

UC Davis

UC Davis Previously Published Works

Title

Uncivil Reactions to Sexual Assault Online: Linguistic Features of News Reports Predict Discourse Incivility

Permalink

<https://escholarship.org/uc/item/7rq63664>

Author

Stevens, Hannah R

Publication Date

2021-08-30

Peer reviewed

Uncivil Reactions to Sexual Assault Online: Linguistic Features of News Reports Predict
Discourse Incivility

Hannah R. Stevens, Irena Acic, & Laramie D. Taylor

University of California, Davis

Author Note:

This work has been peer reviewed and scheduled for publication in *Cyberpsychology, Behavior, and Social Networking*. Please Cite As:

Stevens, H. R., Acic, I., & Taylor, L. D. (In Press). Uncivil Reactions to Sexual Assault Online: Linguistic Features of News Reports Predict Discourse Incivility. *Cyberpsychology, Behavior, and Social Networking*.

Abstract

Reports of sexual assault have been found to elicit online discourse incivility. The present study employs a computerized coding tool to examine linguistic characteristics of news media that are likely to influence discourse incivility—specifically, negative emotion, disagreement, and discussion about power relations. Additionally, machine learning was harnessed to measure the levels of comment toxicity, insult, profanity, threat, and identity attack in *Reddit* and *Twitter* posts sharing news reports of sexual assault. Findings reveal that linguistic features of news articles interact with platform community norms to predict rape culture as expressed within online responses to reports of sexual assault.

Uncivil Reactions to Sexual Assault Online: Linguistic Features of News Reports Predict Discourse Incivility

Five years before the advent of the #MeToo movement, news reports of a rape case in Steubenville, Ohio spread across the United States. A 16-year-old victim had been repeatedly sexually assaulted by high-school football players who also made a video of the assault and discussed it on social media [1]. Some news stories included horrific details (e.g., a graphic video of the teenage boys laughing about how the victim was “so raped”) [2]. The public response was mixed, however. Frequently, rather than resolving to address systemic sexual violence, members of the public scrutinized specifics of the case, reacting with incivility toward the alleged victim and her assailants [3,4]. Despite the internet’s potential to facilitate the democratic ideal of constructive debate, rape culture, a phenomenon wherein society trivializes systemic rape, pervades online forums [5]. Online rape culture is revealed, in part, in online discussions of sexual assault in which users react with incivility—blaming victims and defending accused rapists [6,7]. Manifestations of rape culture as uncivil reactions to discourse around sexual assault raise questions about what features of news media elicit this behavior.

Discourse Incivility

Incivility presents itself in many forms; scholars characterize it as uncourteousness, irrational public expressions, refusal to acknowledge the legitimacy of opposing opinions, defensive reactions, disrespectful behaviors, disrespect for traditions of democracy, and violations of politeness norms [8-12]. Lerner and Miller’s just-world bias sheds light on incivility [13]. Just-world bias is the comforting illusion that good happens to good people, and that bad people deserve their adversity. Literature reveals a link between Just-World Bias and incivility.

For instance, research has shown that individuals react with empathetic anger, hostility, and aggression to unfair situations [14,15]. Consequently, when bad things happen to good people (e.g., the innocent become rape victims), individuals (1) experience cognitive dissonance and (2) try to relieve cognitive discomfort by rationalizing what happened (e.g., blaming victims) [6]. Just-world bias functions alongside stereotypes about which sexual assault and harassment victims are more worthy of empathy—namely, blameless victims violently assaulted by deviant perpetrators. [16-19].

All news reports of sexual assault likely elicit some cognitive dissonance, yet it is unclear why some such reports elicit discourse incivility. This is consequential, as incivility perpetuates itself. For example, a recent content analysis of comments responding to news reports of sexual assault found that hostile attributions of blame (i.e., victim-blaming comments) generated hostile replies [7]. In this work, we examined characteristics of news media that influenced discourse incivility.

Defense Motives and Power

Perceptions of news accounts of sexual violence are shaped by many considerations, including characteristics of those accounts. Media gatekeepers frequently report on sexual assaults through a limited lens that perpetuates stereotypes about sexual violence [1,20] and fails to address broader societal trends[21,22]. For example, a qualitative analysis of news coverage of the Steubenville rape case found prevalent rape myth themes [1]. Likewise, an analysis of news reports of a USA gymnastics sex abuse scandal showed that gatekeepers generally employed episodic frames, focusing on specific situations, instead of including the context of the prevalence of child sexual abuse by those in positions of power [21]. Ignoring patterns of

sexual violence may sustain and reinforce rape myths, including victim blaming [23,24]. Such media-perpetuated myths about the role of power dynamics, sexuality, and gender in sexual violence influence individuals' responses to sexual assault [25,26].

In concert with stereotypes and myths, audience motives can shape their reactions. When individuals are motivated to process a message, they give it more effortful attention [27]. One particularly salient motivation is the defense motive [28], which is the drive to defend core beliefs and pre-existing stereotypes [29]. Research has repeatedly shown that threatening one's belief in a just world elicits discourse incivility [30,31,6]. Scholars have investigated how defense motivations translate into attributions of blame, finding that observers attribute to minimize their own fear of being a victim or an accused perpetrator in a similar situation [6]. Salient power dynamics can prompt defense motives, motivating individuals to defend their ingroups and act with hostility toward outgroups [32]. Thus, we posit **H1: *Article discussions about group power dynamics will be associated with post incivility.***

Disagreement and Dissonance

A second source of incivility may be uncertainty around what happened in cases of sexual assault. Disagreement causes discomfort (i.e., cognitive dissonance). To ease dissonance-induced discomfort, individuals are motivated to resolve conflicting beliefs [33]. In online reports of sexual assault, individuals may resolve conflicting accounts by siding with in-group members and blaming outgroup members [6]. While disagreement may prompt defensive reactions, most journalistic codes require journalists to report news stories with objectivity [34], often interpreted as presenting all sides of a story. Yet, the conflict between entities, specifically the victim and perpetrator, can elicit uncertainty about what *really* happened. Indeed, extant

research shows a positive relationship between the disagreement in news reports and discourse incivility. For example, an analysis of Tweets discussing climate change found that highly disputed topics positively predicted discourse incivility [35]. Accordingly, we predict **H2**: *Article disagreement will increase post incivility.*

Negative Emotion

Negative information in reports of sexual assault may also provoke post incivility; such uncivil reactions may be a coping mechanism for managing distressing emotions. Underpinned by just-world bias [13,36], upsetting information in sexual assault news can motivate individuals to defend their belief in a just world [37-39]. Furthermore, increased negative information motivates individuals to exert more cognitive energy attending to a message [40], and individuals scrutinize negative news more than positive news [41]. Similarly, Trussler and Soroika investigated features of news articles that attract reader attention; they invited participants to the lab to read a news story for a bogus eye-tracking task [42]. When presented with 50 news stories, participants were more likely to select—and carefully read—negative news articles than positive and neutral ones. We predict such careful attention to upsetting information will trigger defense motives and incivility. We posit **H3**: *Article negative emotion will be associated with post incivility.*

Platform Community Norms

Organizational behavior scholars have long theorized that incivility stems from interactions between situational and community norms, shaped by organizational policies and measures. For example, low expectations of procedural justice might interact with employee anger to elicit uncivil behavior [43,44]. Similarly, platform capabilities may shape community

norms—both online and offline—that enable or impede incivility (e.g., Reddit moderator rules and Twitter’s “hateful conduct policy” [45,46]). For example, the anonymity of many online communication spaces (e.g., Reddit) allows users to act without damaging their reputation, particularly when making uncivil comments [47,48]. When anonymous, users may post hostile responses to news reports of sexual assault, (e.g., the shaming and blaming of victims) [7]. Research has shown that removing anonymity from an online forum can decrease discourse incivility [49]. Yet while anonymity may facilitate incivility in certain situations, identifiability may perpetuate incivility in others. For example, scholars have argued that identifiability to an in-group audience can prompt expressions of commitment to one’s in-group [50].

Community norms are contextually-dependent, and they vary—both within online culture broadly and within the micro-cultures of the online platforms. Accordingly, distinct online communities respond to and censure different dimensions of incivility differently [51,45]. Thus, defensive comment incivility may vary by platform. Our related question here is **RQ1: *How does incivility vary by online platform?***

Method

Data Collection

An anonymous social platform (*Reddit*) and an identifiable social platform (*Twitter*) were included in the sample because of their popularity as venues for sharing news [52,53]. An investigation of the most popular news *subreddits* (forums dedicated to sharing news) found that the most popular US news subreddit was ‘*/r/news*.’ We felt it was important to sample from a *subreddit* with all news, as opposed to subreddits specific to sexual assault, to mitigate sampling bias.

We began with a search of each platform using the terms “sexually assaulted,” “sexually assaulting,” and “sexual assault.” To extract posts sharing news reports of sexual assault, we took a systematic random sample of 50 posts shared to each platform every week for five years (from January 2015 to December 2019, totaling 24,000 posts). The *GetOldTweets3* and *Selenium Python3* libraries were used to scrape all Tweets and Reddit posts [54, 55].

To facilitate the representativeness of the news articles, researchers then manually filtered through a shuffled list of 50 English posts from each week per platform to randomly extract 10 posts from each week (five from *Reddit* and five from *Twitter*) that shared a U.S. news report of a sexual assault. We only included reports of sexual assault, and excluded all other news frames (e.g., stories about generalized rape culture). When a news article had been deleted, we sampled a different post for that week. Researchers manually extracted the text from the news articles; a similar approach has been used in media effects research to investigate the relationship between news sentiment and tweet sentiment [56]. The resulting sample contained 480 unique news articles and the post language associated with each.

Sentiment Analysis Procedures

Individuals express incivility in different ways [57], requiring an operationalization that accounts for its various forms. We harnessed machine learning to provide a continuous measure of five prevalent kinds of comment incivility. Specifically, we employed Perspective API, an AI tool created by Google’s Counter-Abuse Technology team, to measure levels of comment toxicity, insult, profanity, threat, and identity attack [58] (see Table 1). Google Perspective API generates a comment score between 0 and 1 for each dimension of incivility. Higher scores show an increased likelihood of a reader perceiving the comment as that dimension of incivility.

Notably, a single comment receives different scores for each dimension. For example, the

comment “If I were her father, I’d have killed her” may receive a score of 0.75 for toxicity and .95 for threat, showing a 75% likelihood that a reader will perceive the comment as toxic and a 95% likelihood that a reader will perceive the same comment as threatening.

After measuring different incivility dimensions, we analyzed the news articles’ features using the Linguistic Inquiry and Word Count computerized coding tool [59]. We extracted three message characteristics hypothesized to predict incivility:

1. We analyzed article disagreement language to measure differences between different sides of a story; these words are used to differentiate between ideas (e.g., “but,” “hasn’t”).
2. We measured negative emotion language to assess levels of anger, anxiety, and sadness in articles (e.g., “afraid,” “depressed”).
3. We measured article words referencing power relations between groups (e.g., “famous,” “bully”).

LIWC is observational in nature, and accounts for long vs. short texts by calculating the percentage of lexicon category words relative to all words in a text. For example, we might discover that 15/745 (2.04%) of words in an article were LIWC *negative emotion* lexicon words. The LIWC output would then assign that particular article a *negative emotion* value of 2.04.

Results

Repeated Measures ANOVA

We first gauged whether to model the dimensions of incivility together or separately using a repeated measures analysis of variance (ANOVA) to determine whether significant differences existed among toxicity, identity attack, insult, profanity, and threat. We ran

Mauchly's test to check the assumption of sphericity [60]. Based on an alpha of 0.05, variance of scores between repeated measurements differed significantly from one another, $p < .001$, suggesting the assumption of sphericity was violated. Accordingly, results were calculated using the Greenhouse-Geisser correction to adjust for sphericity violation [61]. The main effect for the within-subjects factor was significant, $F(4, 1916) = 131.02, p < .001$, suggesting significant differences among toxicity, identity attack, insult, profanity, and threat (see Tables 2-3).

Tukey comparisons were used to test the marginal mean differences in each combination of within-subject effects (i.e., incivility dimension) based on an alpha of 0.05. There were significant differences between each combination of within-subjects dimensions of incivility, except for insult and profanity (see Table 4).

Linear Regression Analyses

We conducted five linear regression analyses to assess whether linguistic indicators of power differentials, negative emotions, and disagreement in news articles predicted each of the five dimensions of incivility expressed in posts associated with those news articles. We also assessed whether our predictor variables interacted with platform (Reddit vs. Twitter).

Discussion of Group Power Dynamics

There were no main effects of power dynamics on dimensions of incivility. However, the power dynamics in news articles in interaction with platform did significantly predict several modes of incivility, namely profanity, insult, threat, and toxicity. In each case, this reflected a stronger relationship between discussion about group power dynamics and post incivility on Reddit than on Twitter (see Table 5).

Linguistic Disagreement

Our second hypothesis predicted article disagreement, operationalized as linguistic differentiation, would increase post incivility. There were no main effects of disagreement on dimensions of incivility. However, the interaction between disagreement and platform had a significant effect on profanity, $p = .020$, and toxicity, $p = .026$, reflecting a stronger relationship between disagreement language and post incivility on Twitter compared to Reddit (see Table 5).

Negative Emotion

Hypothesis three predicted article negative emotion will be associated with post incivility. Negative emotion significantly predicted identity attack, $p = .035$; contrary to our prediction, results revealed increased negative emotion decreased identity attack. While there were no other main effects of negative emotion, it interacted with platform to predict profanity, $p = .024$, reflecting a stronger relationship between negative emotion and profanity on Twitter than on Reddit (see Table 5).

Platform

Platform was found to significantly predict two dimensions of incivility—threat and identity attack. Moving from Twitter to Reddit significantly decreased the mean value of threat, $p < .001$ and identity attack, $p = .004$, showing a greater prevalence of both on Twitter than on Reddit (see Table 5).

Discussion

The present study sheds light on rape culture by examining the antecedents of incivility associated with news reports of sexual assault in online forums. Offering partial support for our predictions, platform exerted a main effect on online threats and identity attacks. Extant

organizational behavior literature provides a useful framework for examining how social network community norms are shaped by platform attributes and capabilities; for example, Reddit's automoderator bot allows moderators to customize rules for discourse [45]. Different automated rules may facilitate community norms discouraging incivility. Contrary to our prediction that article negative emotion would elicit identity attacks, negative emotion decreased identity attacks. Overall, there was no simple, unitary explanation for incivility, and incivility dimensions should be unraveled and studied separately.

While there were no main effects of article group power dynamics and linguistic disagreement on any mode of incivility, these linguistic features of news reports interacted with platform to predict varying modes of incivility. News discussion of group power dynamics interacted with platform, producing greater profanity, insult, threat, and toxicity on Reddit than on Twitter. Conversely, linguistic disagreement in articles shared to Twitter exerted a stronger effect on profanity and toxicity. The relationship between article negative emotion and profanity was also greater on Twitter than on Reddit. In sum, incivility manifested differently on Reddit compared to Twitter, necessitating nuanced measures of incivility on different platforms. Contextualizing these findings with investigations of workplace incivility shed light on the role of online community norms in manifestations of discourse incivility.

Although there are many causes of rape culture, findings reveal that defense motives to sustain preexisting stereotypes may be one driving force behind rape culture as expressed within online responses to reports of sexual assault. Instead of viewing rape culture as a widespread, societal problem, individuals may react defensively with discourse incivility to the cognitive discomfort elicited by news article discussion of group power dynamics, negative emotion, and

linguistic disagreement—thus, reinforcing existing stereotypes and myths about sexual assault [1,16,17]. Furthermore, distinct online community norms, shaped by a platform's affordances, can enable or impede such incivility.

These findings are consistent with organizational behavior theory suggesting that incivility stems from interactions between situational and community norms, which are shaped by organizational policies and measures [43,44]. Distinct community norms, driven by platform capabilities, either enable or impede varying dimensions of incivility. For example, low expectations of being banned from Twitter for profanity might enable a profane response to a report of sexual assault containing heightened linguistic disagreement. Conversely, a user might filter their profanity on Reddit if the subreddit rules stipulate they will be banned as a consequence.

While findings illuminate the relationship between linguistic style and online discourse incivility on two platforms, this study is not without limitations. The LIWC computerized coding tool does not allow the nuanced coding that human coders can generate. We endeavored to reduce this potential limitation by using a well validated sentiment analysis procedure [59]. Future work might employ human coders to analyze specific attributes that make victims more worthy of empathy [16,17].

Also, widespread concern about the role of anonymity in discourse incivility has prompted some media outlets to prohibit anonymous comments [62,63]. Our results revealed that incivility is multi-dimensional, and anonymity may not be the underlying force driving discourse incivility. We focused on news articles shared to only two online platforms, namely *Reddit* and *Twitter*, and therefore cannot make general statements about expressions of rape culture in other

discussion forums. Likewise, this work cannot make conclusions about expressions of rape culture in response to self-disclosures of sexual assault.

Community norms, both within online culture broadly and within the various micro-cultures of the online platforms, likely respond to and censure different dimensions of incivility differently [45,51]. Given the significant role of platform community norms in predicting incivility modes, future work should investigate how specific platform affordances predict manifestations of incivility.

Conclusion

In summary, linguistic features of news provoke different modes of incivility, which manifest uniquely by platform. Future investigations of the effects of linguistic features of news reports of sexual assault require nuanced measures of incivility on different platforms. Our work raises ethical implications for the fields of journalism, suggesting specific linguistic features of news—specifically, linguistic disagreement and discussion about group power dynamics—prompt incivility.

References

1. Kosloski AE, Diamond-Welch BK, Mann O. The presence of rape myths in the virtual world: A qualitative textual analysis of the Steubenville sexual assault case. *Violence and gender* 2018.
2. Horton A. (2019) Roll Red Roll: behind the chilling documentary on high school rape. *The guardian*.
<https://www.google.com/url?q=https://www.theguardian.com/film/2019/apr/09/roll-red-roll-behind-the-chilling-documentary-on-high-school&sa=D&source=editors&ust=1622426264132000&usg=AOvVaw1pZQZisaKQ9-Q3CO2ZWjsK>. Published April 9, 2019. Accessed May 31, 2021.
3. Levy A. (2013) Trial by Twitter. *New Yorker* (New York, N.Y.: 1925).
<https://www.newyorker.com/magazine/2013/08/05/trial-by-twitter>. Published July 29, 2013. Accessed May 31, 2021.
4. Penny L. (2019) Laurie Penny on Steubenville: this is rape culture’s Abu Ghraib moment.<https://www.newstatesman.com/laurie-penny/2013/03/steubenville-rape-cultures-abu-ghraib-moment> (accessed May31 2021).
5. Henry N, Powell A. Embodied harms: gender, shame, and technology-facilitated sexual violence: Gender, shame, and technology-facilitated sexual violence. *Violence against women* 2015; 21:758–779.
6. Grubb A, Harrower J. Attribution of blame in cases of rape: An analysis of participant gender, type of rape and perceived similarity to the victim. *Aggression and violent behavior* 2008; 13:396–405.
7. Zaleski KL, Gundersen KK, Baes J, et al. Exploring rape culture in social media forums. *Computers in human behavior* 2016; 63:922–927.
8. Barber BR, Mattson K, Peterson J. (1997) The state of electronically enhanced democracy: a survey of the internet. New Brunswick, NJ: Walt Whitman Center for Culture and Politics of Democracy.
9. Fritz JH. (2009) Rudeness and incivility in the workplace. In: *Friends and Enemies in Organizations*. London: Palgrave Macmillan UK, pp. 168–194.
10. Muddiman A, Stroud NJ. News values, cognitive biases, and partisan incivility in comment sections: Uncivil comments. *The Journal of communication* 2017; 67:586–609.
11. Mutz DC, Reeves B. The new videomalaise: Effects of televised incivility on political trust. *The American political science review* 2005; 99:1–15.
12. Papacharissi Z. Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society* 2004; 6:259–283.
13. Lerner MJ, Miller DT. Just world research and the attribution process: Looking back and ahead. *Psychological bulletin* 1978; 85:1030–1051.
14. Batson CD, Kennedy CL, Nord L-A, et al. Anger at unfairness: is it moral outrage? *European journal of social psychology* 2007; 37:1272–1285.
15. Chory-Assad RM, Paulsel ML. Classroom justice: student aggression and resistance as reactions to perceived unfairness. *Communication education* 2004; 53:253–273.
16. Eelmaa S, Murumaa-Mengel M. (2021) Who is worthy of help? Constructing the stereotype of the “ideal victim” of child sexual abuse. *SAGE Advance* 2021.

17. Christie N. (1986) *The Ideal Victim*. In: *From Crime Policy to Victim Policy*. London: Palgrave Macmillan UK, pp. 17–30.
18. Siefkes-Andrew AJ, Alexopoulos C. Framing blame in sexual assault: An analysis of attribution in news stories about sexual assault on college campuses. *Violence against women* 2019; 25:743–762.
19. Taylor LD, Alexopoulos C, Ghaznavi J. Touchy subjects: Sex in the workplace on broadcast, cable, and internet television. *Sex roles* 2016; 75:476–489.
20. Hindes S, Fileborn B. Girl power gone wrong?: #MeToo, Aziz Ansari, and media reporting of (grey area) sexual violence. *Feminist media studies* 2020; 20:639–656.
21. Smith LR, Pegoraro A. Media framing of Larry Nassar and the USA gymnastics child sex abuse scandal. *Journal of child sexual abuse* 2020; 29:373–392.
22. Yu W. News portrayals of child sexual abuse in China: Changes from 2010 to 2019. *Journal of child sexual abuse* 2021; 1–22.
23. Adolfsson K. (2018) *Blaming Victims of Rape: Studies on Rape Myths and Beliefs about Rape*. 2018.
24. Manoussaki K, Veitch F. Ambivalent sexism, right wing authoritarianism and rape myth acceptance in Scotland. *International Journal of Gender and Womens' Studies* 2015; 3.
25. Anderson I, Doherty K. (2007) *Accounting for Rape. Psychology, Feminism and Discourse Analysis in the Study of Sexual Violence*. London & New York: Routledge.
26. Sun C, Bridges A, Johnson JA, et al. Pornography and the male sexual script: An analysis of consumption and sexual relations. *Archives of sexual behavior* 2016; 45:983–994.
27. Chaiken S, Trope Y. (1999) *Dual-process theories in social psychology*. New York: Guilford Press.
28. Todorov A, Chaiken S, Henderson MD. (2012) *The heuristic-systematic model of social information processing*. In: *The Persuasion Handbook: Developments in Theory and Practice*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc., pp. 195–212.
29. Chen S, Duckworth K, Chaiken S. Motivated heuristic and systematic processing. *Psychological inquiry* 1999; 10:44–49.
30. Burger JM. Motivational biases in the attribution of responsibility for an accident: A meta-analysis of the defensive-attribution hypothesis. *Psychological bulletin* 1981; 90:496–512.
31. Gyekye SA, Salminen S. The self-defensive attribution hypothesis in the work environment: Co-workers' perspectives. *Safety science* 2006; 44:157–168.
32. Korostelina K. Intergroup identity insults: A social identity theory perspective. *Identity (Mahwah, N.J.)* 2014; 14:214–229.
33. Festinger L. *A Theory of Cognitive Dissonance*. *The American journal of psychology* 1957; 2.
34. Ward SJA. (2020) *Truth and Objectivity*. In: *The Routledge Handbook of Mass Media Ethics*. Routledge, pp. 101–114.

35. Anderson AA, Huntington HE. Social media, science, and attack discourse: How Twitter discussions of climate change use sarcasm and incivility. *Science communication* 2017; 39:598–620.
36. Lerner MJ, Simmons CH. Observer's reaction to the "innocent victim": Compassion or rejection? *Journal of Personality and social Psychology* 1966; 4.
37. Landström S, Strömwall LA, Alfredsson H. Blame attributions in sexual crimes: Effects of belief in a just world and victim behavior. *Nordic psychology* 2016; 68:2–11.
38. Luginbuhl J, Mullin C. Rape and responsibility: How and how much is the victim blamed? *Sex roles* 1981; 7:547–559.
39. Strömwall LA, Alfredsson H, Landström S. Blame attributions and rape: Effects of belief in a just world and relationship level: Blame attributions and relationship level. *Legal and criminological psychology* 2013; 18:254–261.
40. Vaish A, Grossmann T, Woodward A. Not all emotions are created equal: the negativity bias in social-emotional development. *Psychological bulletin* 2008; 134:383–403.
41. Meffert MF, Chung S, Joiner AJ, et al. The effects of negativity and motivated information processing during a political campaign. *The Journal of communication* 2006; 56:27–51.
42. Trussler M, Soroka S. Consumer demand for cynical and negative news frames. *The International Journal of Press/Politics* 2014; 19:360–379.
43. Schilpzand P, De Pater IE, Erez A. Workplace incivility: A review of the literature and agenda for future research: WORKPLACE INCIVILITY. *Journal of organizational behavior* 2016; 37:S57–S88.
44. Walsh BM, Magley VJ, Reeves DW, et al. Assessing workgroup norms for civility: The development of the civility norms questionnaire-brief. *Journal of business and psychology* 2012; 27:407–420.
45. Jhaver S, Birman I, Gilbert E, et al. Human-machine collaboration for content regulation: The case of reddit automoderator. *ACM transactions on computer-human interaction: a publication of the Association for Computing Machinery* 2019; 26:1–35.
46. Twitter Help Center. (2021) Twitter's policy on hateful conduct. <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (accessed May31 2021).
47. Borah P. Does it matter where you read the news story? Interaction of incivility and news frames in the political blogosphere. *Communication research* 2014; 41:809–827.
48. Hutchens MJ, Cicchirillo VJ, Hmielowski JD. How could you think that?!?!: Understanding intentions to engage in political flaming. *New media & society* 2015; 17:1201–1219.
49. Santana AD. Virtuous or Vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism practice* 2014; 8:18–33.
50. Klein O, Spears R, Reicher S. Social identity performance: extending the strategic side of SIDE. *Personality and social psychology review: an official journal of the Society for Personality and Social Psychology, Inc* 2007; 11:28–45.

51. Cho V, Jimerson JB. Managing digital identity on Twitter: The case of school administrators. *Educational management administration & leadership* 2017;
52. Weninger T, Zhu XA, Han J. (2013) An exploration of discussion threads in social news sites: A case study of the Reddit community. 2013.
53. Kwak H, Lee C, Park H, et al. (2010) What is Twitter, a social network or a news media? In: *Proceedings of the 19th international conference on World wide web - WWW '10*. New York, New York, USA: ACM Press.
54. Mottl D. (2019) GetOldTweets3 0.0.11.<https://pypi.org/project/GetOldTweets3/> (accessed May31 2021).
55. Huggins J. The Selenium Browser Automation Project.<https://www.selenium.dev/documentation/en/> (accessed May31 2021).
56. Stevens HR, Oh YJ, Taylor LR. (2021) Desensitization to Fear-Inducing COVID-19 Health News on Twitter: Observational Study. *JMIR Infodemiology*.
57. Anderson AA, Yeo SK, Brossard D, et al. Toxic talk: How online incivility can undermine perceptions of media. *International journal of public opinion research* 2018; 30:156–168.
58. Anon. Google Perspective API.<https://support.perspectiveapi.com/s/about-the-api-attributes-and-languages> (accessed May31 2021).
59. Pennebaker JW, Boyd RL, Jordan K. (2015) The development and psychometric properties of LIWC2015. 2015.
60. Mauchly JW. Significance test for sphericity of a normal n-variate distribution. *The annals of mathematical statistics* 1940:204–209.
61. Greenhouse SW, Geisser S. On methods in the analysis of profile data. *Psychometrika* 1959; 24:95–112.
62. Crosbie V. Time to get tough: Managing anonymous reader comments. In. *The Online Journalism Review* 1999.
63. Rieder R. No comment: It's time for news sites to stop allowing anonymous online comments. *American Journalism Review* 2010; 32:2–3.

Table 1

Dimension	Definition (Google Perspective API, 2020)	Sample Comments
Toxicity	“A rude, disrespectful, or unreasonable comment that is likely to make people leave a discussion.”	Two drunk girls regret fucking frat guys while they were boozed up at a party.
Insults	“Insulting, inflammatory, or negative comment towards a person or a group of people.”	What a piece of shit...This is why I hate men.
Profanity	“Swear words, curse words, or other obscene or profane language.”	“Fucking Whores.”
Threat	“Describes an intention to inflict pain, injury, or violence against an individual or group.”	“If I were her father, I’d have killed her.”
Identity Attack	“Negative or hateful comments targeting someone because of their identity.”	“What the fuck? Something wrong here... (checks article) Ohhhh... he is black. Now I get it.”

Note. Google Perspective API offers a machine learning method of assessing varying dimensions of incivility.

Table 2*Repeated Measures ANOVA Results*

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	η_p^2
Within-Subjects						
Within Factor	4	10.05	2.51	131.02	< .001	0.21
Residuals	1916	36.73	0.02			

Table 3*Means Table for Within-Post Variables*

Variable	<i>M</i>	<i>SD</i>
Toxicity	0.47	0.26
Identity Attack	0.28	0.19
Insult	0.33	0.23
Profanity	0.32	0.27
Threat	0.39	0.25

Note. $n = 480$.

Table 4

The Marginal Means Contrasts for each Combination of Within-Subject Variables for the Repeated Measures ANOVA

Contrast	Difference	SE	<i>t</i>	<i>p</i>
Toxicity - Identity Attack	0.18	0.01	23.10	< .001
Toxicity - Insult	0.14	0.01	26.89	< .001
Toxicity - Profanity	0.15	0.01	29.39	< .001
Toxicity - Threat	0.08	0.01	7.91	< .001
Identity Attack - Insult	-0.05	0.01	-6.06	< .001
Identity Attack - Profanity	-0.03	0.01	-3.49	.005
Identity Attack - Threat	-0.10	0.01	-10.56	< .001
Insult - Profanity	0.01	0.01	1.79	.382
Insult - Threat	-0.06	0.01	-5.08	< .001
Profanity - Threat	-0.07	0.01	-5.61	< .001

Notes: df = 479