



City Research Online

City, University of London Institutional Repository

Citation: Camaj, L., Hellmueller, L., Vallejo Vera, S. & Lindner, P. (2024). The Democratic Value of Strategic Game Reporting and Uncivil Talk: A Computational Analysis of Facebook Conversations During U.S. Primary Debates. *Journalism & Mass Communication Quarterly*, 101(2), pp. 428-450. doi: 10.1177/10776990231226403

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/32448/>

Link to published version: <https://doi.org/10.1177/10776990231226403>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**The Democratic Value of Strategic Game Reporting and Uncivil Talk: A Computational
Analysis of Facebook Conversations During U.S. Primary Debates**

Abstract

This study explores discourse features on Facebook pages of news organizations during the 2020 U.S. primary debates using a state-of-the-art machine-learning model. Informing the scholarly debate about the implications of strategic game reporting in online spaces, we find that it is not necessarily linked to uncivil discourse, yet it might deter from relevant conversations. Second, addressing fears about the undesired outcomes of uncivil talk, our data suggest that incivility can coexist with rational discourse in user comments, although this relationship is not pervasive. Implications of these results are discussed in the context of the role of hybrid media for political engagement during electoral campaigns.

Keywords: political debates, incivility, online discourse, strategic reporting, horserace, computational methods

Political debates epitomize highlights in lengthy election campaigns in American politics, especially during the primary season. During these mediated events, candidates exchange their views for policy proposals while offering a chance for voters to engage in reasoning that informs their vote. Modern political debates occur in hybrid media environments where older and newer media operate simultaneously (Chadwick et al., 2017), creating spaces for voters to participate in discussions and deliberations. This study focuses on the discourse features of user comments on news media's Facebook pages during the 2020 primary debates in United States. Empirical research has come to question the role of online spaces to vitalize citizen communication practices conduit to democracy, yet moderated digital news platforms can offer spaces where quality conversations can flourish under the right conditions (Rowe, 2015). Hence, this study is nested within two layers of mediated political debates: first it examines news coverage of the televised debates on social media and links it to the user comments responding to this news coverage posted asynchronously.

This study contributes to the emerging literature that analyzes viewer discourse on social media during televised debates (Camaj, 2021; Robertson et al., 2019; Ventura et al., 2021) in two crucial ways. First, we explore (1) news coverage with a special focus on strategic game reporting and negativity (Esser & Strömbäck, 2012), and 2) conversations primed by news coverage and other commenters. Pre- and post-debate news coverage and commentary serves as a mechanism that stimulates subsequent news consumption and political conversations (Cho & Choy, 2011), and can impact the quality of discussions (Camaj, 2021). In addition, debate viewers who post social media comments can affect each other (Ventura et al, 2021) and are more likely to be opinion leaders affecting wider public opinion on social media (Tremayne & Minooie, 2015). Exploring these two multilayered factors (the news post and user comment

level) that determine the quality of debate related discussions in online spaces will add an important **first contribution of our study toward theorizing televised debate effects that extend beyond direct exposure.**

The second contribution relates to the context of debates. Primary debate audiences in the United States are highly interested and engaged voters who are most likely to determine the outcome of these elections, and potentially change the course of political discourse at large. Although news commentators on social media are highly partisan (Kim et al., 2020), primary debate viewers are less likely to be polarized than debate viewers during general elections, as they share the same political ideology. Yet, this does not mean that their online discussions are more civil as uncivil talk among homogenous groups can serve as a mobilization tool (Berry & Sobieraj, 2016). Merging content data with user behavioral data on social media across 11 primary political debates, we address the scholarly debate about the implications of strategic game reporting for audience engagement in online spaces (Trussler & Soroka, 2014; Zoizner, 2021) and discuss fears about the undesired democratic outcomes of uncivil talk (Van't Riet & Van Stekelenburg, 2021).

The Democratic Function of Online Discussions During Political Debates

The integration of social media to follow live political debates generates a parallel stream of conversations about candidate performances enabling viewers to tap into overarching public opinion (McKinney et al., 2013). While online conversations can serve multiple behavioral roles for debate viewers (Bucy et al., 2020; Freelon & Karpf, 2015; Robertson et al., 2019), very few studies have explored the discursive quality of debate related user comments. Many viewers tune in to the debates because they are politically motivated partisans (McKinney & Warner, 2013), while social media engagement with debate related content represents a cognitive and expressive

process for strong partisans (Jennings et al., 2020). This study assumes that if televised debates stimulate news consumption and political conversation (Cho & Choy, 2011) and cognitive elaborations (McGregor & Mourão, 2017), there might be democratic value in such discussions. Relevant to this paper, Facebook pages of news media might be well suited to provide a platform for quality discussions (Ziegele et al., 2020), given that debate coverage on Facebook can lead to user conversations that exhibit elements of rational discourse (Camaj, 2021). However, this line of research has been criticized for its predominant concentration on the negative aspects of online discourses, such as incivility (Gonçalves et al., 2020) or toxicity (Ventura et al., 2021), often at the expense of the constructiveness and other positive aspects of the online comments (Reimer et al., 2021).

Previous research in this domain has been primarily guided by deliberative theory which favors communicative behaviors in which citizens engage with reasoning to build consensus about issues of common interest politely and respectfully (Fishkin, 2009). At the minimum, deliberate discussions need to be informed by reasoning and justification; be constructive; reciprocal; and civil (Steenbergen et al., 2003). The deliberative approach considers incivility to be incompatible with quality discourse as it diminishes discourse constructiveness and user engagement. This literature encompasses a variety of definitions and operationalizations of incivility depending on which norm violations it embodies, such as the violation of respect norm (Coe et al., 2014), politeness norm (Mutz, 2015), the collective democratic norms (Papacharissi, 2004), or the violation of multiple norms (Bormann et al., 2022). Some forms of incivility are more benign such as name calling or vulgarities that violate the politeness and individual respect norm (Coe, Kenski, and Rains (2014), while other forms of incivility are more harmful to the

democratic process, such as comments that include mockery and belittlement of social groups or disrespect for collective traditions of democracy (Papacharissi, 2004).

Incivility is a nuanced concept often bound to context and discourse culture. Given the context of our study in the primary elections in the U.S., that focuses on an ideologically homogenous group discussing electoral choices within mainstream news organizations online platforms, we develop our own measure of incivility guided by previous literature that defines incivility as “features of discussion that convey an unnecessarily disrespectful tone towards the discussion forum, its participants, or its topics” (Coe et al., 2014, p.3) and violate social norms governing personal interactions (Mutz, 2015).

But, rather than relying exclusively on deliberate theories, we acknowledge that conceptual models of the public sphere have moved toward multiplicity spurred by recognition of social complexity and sociocultural diversity (Fraser, 1992). A more radical approach acknowledges antagonistic dimension of political participation and recognizes the conflictual aspect of democracy as the fundamental source of its virtues (Mouffe, 1999). Passionate public discussions that often take the form of incivility are recognized as manifestations of dissent that draw attention to social injustice (Edyvane, 2020). As such, these theorists make a substantive normative case for the democratic value of some forms of incivility, while empirical research is inconclusive regarding the harmful effects of online uncivil talk in the realm of politics (Van’t Riet & Van Stekelenburg, 2021), and emphasizes the potential of incivility to mobilize audiences (Berry & Sobieraj, 2016).

In this paper we argue that, while incivility is not an integral dimension of deliberation and acknowledge that some forms of toxicity are harmful to political discourse, it does not inherently deter from all dimensions of quality discussion. Without judging the comparative

benefits of deliberative and agonistic discourse, we argue that digital spaces offer platforms where many forms of communication that can coexist. Social media platforms, which merge the private and the public domains, have contributed to the “privatization of the public sphere” (Klinger & Svensson, 2015) hence the expressive language in these online spaces encapsulates strong focus on conflict, collective identities and passionate discourse which often deviates from the normative expectations of consensus (Dahlberg, 2007). For example, a recent study found that in the context of live-tweeting a televised candidate debate, strong partisans were more likely to generate more comments supporting their candidate and criticizing and attacking the outgroup candidate (Jennings et al., 2020). This suggests that the online discussions can have positive and negative features simultaneously.

Similarly, in our context of primary debates we might find comments that can express a reasoned argument that is supported by empirically verified evidence, personal and anecdotal observations or shared norms and values (Dahlberg, 2011). At the same time or in the same conversation thread commentors might adopt a disrespectful tone or violate some aspects of social norms while still offering rational argumentation that identify common community goals or propose solutions to community problems (Friess & Eilders, 2015). In other words, we assume here that uncivil language in online comment section does not necessarily depress users’ motivations to engage and express arguments.

Yet, Facebook and other digital news platforms have started integrating incivility detection tools that filter and delete uncivil content indiscriminately. On the other side, research investigating the multi-dimensional nature of online discussions is surprisingly rare, and the conditions under which such discourse appears are not well understood. Guided by previous literature (Jaidka et al., 2019; Steenbergen et al., 2003), this study embraced a multi-dimensional

perspective to understand the co-existing features of online discussions prompted by news coverage of televised debates.

Linking Debate Reporting with Discourse Features in News Comments

While we might be less likely to find quality discussions in real-time and synchronous online spaces where people engage to socialize (Ventura et al., 2021), previous research suggests that moderated and primed digital news spaces can offer platforms where quality discussions can flourish (Rowe, 2015) as relevant forms of political participation that lead to informed decision-making during elections. Guided by priming theory, previous research has explored the influence of pre- and post-debate news coverage and commentary on debate viewers' knowledge and opinions about political candidates (Gross et al. 2019). For example, in an experimental setting, Jennings and colleagues found that issue priming during debate viewing produced an elaborative effect on audiences closing the gap between those people with higher and lower levels of political knowledgeable (Jennings et al., 2022). Similarly, recent research suggests that the quality of discussions in online news comments depends on the structure of the information that precedes and debate news coverage can prime online conversations (Camaj, 2021; Ziegele et al., 2020). Guided by the decades long research in political journalism (Esser & Strömbäck, 2012), in this study we adopt three predominant dimensions of election reporting that might prime conversations in the comment section: (1) depoliticization, (2) personalization and (3) negativity.

Media depoliticization is conceptualized as news coverage of elections that marginalizes “the core of politics – the substance, issues, ideologies, and linkages between real-world problems and proposed solutions” (Esser & Strömbäck, 2012, p. 318). Instead, election coverage predominantly concentrates on the process of political campaigns (Cappella & Jamieson, 1997). The coverage of the campaign process can focus on the *strategy* – “interpretations of candidates’

or parties' motives for actions and positions; their strategies and tactics for achieving political or policy goals"; and *game* (a.k.a horserace) coverage, which center on "who is winning or losing elections, in the battle for public opinion" (Aalberg et al., 2012, p.172).

We are also interested to examine the consequences of *personalization* of election coverage for the quality of user discourse. The rise of 'candidate-centered politics' (Van Aelst et al., 2012) is especially dominant on social media, where candidates adopt a more personalized communication style when discussing their professional, emotional, and private lives. Consequently, news coverage prioritizes individual politicians, and in the context of political debates can include news focus on candidate personality traits and personal lives at the expense of parties, ideologies, or policies (Benoit, 2013).

While many scholars have problematized the democratic consequences of strategic game election reporting (Zoizner, 2021), a few argue that such coverage has the potential to engage audiences with election news (Trussler & Soroka, 2014). Recent studies did not find any direct relationship between strategic game coverage and higher levels of incivility in user comments (Camaj, 2021; Goncalves et al., 2020), while emphasis on candidate character in news coverage of the debates might lead to higher number of comments that express relevant opinions (Camaj, 2021). Based on this literature, we propose the following research question:

RQ1: What is the relationship between news frames (strategic game reporting) and discourse features in news user comments on Facebook during televised political debates?

Additionally, news reporting of politics in general and election news more specifically have a bias towards *negativity*, confrontation and conflict (Esser & Strömbäck, 2012). During televised debates, candidates can attack their opponent's policy positions or policy records; they

might emphasize their opponent's character flaws and lack of leadership and attack their opponents negative campaigning and dirty tricks (Benoit, 2013). In turn, such nuanced negative frames are likely to be picked up by media coverage.

There is widespread concern that political attacks during campaigns may have problematic effects, although empirical evidence suggests that it can increase voter engagement with election news (Trussler & Soroka, 2014). Yet, this research also indicates that candidate attacks might contribute to uncivil talk. A recent study found that comments in response to debate news focusing on candidate attacks contained twice the rate of impolite comments than those posted under news on acclaims (Camaj, 2021). In a similar context, Rossini and colleagues (2021) found that negative messages posted by political candidates on their Facebook walls were more likely to receive uncivil comments than advocacy posts, confirming previous findings that the public might be prone to adopt the uncivil rhetoric expressed by politicians. Adding to this line of research, we explore two features of negative reporting of election campaigns, general negativity in news posts and news post that specifically concentrate on candidate attacks. We are guided by the following hypothesis and research question:

Hypothesis 1: Debate coverage emphasizing candidate attacks will have a positive association with uncivil discourse in comments posted on media's Facebook pages.

RQ2: What is the relationship between negative reporting and discourse features in news user comments on Facebook during televised political debates?

Linking Incivility with Deliberate Discourse Features in User Comments

In addition to the elite influences, recent scholarship argues that debate audiences' online expressions are likely to also exert significant influence on debate viewers (Camaj & Northup, 2019), suggesting that the rise of multiscreening and online streaming chats enable audiences to

prime their peers who engage in online conversations or those who consume it while watching televised debates. Relevant to this study, a recent study found that debate audiences who were exposed to a real-time streaming chat containing high levels of toxicity during debate watching reported lower affect toward Democrats and a worse viewing experience (Asbury-Kimmel et al., 2021). The question remains, to what degree do uncivil user comments impact the general quality of user discussions in asynchronous conversations in digital news spaces.

To address this question, we are informed by recent scholarship that has explored the idea that incivility might coexist with other characteristics of quality discussions in online spaces where multiple publics have the chance to engage in conversations about public issues (Chen, 2017; Jaidka et al., 2019; Rossini, 2020). A recent study found a positive association between incivility and justified opinion expression in online news sites (Rossini, 2020), reinforcing the idea that some type of uncivil talk has become normalized in the context of informal online discussions and is more acceptable when directed toward political opinions (Muddiman, 2017).

We adopt a multi-level approach to identifying characteristics of discourse cultures in online spaces and argue that uncivil communication can also enable sophisticated forms of online discussions based on the proposition that uncivil talk can have a mobilizing effect especially among homogenous groups who use uncivil talk to discredit the enemy (Berry & Sobieraj, 2016). Yet, we do not assume that uncivil talk can coexist with all dimensions of deliberative discourse, hence wonder what aspects of desirable online discourse are compatible with incivility in online conversations. We pose the following research question:

RQ3: Which aspects of quality discourse are compatible with the presence of uncivil comments posted on media's Facebook pages during televised debates?

Methodological Approach

Sampling

The data comprises 11 debates from the 2020 Democratic Primary elections. We used Crimson Hexagon to collect the data, focusing on the Facebook pages of the mainstream news outlets. We decided to limit the sample to news organizations that moderated the debates for two reasons: first, they are more likely to take ownership of these events, not only by determining questions asked during the debates, but also capitalizing on the debates to engage their audiences beyond the televised spectacle and in their social media platforms. Second, audiences who watch and engage with the Democratic Party primary debates are interested partisans who mostly consume mainstream and centrist media (Pew Research, 2021).

The data were selected using a two-stage sampling strategy (Rowe, 2015). In the first stage, we collected all Facebook posts by news organizations posted from 7 pm until 2 am, to capture pre- and post-debate news coverage (N=1,503). We selected only news posts that focused on debates and excluded the duplicates. The final media sample includes 480 Facebook posts from 9 news organizations: 6 national broadcasters (ABC=93, CBS=28, NBC=74, CNN=91, MSNBC=102, and PBS=13), two newspapers (New York Times=27 and Washington Post=5) and one online news media (Politico=25).

In the second stage, we generated a random sample of user comments posted under the initial Facebook posts generated in stage one. The Crimson Hexagon limited the number of comments to 10,000 per query, which were randomly selected from the total number of comments posted during the selected time frame. For each debate night, we run several queries to maximize the number of allowed comments per news organization within our time frame

(N=177,055). After cleaning the irrelevant posts and duplicates, we kept only comments posted in reply to the relevant debate related news posts (N=42,146).¹

Data Coding

Manual coding at the post level.

Guided by previous work (Aalberg et al., 2012), in this study we operationalize three aspects of depolarization (Esser & Strömbäck, 2012): issue coverage, strategy focus, and game framing. We created different binary categories to code if Facebook news posts focused on policy ($Kappa = .80$), if they mentioned strategic aspect of the debate or race ($Kappa = .65$), if they discussed horse-race aspects of the debate ($Kappa = .73$).

Although previous scholars include candidate personality and style in the strategy frame (e.g. Ferrer-Conill et al., 2021), in this study we follow Esser and Stromback (2012) in differentiating “personalization” as a distinct category that emphasizes news coverage focus on individual politicians, and can include news on candidate personality traits, their performance, and personal lives at the expense of parties, ideologies, or policies. Hence, we developed a binary category that coded if debate related news posts discussed candidate character traits ($Kappa = .77$). To measure negative campaigning in news coverage, Facebook news posts that quoted or paraphrased one of the candidates were coded for the function of the candidate’s message (Bennoit, 2013): acclaims, defense, or attacks ($Kappa = .75$). This measurement was transformed into a dichotomous variable denoting candidate attack. In addition, this study manually coded for the following variables, which were included in our analysis as control variables: media outlet ($Kappa = 1.0$), debate date ($Kappa = 1.0$), multimedia formatting of the

¹ Table C1 in the Appendix C in the online supplementary files provides a summary of the sample distribution across different debates and news organizations. Online repository files for this paper can be found via this anonymous link: https://osf.io/prh7b/?view_only=4c5eacd9262f4bbaa2f419df76b88de9

post content which measured if posts included visuals ($Kappa = .95$), and type of issues mentioned in the post ($Kappa = .75$). The coding was done by three graduate students who were extensively trained in the use of the codebook. After establishing satisfying intercoder reliability on a random sample of 10% of the Facebook posts, they each proceeded to code independently².

Automated Coding at the Comment Level

For the purpose of this study, we developed a multifaceted computational measurement of discourse quality in user comments posted in response to the news organizations' Facebook posts about debates. We used RoBERTa, a Transformers-based machine learning model (Liu et al., 2019; Goyal et al., 2020), to automatically label the sample of comments after training it on randomly sampled comments hand-labeled by two expert coders.

Labeled Dataset. The coding operationalization of discourse features was adopted from Jaidka et al. (2019), that developed a comprehensive computational instrument to capture the online discourse. Coders labeled a random sample of 1913 comments³ after achieving an overall acceptable inter-coder reliability score (Kohen's Kappa ranging from $Kappa=0.74$ to $Kappa=1.00$). First, we coded if comments were relevant to the debate or not. Respect and empathy were coded if comments respectfully acknowledged other viewpoints or manifested positive or empathetic feelings towards others. **Justification measures whether a user comment provides evidence to support its claims with 1. personal experiences, values and feelings and/or 2. is based on facts, data and links. However, we did not examine whether the justification was based on empirical and theoretical evidence or contained misinformation.** Constructiveness measures

² The codebook and instructions for the manually labeled Facebook Posts and detailed intercoder reliability results can be found on the online repository for this paper:

https://osf.io/prh7b/?view_only=4c5eacd9262f4bbaa2f419df76b88de9

³ See table A8 in Appendix A for the distribution of categories in our training set.

whether a user comment attempts to bring about consensus and resolves conflict by offering fact-checking, identifying common ground, or proposes solutions or asks genuine questions. And finally, a comment was labeled as uncivil if it contained abuses and insults and/or threats. All operationalizations encode the presence or absence of a feature (1/0 binary coding).⁴

The XLM-RoBERTa model. We used RoBERTa, a machine learning model supported by a BERT (Bidirectional Encoder Representation from Transformers) architecture, to train our topic classification model. *Transformers* are a deep learning neural network used in Natural Language Processing (NLP) tasks. Transformers take sequential inputs, like words in a sentence, relating all inputs (words) to each other, allowing for high levels of contextual understanding. Given its ability to better “understand” lexical context, RoBERTa consistently outperforms other machine-learning models in text classification tasks (Liu et al., 2019). For example, an alternative approach to classifying online comments was carried out by Jaidka et al. (2019), who implement classical natural language processing tools to predict deliberation features in Tweets. For all deliberative categories, our RoBERTa models outperform these alternative approaches.⁵ In Appendix A for this article, we expand on the use of Transformer models, the technical specificities of RoBERTa, and the hyper-parameters used to train our model.

Performance. We fine-tuned the RoBERTa model using our training set to classify comments according to the above categories. We trained one model for each category, as the categories are not exclusive. We divided our labelled data into 80% training set, 10% test set, and 10% validation set. The overall performance results from cross-validation and out-of-sample accuracy for all our models are statistically significantly different from the no-information rate. Table 1 in Appendix A presents performance statistics for all models, estimated using cross-

⁵ See tables A1 to A7 in Appendix A for comparative statistics.

validation.⁶ All models have an accuracy above 80% (e.g., the models can correctly predict each category 80% of the time). Finally, note the small difference between F1-scores for the best performing category (Macro-F1) and the worst performing category (Micro-F1).⁷ This suggests that the models are good at predicting that a category occurs and does not occur at a similar rate (see table 1). We run further test using an out-of-sample labelled dataset and find similar performance in all categories.⁸

[Table 1 about here]

Analytical Approach. To analyze the relationship between debate reporting in the Facebook post (level 2) with the quality of discussions at the comment level (level 1), we employ multi-level logistic regression modeling to account for nested data (Heck, Thomas & Tabata, 2013). First, we built random intercept models to assess the variations of the log-odds from one cluster to another and determine whether there is evidence of clustering in the data with respect to the dependent variables. The interclass correlation coefficients (ICC) and the statistically significant Wald Z test for the variance of intercepts across level two units suggested enough clustering to justify using HLM. Second, displayed here, we built intermediate models to assess the variations of the lower-level effects from one cluster to another, measuring the direct effects of the post and comment level attributes. The models presented in Table 1 examine the main effects of debate framing, negativity, incivility and other discourse features. Data presented in Table 2 provide comparative results between our measure and alternative measures of uncivil

⁶ We explain how we implement the cross-validation in Appendix A where are also displayed full performance statistics.

⁷ F1 Score is the weighted average of precision and recall.

⁸ The out-of-sample labelled data is a random sample from our corpus that the models have not previously used. In order to assess the external validity of the models, we compared performance of RoBERTa with the sample of 1400 manually coded observations. See Appendix A.

discourse. Unless noted otherwise, all the results are based on the predictions from our RoBERTa model.

Results

Strategy Game Framing, Negativity and Facebook Comments. Overall, results suggest that conversations on Facebook pages of news organizations during political debates contained high degrees of positive features, as most comments in our sample were relevant (69%) and constructive (50%), more than one-third of the comments provided justification for arguments (39.7%), and about 20% of comments explicitly expressed respect and empathy. Yet, similar to previous studies, we found a substantial number of uncivil comments (25.5%), that to some degree varied among different news organizations.

The first step in the analysis was to examine the relationship between features of debate reporting on Facebook with discourse features in news user comments. The first two research questions explored if there is a significant relationship between election news framing, negative reporting and discourse features in news comments on Facebook during televised political debates. Overall, results presented in table 2 did not find many significant relationships between strategic game reporting nor issues coverage with discourse features. The only significant relationship emerged between posts that emphasized candidate personality and candidate attacks and the relevance of comments posted under those Facebook posts. In other words, Facebook posts that focused on candidate personality had a 12% chance of containing less relevant comments compared to Facebook posts that did not focus on candidate personality (OR=.88; 95% CI=.80 to .97). In addition, Facebook posts that focused on candidate attacks also decreased the chance of comments being relevant by 11% (OR=.89; 95% CI=.79 to .99). Table 2 presents

consistent data suggesting that overall debate news coverage might not be a significant predictor of uncivil language in comment section.

[Table 2 about here]

The Relationship between Discourse Features in Comment Section. The second goal of this study was to examine the relationship between different features of online conversations in user comments posted in reaction to televised debates. Our third questions explored which features of deliberative discourse can co-exist with uncivil discourse in the comment section of media's Facebook pages. We find that about one third of conversation threads that contained uncivil comments were also relevant (27%) and displayed comments with rational justifications (28%); but fewer threads with uncivil comments were also constructive (22.5%) and a limited number expressed respect and empathy (15%). These data provide some evidence for the coexistence of incivility with discourse quality within comment threads.

In the next analyses, displayed in table 2, we predict the likelihood of conversations in news comments containing features of incivility with discourse relevance, respect and empathy, constructiveness, and justification. We find that comments that included uncivil language had a significant positive relationship with relevance and justification. The odds of the discourse in comment section being relevant increase by a factor of 1.3 for each unit increase in the log-odds of incivility (OR=1.30; 95% CI=1.23 to 1.37) and conversations containing uncivil language had 1.5 times more chance of containing also opinion justification (OR=1.51; 95% CI=1.42 to 1.59). Since our variables of interest are binary, the results suggests that comments containing uncivil language were 30% more likely to cooccur with comments that are relevant and 50% more likely to appear with comments that provide justifications than comments that did not have uncivil language. But, results also suggest that an increase in incivility in the comment threads decreased

the likelihood that the conversations were also respectful by 49% (OR=.49; 95% CI=.48 to .54) and decreased the likelihood that conversations were constructive by about 40% (OR=.61; 95% CI=.58 to .64). Thus, to answer RQ3, our data provide consistent evidence that within the context of debate-related digital discussions, comment threads containing uncivil language are incompatible with “respect and empathy” and discussion constructiveness, but they might coexist with conversations that are relevant and reasoned.

Additionally, we found interesting differences in discourse features on Facebook pages of different news outlets. Compared to MSNBC, conversations in the comment section of ABC were less likely to be relevant, respectful, constructive, and rational. In addition, comments posted on Facebook pages of NBC, PBS and Politico were also less likely to provide rational justifications than comments on MSNBC, but comments on CNN were more likely to be constructive. When observing levels of incivility, we find that ABC and NBS are also more likely to contain uncivil comments than MSNBC. Together, these data suggest that differences in discourse features on Facebook pages are dependent on the characteristics of the audience that tunes in to comment on televised debates. **MSNB is considered leftist and their audiences are more homogenous in their political ideology and probably more politically knowledgeable compared with audiences of centrist media, hence it is reasonable to find lower levels of incivility and higher levels of rational arguments.**

Discussion & Conclusion

This comprehensive study contributes to a better understanding of the role of hybrid media spaces for political engagement during electoral campaigns. It provides insightful results on the implications of contextual features for the quality of online discussions related to televised primary debates in the new media ecology. Our data address previous fears that have connected

strategic and negative news reporting (Cappella & Jamieson, 1997) and incivility in online discussions (Anderson et al., 2014) with undesired democratic outcomes and clarifies their relationship with multiple dimensions of audience discourse in digital spaces.

First, an important lesson to draw from the data of this study is that, in the context of hybrid media election events, strategic game framing and negative reporting might not be linked to incivility in online spaces. Confirming previous research (Camaj, 2020; Goncalaves et al., 2020) with robust analysis, we found that strategic game reporting does not increase audiences' likelihood to use swear words or post uncivil comments in reaction to debate related news stories. A state-of-the-art computational instrument used to measure uncivil comments across 11 political debates did not find any significant relationship between strategic game coverage and incivility. When it comes to the probability that uncivil discourse would unfold in the comment space, it might not matter if Facebook posts covering political debates focus on policy discussions, candidate strategies, or candidate personality.

However, we found that Facebook posts that focus on candidate personalities and negativity might be deterrents to relevant conversations. Posts focusing on candidate personality and candidate attacks were less likely to feature comments that were relevant to the discourse about debating candidates. These data contradict recent findings that suggests that strategic and negative coverage of politics has the potential to engage audiences with election news in general (Iyengar et al., 2004; Trussler & Soroka, 2014) and on social media more specifically (Gonçalves et al., 2020). While news audiences might be more likely to click, read, react or share strategic and negative news coverage in digital spaces, we find that that this engagement might not be conduit to relevant discussions on Facebook. These findings suggest the need to consider the context in which discussions occur could play an important role, and debate coverage in

online spaces provides a specific environment where candidate personalities come into focus (Benoit, 2013). Hence, news organizations might consider adjusting their debate coverage away from candidate attacks and focus on personalities, since it has the potential to decrease audiences' meaningful interactions with news coverage a crucial feature for audience loyalty and satisfaction.

This study also informs the conversation about the role of incivility in online conversations to determine viewer experience with debate content. First, results from our large corpus of data suggest that uncivil comments were likely to appear along relevant comments that attempt to elaborate opinions with specific facts and data or anecdotal reasoning. For example, in a comment the discussant employs offensive language by calling another commentor "Dumb F" but also provides an elaborate reference to the statements made by the former U.S. President, Donald Trump, to support his argument. These results support the thesis that incivility and rational conversations may co-occur in online user comments, since some forms of incivility may still allow for discussions rooted in reason (Chen, 2017; Rossini, 2020). However, we also find that incivility is incompatible with constructiveness and respect and empathy in online discussions. For example, one of the commentor says "Progressives for a democratic Republic? This is how stupid you guys are! Bernie is a commie, it won't be dem or repub. More of an enslavement camp dummies!" Yet, this does not preclude people's potential to engage in reasoned discourse with those they do not respect or those they disagree with (Rossini, 2020). In another example, a commentor says "(@user) Look around at your state. People living in the streets and crapping in the streets. You have the most corrupt state government in the country. Your master's degree means nothing. A person with common sense is much wiser than you."

Previous research (Asbury et al., 2021; Ventura et al., 2021) shows that the overly toxic nature of online comments can create a negative experience for users during political events. We

provide more nuance to these findings. In particular, we look at online discourse as multi-dimensional, capable of having different characteristics simultaneously. Thus, the polarity of a comment can have heterogeneous effects depending on the additional traits of the text. The experience and engagement of users in online spaces will be shaped by these characteristics, both positively and negatively. Our data support recent trends in literature that question the separation between rationality and emotionality in online discussions (Bickford, 2011; Hoggett & Thompson, 2002), given that these conversations can be heated, emotionally charged, but at the same time rational. These results are particularly relevant for news organizations that have struggled to deal with uncivil comments on their digital spaces. Many have taken drastic measures completely closing their comment section or indiscriminately filtering out uncivil comments. We offer alternative ways to measure the quality of online conversations in digital news spaces that captures the multi-dimensional nature of conversations happening informally in the public sphere.

Although, overall, our data indicate that uncivil comments might not necessarily represent conversations that are without substance, we draw caution about the prevalence of comments that are uncivil and substantive at the same time. In comparison and unsurprisingly, we find that the best predictors of reasoned arguments are comments that are also constructive by offering fact-checking, identifying common ground, proposing solutions, or asking genuine questions.

These communicative practices are situational and contextual. In the case of primary debates of the oppositional party, people engaging in news commenting might present homogeneous groups who come together to deliberate on the common enemy (Berry & Sobieraj, 2016). Although our measure of incivility captures the violation of politeness behavior as well as

personal threats and flaming that are more detrimental to discourse participation, our dataset did not contain much of the later. Most uncivil content pertained to the public level of incivility rather than personal level (Muddiman, 2017). A substantial number of uncivil comments and insults in our sample were addressed to President Donald Trump. For example, “The corrupt, lying, cheating witch didn't win the popular vote you also might want to educate yourself on the useless popular vote and so much more before commenting.” Outrage towards political enemy might bring people together with like-minded commentators (Berry & Sobieraj, 2016) who deliberate on the best-suited political contender to beat the opponent. Despite of the uncivil language employed in these conversations, some debate watchers might feel validated and encouraged to participate in reasoned conversations.

The results of our study should be informative to news organizations involved in moderating political debates. While unmoderated real-time commentary in debate streaming chat can be toxic and provide negative experiences for debate watchers (Asbury-Kimmel et al., 2021), our study suggests that post-debate coverage and commentary associated with slow-paced online commenting can produce discussions that, although contain uncivil language, provide reasoned conversations that might contribute to voter learning, engagement, and overall better user experiences. In this context, news organizations can adjust their reporting style to help create an online environment that channels relevant and reasoned discussions by putting an emphasis on issues rather than on strategy and attacks.

Finally, we want to make a last argument for the value of primary debates and the role of mainstream news media to channel meaningful conversations among homogenous partisan who are most likely to determine the ideological path of their party. Especially, these results are important in the wake of upcoming 2024 elections in the United States where the republican

frontrunner, Donald Trump, has refused to participate in primary debates and fueled critics who question the utility of political debates in the political process.

Our results should be interpreted through the lens of other methodological limitations, though. We acknowledge Facebook's restrictions in using their API for data collection by academics which limited us to use third party organizations for data collection. Although Crimson Hexagon (now part of Brandwatch) is widely used as a social media data collection tool in academia, it provides limits the amount of data access to 10,000 comments per query and is not totally transparent regarding how it randomizes the sample selection. Given these restrictions, we could not verify the authenticity of Facebook comments analyzed in this study. Yet, Facebook's commitment to crackdown on bots and other malignant actors during the 2020 elections gives us confidence in our data. And second, we acknowledge the limitations that come with the sample selection of mainstream news media included in this study, and future studies should consider alternative media as spaces for political talk and quality discourse.

Despite these shortcomings, our study offers robust findings that user conversations on broadcasters' Facebook pages in response to televised political debates provide spaces where substantive conversations can happen, potentially amplifying the relevance of political debates for voter learning and opinion formation (Chadwick et al., 2017). Commenters who actively participate in debate related discussions might learn from reasoned and constructive discussions, although some might be heated and uncivil. Additionally, online discussions of political debates can serve as mobilizing tools for politically motivated voters who rally around common causes. This study also points out the need to further examine the indirect effects of televised debates, mediated through post-debate news coverage and conversations that extend beyond the two-hour highly anticipated events.

Data availability statement. The data underlying this article are available in Open Science Framework (OSF) data repository at https://osf.io/prh7b/?view_only=4c5eacd9262f4bbaa2f419df76b88de9.

References

- Aalberg, T., Strömbäck, J., & de Vreese, C. H. (2012). The framing of politics as strategy and game: A review of concepts, operationalizations and key findings. *Journalism*, 13(2), 162–178. <https://doi.org/10.1177/1464884911427799>
- Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The “nasty effect”: Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19(3), 373–387. <https://doi.org/10.1111/jcc4.12009>
- Berry, J. M., & Sobieraj, S. (2016). *The outrage industry: Political opinion media and the new incivility*. Oxford University Press.
- Benoit, W. L. (2013). *Political Election Debates: Informing Voters about Policy and Character*. Lanham, Maryland: Lexington Books.
- Bormann, M., Tranow, U., Vowe, G., & Ziegele, M. (2022). Incivility as a violation of communication norms—A typology based on normative expectations toward political communication. *Communication Theory*, 32(3), 332–362. <https://doi.org/10.1093/ct/qtab018>
- Bucy, E. P., Foley, J. M., Lukito, J., Doroshenko, L., Shah, D. V., Pevehouse, J. C., & Wells, C. (2020). Performing populism: Trump’s transgressive debate style and the dynamics of Twitter response. *New Media & Society*, 22(4), 634–658.
- Camaj, L. (2021). Real time political deliberation on social media: Can televised debates lead to rational and civil discussions on broadcasters’ Facebook pages? *Information, Communication & Society*, 24(13), 1907–1924. <https://doi.org/10.1080/1369118X.2020.1749695>

- Camaj, L. & Northup, T. (2019) Dual-screening the candidate image during Presidential debates: The moderating role of Twitter and need to evaluate for the effects on candidate perceptions. *Journal of Broadcasting & Electronic Media*, 63(1), 20-38, DOI:10.1080/08838151.2019.1574117
- Cappella, J. N., & Jamieson, K. H. (1997). *Spiral of Cynicism: The Press and the Public Good*. New York, NY: Oxford University Press.
- Chadwick, A., O'Loughlin, B., & Vaccari, C. (2017). Why people dual screen political debates and why it matters for democratic engagement. *Journal of Broadcasting & Electronic Media*, 61(2), 220–239. <https://doi.org/10.1080/08838151.2017.1309415>
- Chen, G. M. (2017). *Online Incivility and Public Debate: Nasty Talk*. New York, NY: Palgrave Macmillan.
- Cho, J., & Choy, S. P. (2011). From podium to living room: Elite debates as an emotional catalyst for citizen communicative engagements. *Communication Research*, 38(6), 778–804. <https://doi.org/10.1177/0093650210378518>
- 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.
- Cushion, S., & Thomas, R. (2018). *Reporting Elections: Rethinking the Logic of Campaign Coverage*. Cambridge, UK: Polity Press.
- Dahlberg, L. (2011). Re-constructing digital democracy: An outline of four ‘positions.’ *New Media & Society*, 13(6), 855–872. <https://doi.org/10.1177/1461444810389569>
- Dun, L., Soroka, S., & Wlezien, C. (2021). Dictionaries, supervised learning, and media coverage of public policy. *Political Communication*, 38(1–2), 140–158. <https://doi.org/10.1080/10584609.2020.1763529>

- Edyvane, D. (2020). Incivility as dissent. *Political Studies*, 68(1), 93–109. doi: 10.1177/0032321719831983
- Esser, F., & Strömbäck, J. (2012). Comparing news on national elections. In F. Esser & Th. Hanitzsch (eds) *Handbook of comparative communication research* (pp. 308–326). New York, NY: Routledge.
- Fishkin, J. S. (2009). *When the People Speak: Deliberative Democracy and Public Consultation*. Oxford, UK: Oxford University Press.
- Fraser, N. (1992). Rethinking the public sphere: A contribution to the critique of actually existing democracy. In C. Calhoun (Ed.), *Habermas and the public sphere* (pp. 109-142). Cambridge, MA: MIT Press.
- Freelon, D., & Karpf, D. (2015). Of big birds and bayonets: Hybrid Twitter interactivity in the 2012 presidential debates. *Information, Communication & Society*, 18(4), 390-406.
- Friess, D., & Eilders, C. (2015). A systematic review of online deliberation research. *Policy & Internet*, 7(3), 319–339. <https://doi.org/10.1002/poi3.95>
- Gonçalves, J., Pereira, S., & Torres da Silva, M. (2020). How to report on elections? The effects of game, issue and negative coverage on reader engagement and incivility. *Journalism*, 146488492095836. <https://doi.org/10.1177/1464884920958367>
- Gross, K., Porter, E., and Wood, T. J. (2019). Identifying media effects through low-cost, multiwave field experiments. *Political Communication*, 36(2):272–287.
- Heck, R. H., Thomas, S. L. and Tabata, L. N. (2013). *Multilevel and Longitudinal Modeling with IBM SPSS: Quantitative Methodology Series*. Second Edition. New York, NY: Routledge.

- Jaidka, K., Zhou, A., & Lelkes, Y. (2019). Brevity is the soul of Twitter: The constraint affordance and political discussion. *Journal of Communication*, 69(4), 345–372.
<https://doi.org/10.1093/joc/jqz023>
- Jennings, F. J., Suzuki, V. P. & Hubbard, A. (2021). Social media and democracy: Fostering political deliberation and participation. *Western Journal of Communication*, 85(2), 147-167. DOI: 10.1080/10570314.2020.1728369
- Jennings, F. J., Bramlett, J. C., McKinney, M. S., & Hardy, M. M. (2020). Tweeting along partisan lines: Identity-motivated elaboration and presidential debates. *Social Media and Society*, 6(4). doi:10.1177/2056305120965518
- Kim, J.W., Guess, A., Nyhan, B., & Reifler, J. (2021). The distorting prism of social media: how self-selection and exposure to incivility fuel online comment toxicity. *Journal of Communication*, 71, 6, 922–946, <https://doi.org/10.1093/joc/jqab034>
- McGregor, S. C., & Mourão, R. R. (2017). Second screening Donald Trump: Conditional indirect effects on political participation. *Journal of Broadcasting & Electronic Media*, 61(2), 264–290. <https://doi.org/10.1080/08838151.2017.1309418>
- McKinney, M. S., & Warner, B. R. (2013). Do presidential debates matter? Examining a decade of campaign debate effects. *Argumentation & Advocacy*, 49(4), 238–258.
- McKinney, M. S., Houston, J. B., & Hawthorne, J. (2013). Social watching a 2012 republican presidential primary debate. *American Behavioral Scientist*, 58, 556–573.
doi:[10.1177/0002764213506211](https://doi.org/10.1177/0002764213506211)
- Mouffe, C. (1999). Deliberative democracy or agonistic pluralism? *Social Research*, 66(3), 745–758.

- Muddiman, A. (2017). Personal and public levels of political incivility. *International Journal of Communication*, 11, 3182–3202. <https://ijoc.org/index.php/ijoc/article/view/6137>
- 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>
- Mutz, D. C. (2015). In-your-face politics: The consequences of uncivil media. Princeton, NJ: Princeton University Press. <https://doi.org/10.1515/9781400865871>
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2), 259–283. doi: 10.1177/1461444804041444
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The Development and Psychometric Properties of LIWC2015*. <https://repositories.lib.utexas.edu/handle/2152/31333>
- Reimer, J., Häring, M., Loosen, W., Maalej, W., & Merten, L. (2021). Content analyses of user comments in journalism: A systematic literature review spanning communication studies and computer science. *Digital Journalism*, 1–25. <https://doi.org/10.1080/21670811.2021.1882868>
- Robertson, C. T., Dutton, W. H., Ackland, R., & Peng, T.-Q. (2019). The democratic role of social media in political debates: The use of Twitter in the first televised US presidential debate of 2016. *Journal of Information Technology & Politics*, 16(2), 105–118. <https://doi.org/10.1080/19331681.2019.1590283>

- Rossini, P. (2020). Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. *Communication Research*.
<https://doi.org/10.1177/0093650220921314>
- Rossini, P., Sturm-Wikerson, H., & Johnson, T. J. (2021). A wall of incivility? Public discourse and immigration in the 2016 U.S. Primaries. *Journal of Information Technology & Politics*, 18(3), 243–257. <https://doi.org/10.1080/19331681.2020.1858218>
- Rowe, I. (2015). Deliberation 2.0: Comparing the deliberative quality of online news user comments across platforms. *Journal of Broadcasting & Electronic Media*, 59(4), 539–555. <https://doi.org/10.1080/08838151.2015.1093482>
- Steenbergen, M. R., Bächtiger, A., Spörndli, M., & Steiner, J. (2003). Measuring political deliberation: A discourse quality index. *Comparative European Politics*, 1(1), 21–48.
[doi:10.1057/palgrave.cep.6110002](https://doi.org/10.1057/palgrave.cep.6110002)
- Tremayne, M., & Minooie, M. (2015). Using social media to analyze candidate performance during televised political debates. *Electronic News*, 9(3), 143–159.
<https://doi.org/10.1177/1931243115593321>
- Trussler, M., & Soroka, S. (2014). Consumer demand for cynical and negative news frames. *The International Journal of Press/Politics*, 19(3), 360–379.
<https://doi.org/10.1177/1940161214524832>
- Van Aelst, P., Sheafer, T., & Stanyer, J. (2012). The personalization of mediated political communication: A review of concepts, operationalizations and key findings. *Journalism*, 13(2), 203–220. <https://doi.org/10.1177/1464884911427802>

Ventura, T., Munger, K., McCabe, K., & Chang, K. C. (2021). Connective effervescence and streaming chat during political debates. *Journal of Quantitative Description: Digital Media*, 1. <https://doi.org/10.51685/jqd.2021.001>

Ziegele, M., Quiring, O., Esau, K., & Friess, D. (2020). Linking news value theory with online deliberation: How news factors and illustration factors in news articles affect the deliberative quality of user discussions in SNS' comment sections. *Communication Research*, 47(6), 860–890. <https://doi.org/10.1177/0093650218797884>

Zoizner, A. (2021). The Consequences of strategic news coverage for democracy: A meta-analysis. *Communication Research*, 48(1), 3–25.
<https://doi.org/10.1177/0093650218808691>