1), 109–134. https://doi.org/10.5117/CCR2020.1.005.KATH

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*(2), 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

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news." *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Weber, I., Laban, A., Masullo, G. M., Gonçalves, J., Torres da Silva, M., & Hofhuis, J. (2020). International perspectives on what's considered hateful or profane online. Center for Media Engagement. https://mediaengagement.org/research/perspectives-on-online-profanity

Ziegele, M., Naab, T. K., & Jost, P. (2020). Lonely together? Identifying the determinants of collective corrective action against uncivil comments. *New Media & Society*, *22*(5), 731–751. https://doi.org/10.1177/1461444819870130