Lurking for the Common Good: Does Observing Civil Cross-Party Dialogue on Social Media Reduce Affective Polarization? A Registered Report

This Version: December 7, 2025

Word Count: 3,800

J. Grady Anderson* Miguel Arcero-Miranda Emma Barrett†
Sarah Bhalla Courtney Boutilier Emma Clark
Ella Garcia Holden Mamula Denise Perez
Francisca Puiatti Alyssa Robinson Maria Clara Sandoval Durski
Carlisle Rainey‡

Abstract

Affective polarization is pervasive in American politics, and has downstream impacts that could be harmful to democracy. As such, scholars have devoted a substantial amount of attention to identifying strategies to lower the temperature in America. Levendusky and Stecula (2021) found that bringing partisans together to engage in face-to-face conversations, political or otherwise, decreases polarization. We argue, however, that many conversations now occur on social media, where people may "lurk", rather than engaging directly. Does observing civil conversation on social media result in a similar reductive effect as does face-to-face dialogue? In this registered report, we propose a novel survey experiment where respondents are randomly assigned to one of three groups: a true control, a "civil" treatment in which respondents observe a respectful conversation between two opposing partisans on X, and an "uncivil" treatment in which respondents observe a disrespectful conversation between two opposing partisans on X. We hypothesize that observing a civil conversation will reduce polarization, and conversely, observing an uncivil conversation will heighten polarization, as measured by the standard feeling thermometer and the newly-introduced Affective Polarization Scale (Campos and Federico 2025).

*J. Grady Anderson is Undergraduate Student, Department of Political Science, Florida State University.
†Emma Barrett is Ph.D. Student, Department of Government, Harvard University.
‡Carlisle Rainey is Associate Professor of Political Science, Florida State University, 540 Bellamy, Tallahassee, FL, 32306. (crainey@fsu.edu). Other co-authors are or were undergraduate students in the Research Intensive Bachelor's Certificate Program at Florida State University.

Introduction

Affective polarization, in which partisans view out-partisans negatively and copartisans positively (Iyengar and Westwood 2015), continues to be a pervasive problem in American politics. Feeling thermometer ratings have demonstrated an increasing disparity in Americans' views of out-partisans in relation to copartisans (Iyengar and Krupenkin 2018).

Affective polarization has devastating effects for democracy at large. Chiefly, affective polarization is a cyclical process, meaning that Republicans and Democrats socially distance themselves, leading to increased levels of affective polarization. This social distancing affects the spousal market (Iyengar, Sood, and Lelkes 2012), friendships (Huber and Malhotra 2017), and even where people move (Gimpel and Hui 2015). Despite signaling consistent support for democracy (Holliday et al. 2024), polarized Americans are prone to look the other way when their party violates democratic norms (Graham and Svolik 2020).

Because affective polarization poses such a threat to democracy, scholars have devoted much attention to finding ways to decrease levels of polarization. One such avenue that scholars have identified is to promote cross-party dialogue. Theoretically, discussion is a cornerstone of democracy. Political discussions allow people opportunities to work through issues through debate and encounter perspectives they may not otherwise encounter (Mutz 2006). Levendusky and Stecula (2021), for example, investigate the effects of in-person, cross-party conversation on levels of polarization. Not only do they find that cross-partisan discussion decreases polarization, but that these effects persist across time. Similarly, Fiskin et al. (2021) find that bringing together opposing partisans for in-person conversation reduces polarization on salient issues like immigration and healthcare.

Levendusky and Stecula (2021) and Fiskin et al. (2021) do not address the effects of discussion on *social media*, which has become an important arena for political discussions. Combs et al. (2023), for example, demonstrated that dialogue in a digital setting can have similar effects as did the in-person conversations in Levendusky and Stecula (2021) and Fiskin et al. (2021). However, we contend that most people do not actively engage in dialogue over social media. Instead, they observe conversations without actively participating in the discussion.

Another consideration is the civility of these discussions that social media users are observing. By nature of occurring in an online setting, these conversations often entail degrees of separation and perhaps anonymity that are not conducive to following communication norms. In fact, researchers have demonstrated an increase in incivility over time on social media such as Reddit (Sun, Wojcieszak, and Davidson 2021). So, it may be that the effects of dialogue on social media may hinge on the (in)civility of the discussion.

We propose a survey experiment to examine whether observing civil cross-party dialogue on social media

can yield similar results as did Levendusky and Stecula (2021) and Fishkin et al. (2021). Similarly, we examine whether observing uncivil cross-party dialogue on social media will further polarize respondents. We expect that (H1) exposure to civil conversations on social media will decrease affective polarization and, following recent evidence on the effects of uncivil dialogue on social media (Allamong et al. working paper), we expect that (H2) viewing uncivil cross-party dialogue will heighten affective polarization.

In our experiment, we will expose participants to either a control condition, in which respondents see a tweet that prompts the ensuing discussion about a perceived "mandate" for President Trump following the 2024 election; a civil conversation between opposing partisans about the perceived mandate, or an uncivil conversation about the perceived mandate between opposing partisans. We deviate from similar studies in one substantial way: while we will measure affective polarization using the feeling thermometer for the purposes of comparisons to prior results, our main outcome variable will be the newly-formalized Affective Polarization Scale (Campos and Federico 2025), which provides a more nuanced understanding of polarization than the feeling thermometer.

Literature Review

Affective polarization in the American public has been the subject of much scrutiny by scholars of political behavior. Researchers have demonstrated this increase in numerous fashions, including in feeling thermometer evaluations of in- and out-partisans (Iyengar and Krupenkin 2018), social distancing measures (Iyengar, Konitzer, and Tedlin 2018), and trust (Lee 2022). Importantly, affective polarization is not benign; it has devastating effects for democracy. In addition to partisans social distancing themselves (Iyengar, Konitzer, and Tedlin 2018), polarized partisans are more likely to undermine support for democratic norms (Kingzette et al. 2021; see, however, Broockman, Kalla, and Westwood 2023), overlook violations of democratic norms from co-partisans (Graham and Svolik 2020), and endorse political violence (Piazza 2023).

Because of the troubling consequences of affective polarization, there has been a concerted effort among academics to depolarize the electorate. One of the most promising avenues for depolarizing is engaging partisans in cross-party conversations. Levendusky and Stecula (2021) investigate the effects of cross-party conversations by physically bringing people together to engage in in-person dialogue with opposing partisans. They find that cross-partisan discussion does decrease affective polarization, both when the topic of discussion is political *and* non-political. Not only does it decrease affective polarization initially, but the effects are long-lasting. Fishkin et al. (2021) similarly find that bringing together opposing partisans for in-person conversations reduces polarization, including on salient issues like immigration and healthcare.

A new avenue for depolarization combines the deliberative democracy effects of Levendusky and Stec-

ula (2021) and Fishkin et al. (2021) with the promise of social media as a public forum. Combs et al. (2023) examine the extent to which cross-partisan conversations on social media exacerbate or mitigate affective polarization. They found that those who engaged in cross-partisan conversations about political topics depolarized compared to those in a placebo group. Similarly, Rossiter and Carlson (2024) found that online conversations between opposing partisans after the 2020 presidential election significantly decreased polarization, even in one of the most contentious periods in American politics.

Whatever the potential may be for social media as a tool to depolarize partisans, there is evidence that content on social media has become increasingly uncivil. Frimer et al. (2023) in their analysis of 1.3 million tweets from members of Congress, find that incivility has increased by 23% since 2009, an effect driven by positive feedback loops in the forms of likes and retweets of uncivil content on social media. According to Sun, Wojcieszak, and Davidson (2021), political communities on social media are more susceptible to incivility than most other communities.

So, while social media can serve as a crucial pathway for depolarizing (Combs et al. 2023; Rossiter and Carlson 2024), the amount of incivility on social media may preclude depolarization. Gao et al. (2024), in their study of 40 million comments on Reddit, find that users send more uncivil comments users who identity with the opposite political party as them, threads that start with uncivil comments gather more responses and are more uncivil, and that users are more likely to drop out from subreddits when exposed to higher levels of incivility. Allamong et al. (working paper) randomly assigned users of their "social media accelerator" to feeds of civil or uncivil discussions with "synthetic users powered by GPT-4". They found that those in the uncivil condition were less comfortable sharing their views on the app and made more posts. Importantly, they also found that exposure to the uncivil condition increased out-party animus.

This suggests that (in)civility on social media may moderate the effects of cross-party conversation on affective polarization. Where civil discussions online may depolarize partisans (Combs et al. 2023; Rossiter and Carlson 2024), uncivil discussions may lead to an increase in affective polarization (e.g., Allamong et al. working paper).

While it is important to understand how cross-partisan discussions affect polarization, particularly on social media, we argue that individuals often do not engage with others on these platforms. Instead, they "lurk", observing discussions without actively participating. Our study then differs from previous studies in two ways. First, we examine the effects of survey respondents observing cross-partisan social media discussions. Second, we conceptualize affective polarization using the newly-formalized Affective Polarization Scale (Campos and Federico 2025), which uses a tripartite measure of othering, aversion, and moralization to draw conclusions of levels of affective polarization. While the Affective Polarization Scale will be our primary outcome measure of affective polarization, we will also collect data on feeling thermometer scores

of respondents in order to compare to prior studies and to serve as a baseline for determining sample size.

In our experiment, participants are exposed to a control, which is a neutral tweet aimed at priming the conversation that those in the civil and uncivil conditions view, or civil or uncivil conversation between a member of their party and the opposing party on social media. We formally state our hypotheses below:

Hypothesis 1: observing a civil conversation will decrease affective polarization, and

Hypothesis 2: observing an uncivil conversation will increase affective polarization.

Research Design

We propose a survey experiment to examine the effects of civil and uncivil social media discussions on affective polarization.

Experimental Conditions

We assign respondents to one of three conditions: *Civil, Uncivil, and Control.* All conditions are exposed to a tweet with a neutral statement, designed to prime the ensuing conversation.

Jake Redding: In his acceptance speech after winning the 2024 Presidential Election, President-elect Trump said that America has given him "an unprecedented and powerful mandate".

Because this tweet priming respondents is consistent across the three conditions, any treatment effects that may be caused by the tweet can be assumed to exist in all conditions, ensuring that any differences observed across conditions can be attributed to the treatment itself rather than the primer.

All tweets were generated by the researchers using tweetgen.com. Pictures of the control and treatments will be included in the appendix.

Civil Condition:

In the civil condition, we first prime respondents on the subject of the discussion they are about to read, as discussed above.

After reading the primer, respondents are instructed to read a social media conversation between Ashley Roberts, a Democrat, and Bill Anderson, a Republican. Similar to the primer, the conversation takes the form of successive tweets. The conversation is as follows.

Ashley Roberts: Listen, I get that Trump won the presidency but didn't he get only 49.8% of the popular vote? It's hard to see this as a 'broad mandate' from the American people.

Bill Anderson: Sure, but at the same time, winning 312 electoral votes isn't nothing. If Harris had won by that much, Democrats would be calling it a decisive victory. So both sides are guilty of this, but this is

how we've always decided elections in the US.

Ashley Roberts: For sure. I just don't get how he feels like he has the right to slash the government in the name of the American people when the margins were so small. He's making it seem like 90% of America is on board with his agenda. Most people, like you and me, are middle of the road.

Bill Anderson: Yeah that makes sense. I'm definitely more middle of the road, but I do think his policies will help everyone. They're not just catering to the GOP. I guess time will tell!

Uncivil Condition

The uncivil condition is primed with the same tweet as the civil and control conditions. However, the text of the conversation changes drastically between the civil and uncivil conditions.

Ashley Roberts: Whatever, Trump might have won the presidency but he barely scraped by with 49.8% of the popular vote. There's absolutely no way this is a "broad mandate" from the American people.

Bill Anderson: Cry about it. Winning 312 electoral votes is ALL that matters. If Harris had won by that much, Democrats would be calling it a decisive victory. You're ignoring what matters. This is how we've ALWAYS decided elections in the US.

Ashley Roberts: I don't care. He has no right to slash the government in the name of the American people when the margins were so small. He's making it seem like 90% of America is on board with his agenda. Most people realize this is crazy.

Bill Anderson: That's just nonesense. Anyone with half a brain can see that his policies will help everyone. They're not just catering to the GOP. You'll see.

Control Condition

The control condition only sees the primer tweet about the mandate for President Trump.

Outcomes

We have two primary outcomes to measure affective polarization. We will utilize the feeling thermometer to measure affective polarization. Here, we deviate from Levendusky and Stecula (2021) in one meaningful way: where Levendusky and Stecula combine a variety of outcomes as their measurement (feeling thermometer, trait ratings, trust, and willingess to interact), we focus exclusively on the feeling thermometer due to budgetary constraints.

Following related work on affective polarization (Druckman et al. 2022; Broockman, Kalla, and Westwood 2023), we will utilize the out-party feeling thermometer ratings and the difference between in-party and

out-party ratings. While the out-party feeling thermometer has been widely used in research on affective polarization, we use the difference in feeling thermometer scores for the in- and out-party, as the difference between in- and out-party feeling thermometer scores yields greater statistical power when compared to the traditional out-party feeling thermometer (Broockman, Kalla, and Westwood 2023).

Finally, we use the newly-introduced Affective Polarization Scale (hereafter APS, Campos and Federico 2025) as an additional measure of affective polarization. The APS provides a nine-item battery structured around the three components of Finkel et al.'s (2020) "political sectarianism": othering, aversion, and moralization.

We discuss below our pre-post design for the feeling thermometer outcomes and our quasi pre-post design for the APS.

Pre-Treatment Measures

We include pre-treatment measures of both the feeling thermometer and APS. This follows guidance from Clifford, Sheagley, and Piston (2021), who demonstrate that pre-post designs dramatically improve statistical power without altering treatment effects. Additionally, attenuation effects of pre-post designs are relatively small when compared to post-only designs (Clifford, Sheagley, and Piston 2021; Jordan, Ollerenshaw, and Trexler working paper). Thus, the pre-post design will allow us to maintain statistical power with a reasonable sample size.

The feeling thermometer questions are identical in the pre-treatment as they are asked post-treatment. We ask respondents: "How would you rate your feelings toward [Democratic/Republican] Party voters?" The question is repeated for the in-party and the out-party.

Following Cutler, Pietryka, and Rainey (working paper), we utilize a quasi pre-post design for the APS. In the pre-treatment, we use a three-question alternative-APS battery for the pre-treatment measure. Campos and Federico (2025) note that the three items we utilize in the pre-treatment battery are closely related to the final nine-item battery that they formalize.

While attenuation effects should not be a major concern in pre-post designs (Clifford, Sheagley, and Piston 2021; Jordan, Ollerenshaw, and Trexler working paper), the three-item alternative-APS battery utilizes different questions, alleviating concerns over attenuation effects. Respondents can respond on a 7-point agree/disagree scale. The three-questions utilized in the alternative-APS scale are shown in Table 1.

Post-Treatment Measures

The post-treatment measure of the feeling thermometer follows the pre-treatment measure. We ask respondents to rate the in-party and out-party on the standard 101-point feeling thermometers.

Table 1: Alternative three-item Affective Polarization Scale for Pre-Treatment

Othering	Republicans/Democrats seem out of place when they are around us [Democrats/Republicans].
Aversion	Identifying as a [Republican/Democrat] rather than a [Democrat/Republican] makes someone a bad person.
Moralization	As a [Democrat/Republican], my feelings about politics are connected to my core moral beliefs.

The post-treatment measure of the APS differs from the pre-treatment. Where we use an alternative three-question battery in the pre-treatment, we use the full APS post-treatment. The nine-question battery is shown in Table 2. Questions in the post-treatment APS are randomized.

Table 2: Final 9-item Affective Polarization Scale

Othering	• I feel as though [in-party] are very different from [out-party]. • [Out-party] live in a different world from us [in-party]. • [Out-party] act in ways that us [in-party] could never understand.
Aversion	• As a [in-party], I would not want to be friends with someone who was a [out-party]. • If I found out a friend of mine was a [out-party], I would want to stop spending time with them. • Although I do not agree with their political views, there are people I like who are [out-party].*
Moralization	• My identity as a [in-party] is connected to my core moral beliefs. • My identity as a [in-party] reflects my beliefs about the difference between right and wrong. • My identity as a [in-party] is rooted in moral principles.

Note: (*) denotes reverse coded items.

Sample and Power Analysis

Utilizing Rainey's Power Rules (Rainey working paper), we can use pilot data to determine the sample size needed per group to detect a 2-point change on the feeling thermometer, which we consider to be a small but substantive meaningful effect.

Our pilot study has an average of 91 respondents per condition. We estimated the standard error of the average treatment effect for the uncivil condition relative to the control condition ($\text{SE}_{uncivil} = 2.11$).*

$$n_{\text{needed per condition}} = n_{\text{pilot}} \left[\frac{2.5}{\tilde{\tau}} \left(\sqrt{\frac{1}{n_{\text{pilot}}} + 1} \right) SE_{\hat{\tau}}^{\text{pilot}} \right]^2$$

Plugging in the standard error, average respondents per condition in the pilot, and our minimal detectable effect, we find that we need 773 respondents per condition to estimate a 2-point difference in the 101-point feeling thermometer.

*The standard error for the uncivil condition relative to the control condition was greater than the standard error for the civil condition. $\text{SE}_{civil} = 2.10$. Here, we utilize the standard error for the uncivil condition for a conservative estimate.

We plan to recruit 2,319 respondents through PureSpectrum and will pay $1 to each respondent for completing the survey. While platforms such as PureSpectrum are not representative of the U.S. population, the literature on survey experiments indicates that results from crowdsourcing platforms such as PureSpectrum are comparable to population-based samples (Mullinix et al. 2015; Berinsky, Huber, and Lenz 2012; Coppock 2019).

Proposed Statistical Models

We plan to estimate the average treatment effect of the civil and uncivil treatments using ordinary least squares and the standard errors using HC2 Huber-White sandwich estimator (MacKinnon and White 1985). We plan to control for pre-treatment measures. To evaluate our hypotheses, we will use 90% Wald confidence intervals (McCaskey and Rainey 2015).

Null Results

If our results are not statistically significant, we will follow Rainey's (2014) equivalence-testing approach to assess if the data allow us to reject meaningful effects. Using the standard error ($SE_{uncivil} = 2.11$) from our pilot data and a 90% confidence interval, we establish a width of 6.92 points on the feeling thermometer scale. If the treatment effect is not statistically significant, we will be well-powered to reject effects of approximately 7 points on the feeling thermometer scale.

Results

Hypothesis 1

[Results to be added here.]

Hypothesis 2

[Results to be added here.]

Exploratory: Average Treatment Effects for the Affective Polarization Scale

We do not have theoretical expectations for the average treatment effects of the civil and uncivil conditions on the Affective Polarization Scale (APS). As such, we consider any analyses related to the APS scale exploratory. We will estimate the average treatment effect for each of the three APS subscales as well as the full APS. For each, we will control for pre-treatment measures of the APS.

Conclusion and Discussion

[To be added upon data collection and analysis.]

Ethical Standards

The pre-test was reviewed and approved by the Institutional Review Board at Florida State University (STUDY00006136). Any substantial changes from the pre-test to the full test will be reviewed by IRB.

References

Allamong, M., Hillygus, D. S., Bail, C., & Volfovsky, A. (2025, May). Incivility in Cross-Partisan Interactions on Social Media. In 80th Annual AAPOR Conference. AAPOR.

Broockman, D. E., Kalla, J. L., & Westwood, S. J. (2023). Does affective polarization undermine democratic norms or accountability? Maybe not. *American Journal of Political Science, 67*(3), 808-828.

Campos, N., & Federico, C. (2025). A new measure of affective polarization. *American Political Science Review*, 1-19.

Clifford, S., Sheagley, G., & Piston, S. (2021). Increasing precision without altering treatment effects: Repeated measures designs in survey experiments. *American Political Science Review, 115*(3), 1048-1065.

Combs, A., Tierney, G., Guay, B., Merhout, F., Bail, C. A., Hillygus, D. S., & Volfovsky, A. (2023). Reducing political polarization in the United States with a mobile chat platform. *Nature Human Behaviour, 7*(9), 1454-1461.

Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2022). (Mis) estimating affective polarization. *The Journal of Politics, 84*(2), 1106-1117.

Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., ... & Druckman, J. N. (2020). Political sectarianism in America. *Science, 370*(6516), 533-536.

Fishkin, J., Siu, A., Diamond, L., & Bradburn, N. (2021). Is deliberation an antidote to extreme partisan polarization? Reflections on "America in one room". *American Political Science Review, 115*(4), 1464-1481.

Frimer, J. A., Aujla, H., Feinberg, M., Skitka, L. J., Aquino, K., Eichstaedt, J. C., & Willer, R. (2023). Incivility is rising among American politicians on Twitter. *Social Psychological and Personality Science, 14*(2), 259-269.

Gao, Y., Qin, W., Murali, A., Eckart, C., Zhou, X., Beel, J. D., ... & Yang, D. (2024, May). A Crisis of Civility? Modeling Incivility and Its Effects in Political Discourse Online. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 18, pp. 408-421).

Gimpel, J. G., & Hui, I. S. (2015). Seeking politically compatible neighbors? The role of neighborhood partisan composition in residential sorting. *Political Geography, 48*, 130-142.

Graham, M. H., & Svolik, M. W. (2020). Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *American Political Science Review, 114*(2), 392-409.

Holliday, D. E., Iyengar, S., Lelkes, Y., & Westwood, S. J. (2024). Uncommon and nonpartisan: Antidemo-

cratic attitudes in the American public. *Proceedings of the National Academy of Sciences, 121*(13), e2313013121.

Huber, G. A., & Malhotra, N. (2017). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics, 79*(1), 269-283.

Iyengar, S., Konitzer, T., & Tedin, K. (2018). The home as a political fortress: Family agreement in an era of polarization. *The Journal of Politics, 80(4)*, 1326-1338.

Iyengar, S., & Krupenkin, M. (2018). The strengthening of partisan affect. *Political Psychology, 39*, 201-218.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly, 76*(3), 405-431.

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. American *Journal of Political Science, 59*(3), 690-707

Jordan, D., Ollerenshaw, T., & Trexler, A. (2025). New Evidence and Design Considerations for Repeated Measure Experiments in Survey Research (No. q6czp_v1). Center for Open Science.

Kingzette, J., Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2021). How affective polarization undermines support for democratic norms. *Public Opinion Quarterly, 85*(2), 663-677.

Lee, A. H. Y. (2022). Social trust in polarized times: How perceptions of political polarization affect Americans' trust in each other. *Political Behavior, 44*(3), 1533-1554.

Levendusky, M. S., & Stecula, D. A. (2021). We need to talk: How cross-party dialogue reduces affective polarization. Cambridge University Press.

Mutz, D. C. (2006). Hearing the other side: Deliberative versus participatory democracy. Cambridge University Press.

Piazza, J. A. (2023). Political polarization and political violence. *Security Studies, 32*(3), 476-504.

Rainey, C. (2025). Power Rules: Practical Advice for Computing Power (and Automating with Pilot Data) (No. 5am9q_v2). Center for Open Science.

Rossiter, E. L., & Carlson, T. N. (2024). Cross-partisan conversation reduced affective polarization for republicans and democrats even after the contentious 2020 election. *The Journal of Politics, 86*(4), 1608-1612.

Sun, Q., Wojcieszak, M., & Davidson, S. (2021). Over-time trends in incivility on social media: evidence from political, non-political, and mixed sub-reddits over eleven years. *Frontiers in Political Science, 3*, 741605.

Appendix

Civil Condition

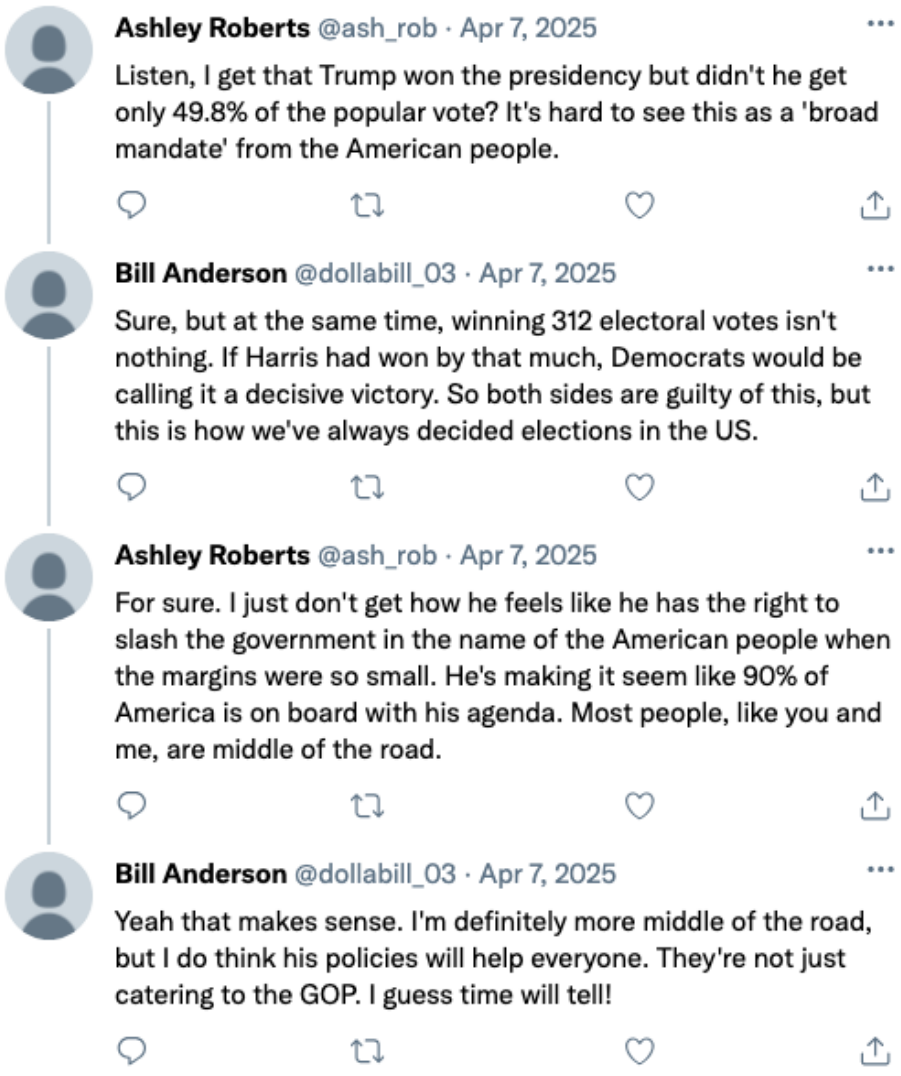

Figure 1: Civil Condition

Uncivil Condition

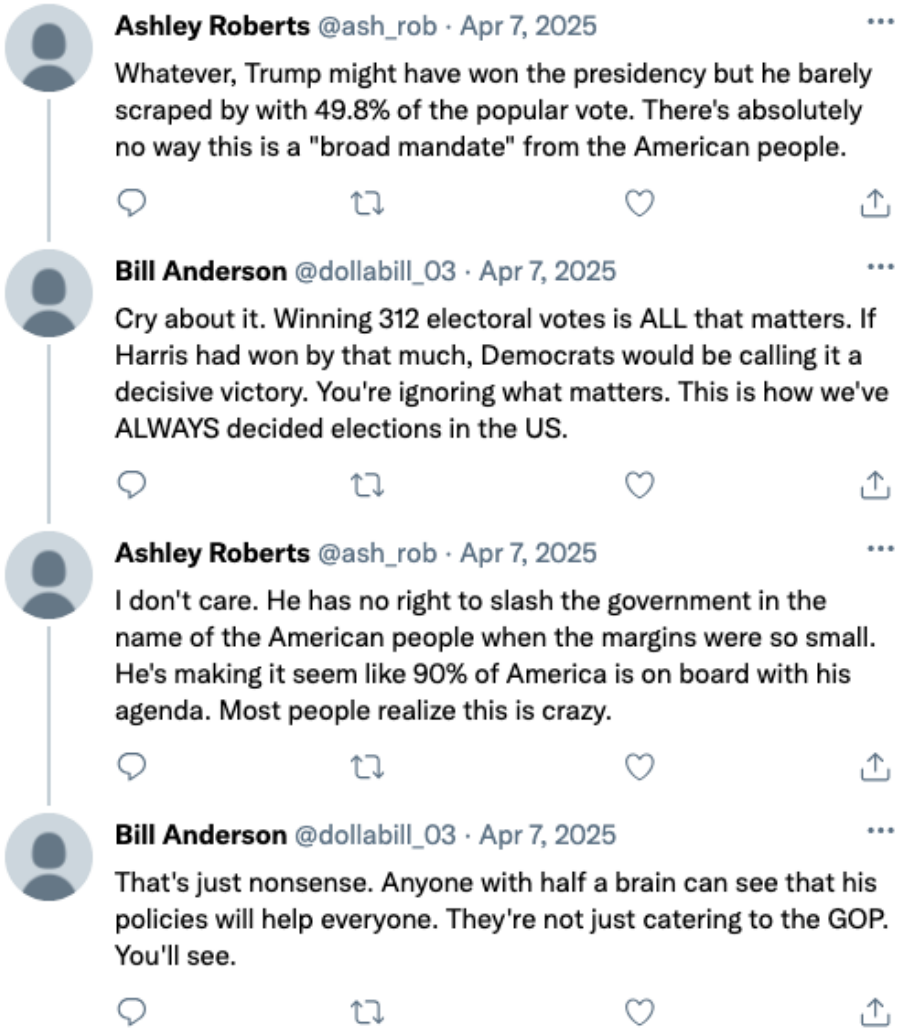

Figure 2: Uncivil Condition

Control Condition

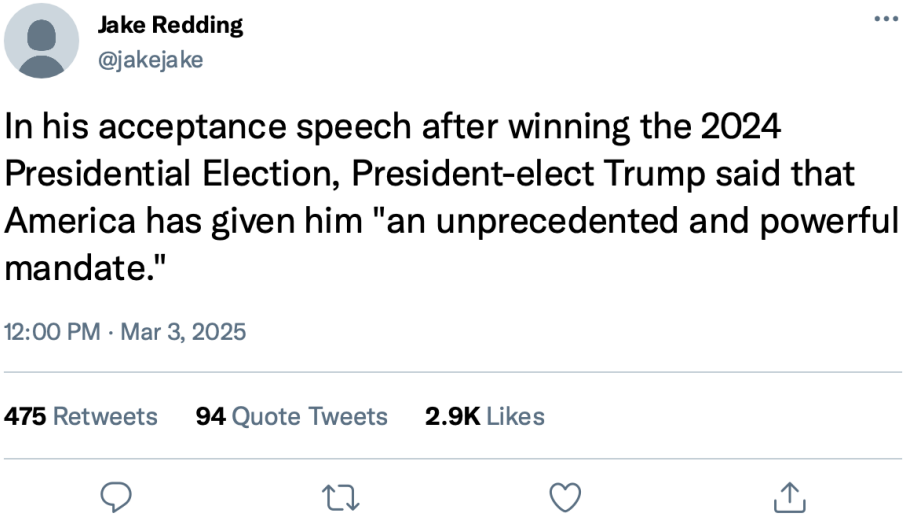

Figure 3: Control Condition (Also seen by treatment conditions)