

Polarization or Backlash? Voters' Reactions to Real Politicians' Tweets
Framing White Supremacy and Radical Islam as Terrorist Threats

Kirill Zhirkov and Rachel Smilan-Goldstein

University of Virginia

[Draft. Please don't cite or circulate]

Abstract

Progressive politicians in the United States increasingly discuss white supremacy as a form of terrorism. In recent decades, though, discussions of terrorist threat focused almost exclusively on radical Islam. We explore two potential reactions to this new rhetoric using two survey experiments. In both studies, respondents are exposed to real tweets of Democratic and Republican politicians discussing white supremacy and radical Islam, respectively, as security threats. One potential reaction is polarization: partisans diverging in their assessments of Republicans' anti-Muslim appeals in response to messages from Democratic politicians linking white supremacy to terrorism. The other is backlash: white respondents increasing their acceptance of anti-Muslim rhetoric after being exposed to tweets framing white supremacy as a threat. Overall, we find a backlash effect that is driven by Democrats and those high in group empathy. Our results highlight potential side effects of political communication about terrorism.

Keywords: backlash, group empathy, polarization, political communication, terrorism

The consequences of invoking terrorist threat are relatively well studied in political communication. Psychological consequences of exposure to news depicting terrorist events include heightened outgroup prejudice (Das et al., 2009), more punitive attitudes towards law-breakers (Fischer et al., 2007), and greater spread of fear (Makkonen et al., 2020). Politically, terrorist events can decrease support for civil liberties (Davis & Silver, 2004), change evaluations of political leaders (Albertson & Gadarian, 2016; Merolla & Zechmeister, 2013), and impact foreign policy opinions (Gadarian, 2010; Huddy et al., 2005; Malhotra & Popp, 2012). Overall, political actors may have strong incentives to strategically frame certain groups or events in terms of terrorist threat.

Indeed, rhetoric implicating terrorism and homeland security is not new in American politics. However, in recent history, it was directed almost exclusively against Muslims. Content analyses of U.S. news media have revealed consistent patterns in news coverage of terrorist events emphasizing Muslim identity of perpetrators (Bleich & Van der Veen, 2021; Lajevardi, 2021; Powell, 2011). And exposure to media representations of Muslims in the context of terrorism and violence increases negative stereotypes and Islamophobia (Saleem et al., 2017; Von Sikorski et al., 2021). Ultimately, stereotypes linking Muslims to violence have been mobilized to support policies such as the “war on terror” (Sides & Gross, 2013).

Whereas political appeals connecting Islam to terrorism have a history in American political communication, seeing extreme forms of white identity as a national security threat is a relatively new phenomenon in mainstream politics. In 2021, the Office of the Director of National Intelligence issued a report warning that threats from domestic violent extremists motivated by hatred of minority populations presented a major security threat.¹ These

¹ <https://www.dni.gov/files/ODNI/documents/assessments/UnclassSummaryofDVEAssessment-17MAR21.pdf>

assessments are now reflected in the rhetoric of prominent Democratic politicians. For instance, in his first presidential address to Congress, President Joe Biden described white supremacy as a domestic terror threat that the country could not ignore.² Similar statements can be found in the social media posts of other high-profile Democrats, including members of Congress.

In this paper, we explore voters' reactions to new rhetoric that frames white supremacy as a security threat. Our design takes advantage of the way these appeals mirror more common messages linking terrorism with Muslims. Specifically, we explore how exposure to messages about white supremacy as a terrorist threat (that come almost exclusively Democratic politicians) impacts perceived acceptability of more common appeals connecting terrorism to radical Islam (coming primarily from Republican politicians). Building upon existing literature, we propose two alternative hypotheses. One is partisan polarization: when Democrats and Republicans view partisan messaging on white supremacy as a security threat, they diverge even further in their assessments of anti-Muslim appeals. The other is white backlash: exposure to messages linking white supremacy to terrorism increases perceived acceptability of anti-Muslim appeals among whites. We also explore how messaging against white supremacy that emphasizes the security aspect may amplify the effects of group empathy.

We test these hypotheses in two original survey experiments, in which we ask respondents about perceived acceptability of anti-Muslim tweets by Republican politicians before or after exposing them to tweets by Democratic politicians that connect white supremacy to terrorism. We find strong evidence of backlash, which—surprisingly—is produced by Democrats and those high in group empathy. In other words, political messaging framing white supremacy as a security threat has the potential to demobilize those who otherwise are the first to

² <https://www.usatoday.com/story/news/politics/2021/04/28/biden-calls-white-supremacy-terrorism-speech-congress/4884034001/>

call out anti-Muslim appeals. We discuss the implications of our findings for political communication and American politics more generally.

Partisan Polarization

How does exposure to partisan messages linking certain groups or events to terrorist threat impact citizens' opinions? The default prediction from existing American public opinion literature is polarization. Citizens accept in-party cues by adjusting their political positions (for a review, see Leeper & Slothuus, 2014), and simultaneously reject cues coming from the out-party (Goren et al., 2009), resulting in opinion polarization.

The polarizing effects of political communication should be particularly high under the current levels of elite partisan polarization in U.S. politics (McCarty et al., 2006; Theriault, 2008). Indeed, elite-level polarization changes how citizens evaluate political issues by shifting the focus toward partisan endorsements and away from substantive information (Druckman et al., 2013). Given affective polarization in the mass public (Iyengar et al., 2012), which is largely driven by negative partisanship (Abramowitz & Webster, 2018; Bankert, 2021), partisans' propensity to respond to out-party messages by doubling-down on their pre-existing beliefs can be even stronger.

Therefore, the *polarization hypothesis* is that Democrats should consider an anti-Muslim message coming from a Republican politician less acceptable after seeing a message from a Democratic politician linking white supremacy to terrorism; meanwhile, there should be opposite effect among Republicans.

White Backlash

Partisanship, however, is only one group identity that powerfully shapes Americans' behaviors and opinions—another is race. The term “white backlash” has been used to describe

the process of white Americans adopting more conservative politics and starting to vote Republican in response to immigration, changing demographics, and related group status threats (Abrajano & Hajnal, 2015). Perceived group status threat also makes white Americans more likely to approve of explicit anti-minority appeals and view more favorably politicians making such appeals (Christiani, 2021). In other words, the politics of white Americans is increasingly influenced by their racial identity—and it pushes them to the right (Jardina, 2019).

Even though U.S. parties are increasingly sorted on the basis of race (Mason, 2018; Zhirkov & Valentino, 2022), whites still constitute a plurality of the Democratic coalition. And when white Democrats face cross-pressures with regard to their political identities, they often follow race over party. For instance, racial resentment makes a better predictor of opposition to affirmative action among white Democrats than it does among white Republicans (Feldman & Huddy, 2005). In addition, white Democrats report colder feelings to their own party after being exposed to Democratic political ads designed to court Latino voters (Ostfeld, 2019). Most recently, white identity has been found to be more strongly associated with opposition to immigration among white Democrats than among white Republicans (Perez et al., 2021).

Therefore, the *backlash hypothesis* is that whites, independently of their partisanship, should see an anti-Muslim message coming from a Republican politician as more acceptable after seeing a message from a Democratic politician linking white supremacy to terrorism. See Table 1 for a comparison between the polarization and backlash hypotheses.

Table 1
Summary of the Polarization and Backlash Hypotheses: Expected Effects of Exposure to a Message Linking White Supremacy to Terrorism on Acceptability of Anti-Muslim Appeals among White Partisans

Respondents	Effect on acceptability	
	Polarization	Backlash
White Democrats	–	+
White Republicans	+	+

Group Empathy

Polarization and backlash are not necessarily mutually exclusive. They may coexist as individuals react to political messages differently depending on their psychological predispositions. Here we focus on group empathy, an individual-difference variable that explains willingness to side with outgroups when they are attacked or treated unjustly (Sirin et al., 2021). According to past evidence, the group empathy mechanism can be strong enough to override prejudicial and exclusionary reactions even when the outgroups in question are seen as threatening (Sirin et al., 2016). Derogatory speech directed at minority groups provokes those high in group empathy to intervene in order to stop it (Sirin et al., 2017).

Therefore, the *group empathy hypothesis* postulates an interactive effect: there should be a negative relationship between group empathy and perceived acceptability of an anti-Muslim message from a Republican politician, and it should be stronger when it is seen after a message from a Democratic politician linking white supremacy to terrorism. In other words, exposure to political appeals invoking terrorist threat should reinforce the negative effect of group empathy on the perceived acceptability of anti-Muslim messages.

Study 1: Student Sample

Data and Method

To test our hypotheses, we carried out an original online survey-experimental study in March and April 2020. Our participants were political science major undergraduate students from a large public university in the southeastern United States. The survey was completed by 460 non-Hispanic white respondents.³ Male–female ratio was 41.7% to 58.3%. In terms of

³ The questionnaire was also completed by 235 non-white respondents, including Hispanics. In the main analysis, we focus on non-Hispanic whites due to the content of our hypotheses. However, we present a comparison of effects among white and non-white respondents later in the paper.

socioeconomic status, 27% of students described their families as upper class, 58.6% as middle class, 11.8% as working class, and 2.6% as lower class. Partisanship was 49.5% Democratic, 22.6% Republican, and 27.9% independent.

The survey experiment proceeded as follows. The questionnaire started with the group empathy battery. Respondents were presented with statements like: “I would describe myself as a pretty soft-hearted person towards people of other racial or ethnic groups,” and then asked how well these statements described them. Each respondent rated the total of four statements using a scale from 1 = *Does not describe me* to 5 = *Describes me extremely well*.⁴

After completing the group empathy battery, respondents were presented with screenshots of real tweets by Democratic and Republican members of Congress (see Figure 1 for sample tweets). Tweets by Democratic representatives addressed white supremacy as a national security threat whereas tweets by Republican representatives addressed radical Islam as a national security threat. Half of respondents saw a Republican politician’s tweet first (control condition) whereas the other half saw a Republican politician’s tweet after a Democratic politician’s tweet (treatment condition).⁵ Tweets presented to respondents were drawn randomly from a predefined selection of six tweets by Alexandria Ocasio-Cortez (Democratic representative) and six tweets by Louie Gohmert and Steve King (Republican representatives).⁶ This was done in order to minimize potentially idiosyncratic effects of each specific tweet. The selection of representatives was purposeful for the sake of external validity: these members of

⁴ See Supplementary Materials for the full group empathy battery.

⁵ In the control condition, respondents saw a Democratic tweet after a Republican tweet. We use the corresponding comparison (perceived appropriateness of a Democratic tweet on white supremacy before vs. after seeing a Republican tweet on radical Islam) to test whether any cross-exposure to out-party rhetoric, rather than anti-prejudice messaging specifically, can cause backlash.

⁶ When presented with tweets, respondents were informed of representatives’ partisanship in order to account for potential differences in political knowledge.

Congress were the ones who tweeted on the chosen issues (white supremacy for Democrats and radical Islam for Republicans) most often. Respondents were asked the following question: “Do you think that the language that the politician used in the tweet is appropriate for an elected official -- regardless of the point that they are trying to make?” Response scale was from 1 = *Completely inappropriate* to 7 = *Completely appropriate*.⁷

Figure 1

Tweet Examples



Alexandria Ocasio-Cortez
@AOC

White supremacy is a fundamentally violent ideology that's been growing at an alarming rate, fueling one mass shooting after another.

It's about time we recognized white nationalism for the major terrorist threat that it is.

(a) Democratic tweet



Louie Gohmert
@replouiegohmert

Could Someone Please Inform the Federal Judiciary That Jihadists Are Muslim? - bit.ly/2nvvlpF via @NRO

(b) Republican tweet

Results

Results are presented in Table 2 by respondent partisanship. Hereinafter, for the purpose of the analyses we treat respondents who are party-leaning independents as partisans. We find a significant positive effect of the treatment condition for Democratic respondents. When respondents see a Republican politician's tweet (on radical Islam as a security threat) after a Democratic one (on white supremacy as a security threat), they rate a Republican politician's tweet as more appropriate—compared with when the same Republican tweet is seen first. The effect is also relatively large in terms of magnitude: approximately 0.8 on a 7-point scale. These findings support the backlash hypothesis. Instead of polarizing respondents' opinions across party lines, exposure to messages linking white supremacy to terrorism leads Democrats to see rhetoric against Muslims from the Republican side as more acceptable. We find no significant

⁷ See Supplementary Material for the full list of tweets.

treatment effects for Republican respondents. The average level of perceived acceptability of anti-Muslim tweets is substantially higher among Republicans than among Democrats—but the means are still low enough to not be concerned about a potential ceiling effect.

Table 2

Mean Levels of Perceived Appropriateness of Anti-Muslim Tweets by Experimental Condition and Treatment Effects by Respondents' Partisanship, Study 1

	Control	Treatment	Difference	<i>N</i>
Democrats	2.60	3.42	0.82 [0.44, 1.21]	284
Republicans	4.54	4.05	-0.49 [-1.05, 0.08]	147

Note. 95% confidence intervals in brackets (difference estimates only).

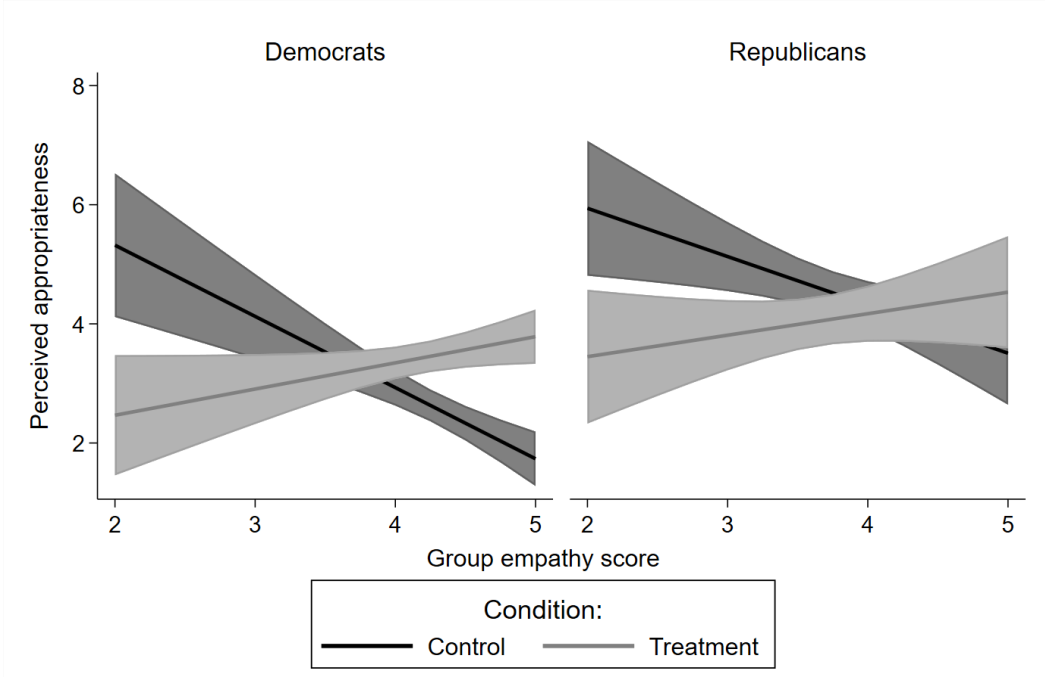
We implement a test to see whether the increase in acceptability of Republican messaging is indeed a reaction to messaging about white supremacy. Specifically, we estimate the potential mirroring effect by comparing perceived acceptability of Democratic tweets depending on whether they are seen before or after Republican ones. We find no significant effects (see Table S2 in Supplementary Material). In other words, the increase in acceptability is unique to the situation in which anti-Muslim messages are seen after ones linking white supremacy to terrorism.

Next, we move to explore the other hypothesized effect of messages linking white supremacy to terrorism: activation of group-empathic concerns. Specifically, we estimate OLS regression models that include interactive effects of empathy score and experimental treatment. Following the dominant guidelines (Brambor et al., 2005), we present results of the interactive model graphically in Figure 2.⁸ Recall that we have expected the discussion of white supremacy

⁸ Even though group empathy scores can potentially range from 1 to 5, there are no respondents with scores lower than 2. Therefore, we limit the graph's range to the empirically observed range (from 2 to 5).

as a security threat to reinforce the effect of group empathy. In this case, the negative effect of empathy score on perceived appropriateness of anti-Muslim rhetoric should have become even stronger compared to the control. However, the analysis returns an opposite result: the message connecting white supremacy to terrorism erases the effect of group empathy. The effect of empathy on perceived appropriateness of anti-Muslim tweets, which is negative as expected in the control condition, becomes insignificant in the treatment condition for both Democratic and Republican respondents.⁹

Figure 2
Relationship between Group Empathy and Perceived Appropriateness of Anti-Muslim Tweets by Experimental Condition and Respondents' Partisanship, Study 1



Note. Predictions with 95% confidence intervals. See Table S1 in Supplementary Material for coefficients.

⁹ Prediction lines for the anti-prejudice condition in Figure 2 may look like positive effects but the corresponding estimates are not significant on the 95% confidence level for both Democratic respondents ($p = .122$) and Republican respondents ($p = .319$).

Results of the interactive analysis have an important substantive interpretation. The overall increase in perceived appropriateness of anti-Muslim tweets is produced by those highest in group empathy. The apparent mechanism is demobilization: high-empathy respondents who otherwise would be the first to call out anti-Muslim rhetoric from Republican politicians are less likely to do so after being exposed to a message from a Democratic politician that talks about white supremacy in terms of security threat.

Discussion

The results of Study 1 have shown support for the backlash hypothesis. Exposure to messages linking white supremacy to terrorism does not polarize opinion between Democrats and Republicans. Instead, respondents who see the Democratic message against white supremacy first rate the Republican anti-Muslim message as more acceptable. We have also found an interactive effect between the treatment and group empathy. However, it works differently from what we have hypothesized: instead of reinforcing the negative effect of empathy on perceived appropriateness of anti-Muslim rhetoric, exposure to messaging about white supremacy as a security threat erases it.

The Study 1 sample limits the degree to which its findings can be generalized. First, it comes from college students who are known to differ from the general population on a number of important dimensions (Sears, 1986; but cf. Druckman & Kam, 2011). Second, our sample contains relatively few Republicans, making estimated effects for that group less reliable due to low statistical power. We aim to overcome these limitations in Study 2, which is effectively a replication of Study 1 on a general population sample.

Study 2: General Population Sample

Data and Method

We ran our replication study with a sample of U.S. adults in July 2020. Participants were recruited using the Lucid panel that was shown to approximate American national demographics (Coppock & McClellan, 2019). The survey was completed by 700 non-Hispanic white respondents.¹⁰ The sample characteristics were as follows. Mean age was 44 years old. Gender ratio was 48.4% male to 51.6% female. College education was reported by 40.1% of respondents. Median income was between \$40,000 and \$44,999. In terms of partisanship, 36.6% of respondents were Democrats, 38.5% were Republicans, and 24.9% were independents.

The experimental procedure replicated the one used in Study 1. Survey participants started by completing the group empathy battery and then were presented with the tweets. Similar to Study 1, half of respondents saw a Republican politician's tweet first (control condition) whereas the other half saw a Republican politician's tweet after that of a Democratic politician (treatment condition).¹¹ Tweets were rated in terms of perceived appropriateness.

A Democratic tweet and a Republican tweet presented to each respondent were, again, randomly chosen from predefined lists of real tweets (six for each party) by members of Congress. At the same time, the selection of tweets for Study 2 was different. In Study 1, all Democratic tweets were from Alexandria Ocasio-Cortez whereas Republican tweets were from Louie Gohmert and Steve King. However, this selection could raise concerns about respondents' reactions to the specific politicians rather than to the tweets' content. Therefore, for Study 2 we

¹⁰ The questionnaire was also completed by 276 non-white respondents, including Hispanics.

¹¹ Similar to Study 1, respondents saw a Democratic tweet after a Republican one in the control condition.

used tweets from four different Democratic members of Congress. The selection of Republican tweets also included tweets from more representatives.¹²

Results

We, again, start by documenting the treatment effects (difference between the treatment condition and the control condition) for perceived appropriateness of anti-Muslim tweets. Results are presented in Table 3 by respondent partisanship. Estimates show that the findings of Study 1 are exactly replicated: Messages linking white supremacy to terrorism impacted the perceived appropriateness of a Republican politician’s tweet for Democratic respondents alone. This effect is positive, with exposure to a Democratic tweet on the threat of white supremacy leading Democrats to rate a Republican tweet on the threat of radical Islam as more appropriate. We also confirm that, similar to Study 1, an increase in perceived acceptability is not found when a Democratic tweet is seen after a Republican one (if anything, Democrats rate tweets about white supremacy as less acceptable after seeing ones about radical Islam; see Table S4 in Supplementary Material).

Table 3
Mean Levels of Perceived Appropriateness of Anti-Muslim Tweets by Experimental Condition and Treatment Effects by Respondents’ Partisanship, Study 2

	Control	Treatment	Difference	<i>N</i>
Democrats	3.11	3.72	0.61 [0.10, 1.11]	231
Republicans	4.94	4.94	<0.01 [-0.41, 0.42]	299

Note. 95% confidence intervals in brackets (difference estimates only).

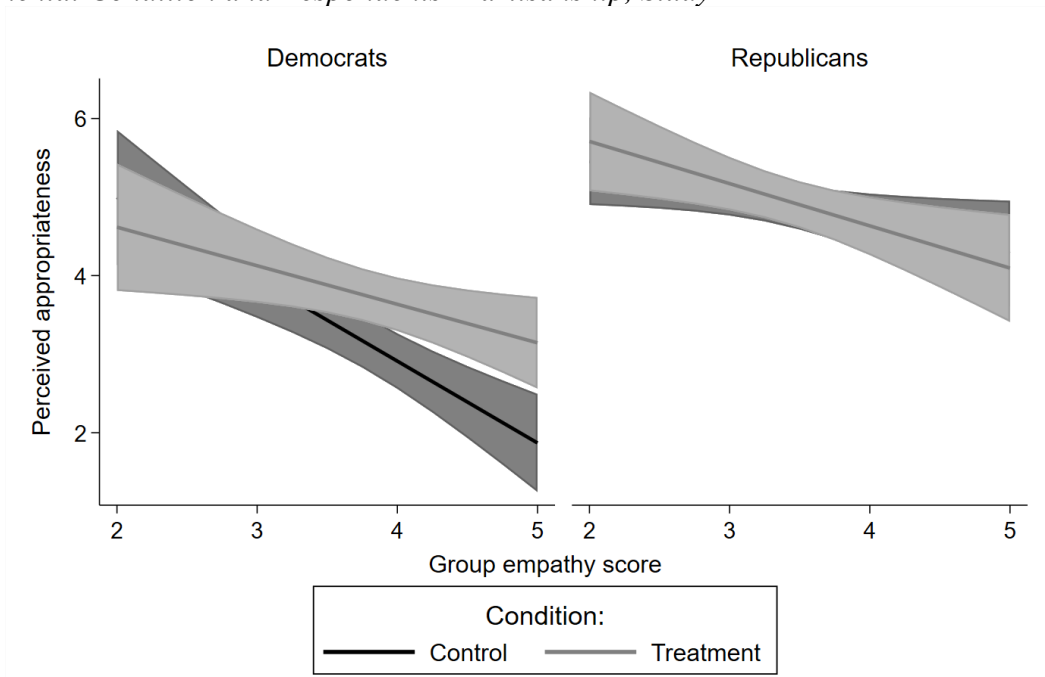
Next, we replicate the interactive analysis. Results are presented in Figure 3. They, again, replicate those reported in Study 1. The negative effect of group empathy on perceived

¹² See Supplementary Material for the full list of tweets.

appropriateness of anti-Muslim tweets from Republican representatives turns insignificant in the treatment condition. Importantly, the significant interaction is found only among Democrats. Among Republicans, group empathy does not show an effect. In other words, the increase in perceived appropriateness of Republican tweets about Islam or Muslims in the treatment condition is indeed produced by those high in group empathy. And this effect seems to be mostly produced by high-empathy Democrats.

Figure 3

Relationship between Group Empathy and Perceived Appropriateness of Anti-Muslim Tweets by Experimental Condition and Respondents' Partisanship, Study 2



Note. Predictions with 95% confidence intervals. See Table S3 in Supplementary Material for coefficients.

We also take advantage of Study 2 design—the fact that respondents were shown tweets from multiple Democratic representatives—to explore potential messenger effects. Specifically, we estimate the experimental effects independently across different tweets and then compare them across politicians. We find no differences between Alexandria Ocasio-Cortez and other

Democratic representatives ($F_{1,450} < 0.01, p = .972$), as well as no gender difference ($F_{1,450} = 0.62, p = .430$). Even though these results suggest no messenger effect, they should be treated as only suggestive due to relatively small sample sizes.

Discussion

The results of Study 2 have replicated the findings of Study 1. First, our experimental results show backlash rather than polarization: respondents who see the message linking white supremacy to terrorism first rate the anti-Muslim message as more acceptable. Second, this effect is largely produced by Democrats high in group empathy. These Democrats are less likely to call out anti-Muslim rhetoric from Republican representatives as inappropriate after being exposed to messages about white supremacy as a security threat.

How White is the Backlash?

Studies 1 and 2 have shown support for the backlash hypothesis: respondents see an anti-Muslim message coming from a Republican representative as more acceptable after seeing a message linking white supremacy to terrorism from a Democratic representative. Recall that the hypothesis has been formulated following the notion of “white backlash” and tested among non-Hispanic white respondents. It follows the literature on the growing importance of white identity and the related conservative shift among white Americans. However, the corresponding effect is found among white Democrats but not white Republicans—potentially complicating the backlash story. It is possible that, being cross-pressured by partisan and racial identity, white Democrats are particularly sensitive to messaging on race and religion.

To provide at least a preliminary evaluation of this conjecture, we implement a relatively simple set of analyses comparing the effects found in Studies 1 and 2 for white and non-white Democrats. Such comparisons are possible because, even though up to this point our analyses

have exclusively considered non-Hispanic whites, the questionnaires for both studies were completed by substantial numbers of non-white respondents.¹³

Table 4

Mean Levels of Perceived Appropriateness of Anti-Muslim Tweets by Experimental Condition and Treatment Effects among White and Non-white Democrats, Studies 1 and 2

	Control	Treatment	Difference	<i>N</i>
Study 1				
White Dems	2.60	3.42	0.82 [0.44, 1.21]	284
Non-white Dems	2.63	3.15	0.53 [0.02, 1.03]	182
Study 2				
White Dems	3.11	3.72	0.61 [0.10, 1.11]	231
Non-white Dems	3.75	3.93	0.19 [-0.47, 0.84]	126

Note. 95% confidence intervals in brackets (difference estimates only).

We start by comparing the treatment effects: impact of framing white supremacy in terms of security threat on perceived appropriateness of anti-Muslim tweets by Republican politicians. Results are presented in Table 4. We find a significant backlash effect for non-white Democrats in Study 1, although it is somewhat lower than that for white Democrats. In Study 2, the effect of the message linking white supremacy to terrorism on perceived appropriateness of anti-Muslim tweets is not significant (the point estimate is slightly positive). It is also necessary to note that the perceived acceptability of anti-Muslim appeals among non-white Democrats in Study 2 is relatively high even in the control condition. Since the main differences between the two samples are age and education (Study 1 participants are college students), this result potentially reveals

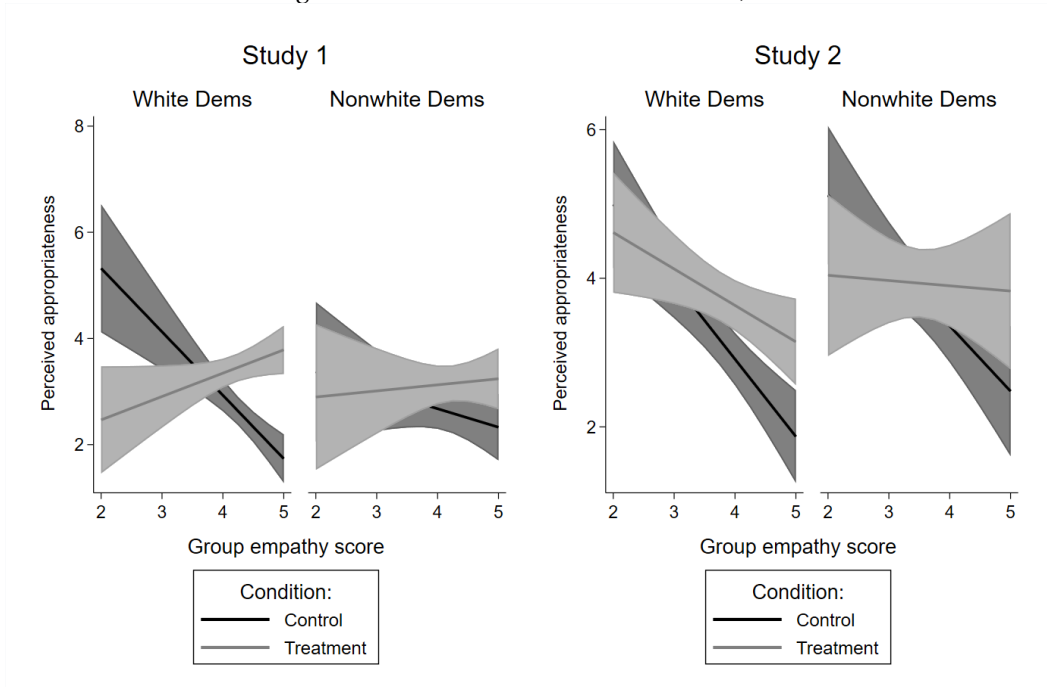
¹³ At the same time, the sample sizes do not allow us to implement more fine-grained analyses of respondents from various ethnic and racial groups collapsed under the “non-white” umbrella.

interesting generational differences among non-white Democrats—but this conjecture needs further testing.

Then, we turn to comparing the interactive effects of the message linking white supremacy to terrorism and group empathy for white and non-white Democrats. Results are presented in Figure 4. Overall, the pattern is the same: exposure to the “white supremacy as a security threat” frame almost reverses (in Study 1) or at least significantly decreases (in Study 2) the effect of group empathy on perceived acceptability of anti-Muslim tweets. This interactive effect for non-whites is somewhat smaller in Study 1 and slightly larger in Study 2. In other words, demobilization of those highest in group empathy is found among non-whites as well.

Figure 4

Relationship between Group Empathy and Perceived Appropriateness of Anti-Muslim Tweets by Experimental Condition among White and Non-white Democrats, Studies 1 and 2



Note. Predictions with 95% confidence intervals. See Table S5 in Supplementary Material for coefficients.

Although results from the two studies are mixed, they suggest that the backlash effect we find is not exclusively white. Non-white Democrats also react to messages that link white supremacy to terrorism by rating anti-Muslim appeals as more acceptable—but effect magnitudes are lower than those among whites. Similar to whites, the effect for non-whites is driven by those high in empathy.

Conclusion

In this paper, we have explored potential reactions to the emerging rhetoric in American political communication that frames white supremacy as a terrorist threat. We specifically focused on how it may impact the acceptability of more common appeals linking terrorism to radical Islam. We have outlined alternative hypotheses—partisan polarization and white backlash—and tested them in two survey-experimental studies, one using a student sample and one using a general population sample. Results of both studies show that exposure to messages from Democratic politicians addressing white supremacy as a security threat increases the perceived acceptability of appeals from Republican politicians linking Muslims to terrorism. In terms of the hypothesized mechanisms, we find evidence of backlash, which is driven by Democrats and those high in group empathy. Seeing a message on white supremacy as terrorism makes those high in group empathy less likely to reject anti-Muslim appeals. Meanwhile, we find no evidence supporting the polarization hypothesis.

The exact reasons for these findings are not fully clear. It is possible that white Democrats are particularly sensitive to messages connecting white supremacy to terrorism because of partisan and racial identity cross-pressures—whereas, for white Republicans, political implications of these identities are aligned. However, the analyses in which we compare white and non-white respondents suggest that the same kind of backlash can also occur among non-

white Democrats. Erasure of the group empathy effect on perceived acceptability of anti-Muslim rhetoric after exposure to tweets about white supremacy—replicated in both student and adult samples—is no less interesting. Discussing white supremacy in terms of domestic terrorism demobilizes respondents high in group empathy, making them less willing to call out anti-Muslim appeals. The mechanism behind this demobilization can be investigated in future studies.

It is also necessary to emphasize that the chosen research design should have made our test favorable to the polarization hypothesis and conservative for the backlash hypothesis. First, white supremacy is an extreme form of white identity that is rejected by white Democrats (Jardina et al., 2020), making backlash against this particular message more difficult to find. Second, cues coming from specific politicians tend to be more polarizing than ones simply associated with party labels (Nicholson, 2012). Thus, it should have been particularly difficult for Democrats to find anti-Muslim tweets from named Republican politicians acceptable, especially ones coming from such controversial figures as Steve King. Still, we find backlash among Democrats—and no polarization.

An important aspect of our experiment, which has implications for interpretation of the findings, is the nature of the stimuli: real-world tweets from Democratic and Republican members of Congress. This is a common design in studies on social media messaging (Ballard et al., 2022; Bowler et al., 2022), that increases external validity by presenting respondents with the same stimuli they can be exposed to in real life. At the same time, this design involves constraints on randomization of certain attributes. For instance, in the list that we have selected, the Democratic tweets are consistently longer than the Republican ones. Similarly, tweets occasionally conflate different issues: for instance, radical Islam with immigration (for Republicans) and white supremacy with gun violence (for Democrats). In addition, our design

makes it difficult to disentangle the message and the messenger effects, especially ones of race and gender (Bonilla et al., 2022). Results from Study 2 alleviate these concerns, but only partially. The relationship between the message content and the messenger identity can be an interesting topic for future studies.

Our findings have important implications for American politics more broadly. Scholars of political communication have recently noticed the growing acceptability of explicit group-based appeals in U.S. campaigns (Valentino et al., 2018). Our survey-experimental results suggest that this phenomenon may be at least partly explained by backlash against progressive messaging on race and religion. Even exposure to messages mentioning white supremacy, which is almost universally condemned across the political spectrum, makes voters more accepting of rhetorical attacks against minority groups. The fact that this backlash is produced exclusively by Democrats indicates that partisan sorting on the basis of race and racial attitudes is not complete. Despite racial conservatives abandoning the Democratic Party (Valentino & Sears, 2005), while those who remain turning more racially liberal (Sides et al., 2018), there are still Democrats who may react negatively to strongly worded progressive messages. Future studies should explore whether such backlash effects extend beyond the case of discussing white supremacy as a form of domestic terrorism.

References

- Abrajano, M., & Hajnal, Z. L. (2015). *White backlash: Immigration, race, and American politics*. Princeton University Press.
- Abramowitz, A. I., & Webster, S. W. (2018). Negative partisanship: Why Americans dislike parties but behave like rabid partisans. *Advances in Political Psychology, 39*, 119–135. <https://doi.org/10.1111/pops.12479>
- Albertson, B., & Gadarian, S. K. (2016). Anxiety over terrorism advantages Hillary Clinton. *Political Communication, 33*(4), 681–685. <https://doi.org/10.1080/10584609.2016.1224419>
- Bankert, A. (2021). Negative and positive partisanship in the 2016 U.S. presidential elections. *Political Behavior, 43*, 1467–1485. <https://doi.org/10.1007/s11109-020-09599-1>
- Ballard, A., DeTamble, R., Dorsey, S., Heseltine, M., & Johnson, M. (2022). Incivility in congressional tweets. *American Politics Research, 50*(6), 769–780. <https://doi.org/10.1177/1532673X221109516>
- Bleich, E., & Van der Veen, A. M. (2021). Media portrayals of Muslims: A comparative sentiment analysis of American newspapers, 1996–2015. *Politics, Groups, and Identities, 9*(1), 20–39. <https://doi.org/10.1080/21565503.2018.1531770>
- Bonilla, T., Filindra, A., & Lajevardi, N. (2022). How source cues shape evaluations of group-based derogatory political messages. *Journal of Politics, 84*(4), 1979–1996. <https://doi.org/10.1086/720308>
- Bowler, S., Carreras, M., & Merolla, J. L. (2022). Trump tweets and democratic attitudes: Evidence from a survey experiment. *Political Research Quarterly*. Advance online publication. <https://doi.org/10.1177/10659129221137348>

- Brambor, T., Clark, W. R., & Matt Golder. 2006. Understanding interaction models: Improving empirical analyses. *Political Analysis*, 14(1), 63–82. <https://doi.org/10.1093/pan/mpi014>
- Christiani, L. (2021). When are explicit racial appeals accepted? Examining the role of racial status threat. *Political Behavior*. Advance online publication. <https://doi.org/10.1007/s11109-021-09688-9>
- Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research and Politics*, 6(1). <https://doi.org/10.1177/2053168018822174>
- Das, E., Bushman, B. J., Bezemer, M. D., Kerkhof, P., & Vermeulen, I. E. (2009). How terrorism news reports increase prejudice against outgroups: A terror management account. *Journal of Experimental Social Psychology*, 45(3), 453–459. <https://doi.org/10.1016/j.jesp.2008.12.001>
- Davis, D. W., & Silver, B. D. (2004). Civil liberties vs. security: Public opinion in the context of the terrorist attacks on America. *American Journal of Political Science*, 48(1), 28–46. <https://doi.org/10.1111/j.0092-5853.2004.00054.x>
- Druckman, J. N., & Kam, C. D. (2011). Students as experimental participants: A defense of the “narrow data base.” In J. N. Druckman, D. P. Greene, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge Handbook of Experimental Political Science* (pp. 41–57). Cambridge University Press. <https://doi.org/10.1017/CBO9780511921452.004>
- Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, 107(1), 57–79. <https://doi.org/10.1017/S0003055412000500>

- Feldman, S., & Huddy, L. (2005). Racial resentment and white opposition to race-conscious programs: Principles or prejudice? *American Journal of Political Science*, 49(1), 168–183. <https://doi.org/10.1111/j.0092-5853.2005.00117.x>
- Fischer, P., Greitemeyer, T., Kastenmueller, A., Frey, D., & Osswald, S. (2007). Terror salience and punishment: Does terror salience induce threat to social order? *Journal of Experimental Social Psychology*, 43(6), 964–971. <https://doi.org/10.1016/j.jesp.2006.10.004>
- Gadarian, S. K. (2010). The politics of threat: How terrorism news shapes foreign policy attitudes. *Journal of Politics*, 72(2), 469–483. <https://doi.org/10.1017/S0022381609990910>
- Goren, P., Federico, C. M., & Kittilson, M. C. (2009). Source cues, partisan identities, and political value expression. *American Journal of Political Science*, 53(4), 805–820. <https://doi.org/10.1111/j.1540-5907.2009.00402.x>
- Huddy, L., Feldman, S., Taber, C., & Lahav, G. (2005). Threat, anxiety, and support of antiterrorism policies. *American Journal of Political Science*, 49(3), 593–608. <https://doi.org/10.1111/j.1540-5907.2005.00144.x>
- Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, 76(3), 405–431. <https://doi.org/10.1093/poq/nfs038>
- Jardina, A. (2019). *White identity politics*. Cambridge University Press. <https://doi.org/10.1017/9781108645157>

- Jardina, A., Kalmoe, N., & Gross, K. (2020). Disavowing white identity: How social disgust can change social identities. *Political Psychology*. Advance online publication. <https://doi.org/10.1111/pops.12717>
- Lajevardi, N. (2021). The media matters: Muslim American portrayals and the effects on mass attitudes. *Journal of Politics*, 83(3), 1060–1079. <https://doi.org/10.1086/711300>
- Leeper, T. J., & Slothuus, R. (2014). Political parties, motivated reasoning, and public opinion formation. *Advances in Political Psychology*, 35, 129–156. <https://doi.org/10.1111/pops.12164>
- Makkonen, A., Oksanen, A., Gadarian, S. K., Herreros, F., Winsvold, M. S., Solheim, O. B., Enjolras, B., & Steen-Johnsen, K. (2020). Fear-triggering effects of terrorism threats: Cross-country comparison in a terrorism news scenario experiment. *Personality and Individual Differences*, 161, 109992. <https://doi.org/10.1016/j.paid.2020.109992>
- Malhotra, N., & Popp, E. (2012). Bridging partisan divisions over antiterrorism policies: The role of threat perceptions. *Political Research Quarterly*, 65(1), 34–47. <https://doi.org/10.1177/1065912910385251>
- Mason, L. (2018). *Uncivil agreement: How politics became our identity*. University of Chicago Press.
- McCarty, N., Poole, K. T., & Rosenthal, H. (2006). *Polarized America: The dance of ideology and unequal riches*. MIT Press.
- Merolla, J. L., & Zechmeister, E. J. (2013). Evaluating political leaders in times of terror and economic threat: The conditioning influence of politician partisanship. *Journal of Politics*, 75(3), 599–612. <https://doi.org/10.1017/S002238161300039X>

- Nicholson, S. P. (2012). Polarizing cues. *American Journal of Political Science*, 56(1), 52–66.
<https://doi.org/10.1111/j.1540-5907.2011.00541.x>
- Ostfeld, M. C. (2019). The new white flight? The effects of political appeals to Latinos on white Democrats. *Political Behavior*, 41, 561–582. <https://doi.org/10.1007/s11109-018-9462-8>
- Perez, E. O., Kuo, E. E., Russel, J., Scott-Curtis, W., Mucoz, J., & Tobias, M. (2021). The politics in white identity: Testing a racialized partisan hypothesis. *Political Psychology*. Advance online publication. <https://doi.org/10.1111/pops.12788>
- Powell, K. A. (2011). Framing Islam: An analysis of U.S. media coverage of terrorism since 9/11. *Communication Studies*, 62(1), 90–112.
<https://doi.org/10.1080/10510974.2011.533599>
- Saleem, M., Prot, S., Anderson, C. A., & Lemieux, A. F. (2017). Exposure to Muslims in media and support for public policies harming Muslims. *Communication Research*, 44(6), 841–869. <https://doi.org/10.1177/0093650215619214>
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3), 515–530. <https://doi.org/10.1037/0022-3514.51.3.515>
- Sides, J., & Gross, K. (2013). Stereotypes of Muslims and support for the war on terror. *Journal of Politics*, 75(3), 583–598.
- Sides, J., Tesler, M., & Vavreck, L. (2018). *Identity crisis: The 2016 presidential campaign and the battle for the meaning of America*. Princeton University Press.
- Sirin, C. V., Valentino, N. A., & Villalobos, J. D. (2016). Group empathy theory: The effect of group empathy on US intergroup attitudes and behavior in the context of immigration threats. *Journal of Politics*, 78(3), 893–908. <https://doi.org/10.1086/685735>

- Sirin, C. V., Valentino, N. A., & Villalobos, J. D. (2017). The social causes and political consequences of group empathy. *Political Psychology, 38*(3), 427–448.
<https://doi.org/10.1111/pops.12352>
- Sirin, C. V., Valentino, N. A., & Villalobos, J. D. 2021. *Seeing us in them: Social divisions and the politics of group empathy*. Cambridge University Press.
<https://doi.org/10.1017/9781108863254>
- Therriault, S. M. (2008). *Party polarization in Congress*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511790652>
- Valentino, N. A., Neuner, F. G., Vandebroek, L. M. (2018). The changing norms of racial political rhetoric and the end of racial priming. *Journal of Politics, 80*(3), 757–771.
<https://doi.org/10.1086/694845>
- Valentino, N. A., Sears, D. O. (2005). Old times there are not forgotten: Race and partisan realignment in the contemporary South. *American Journal of Political Science, 49*(3), 672–688. <https://doi.org/10.1111/j.1540-5907.2005.00136.x>
- Von Sikorski, C., Matthes, J., & Schmuck, D. (2021). The Islamic State in the news: Journalistic differentiation of Islamist terrorism from Islam, terror news proximity, and Islamophobic attitudes. *Communication Research, 48*(2), 203–232.
<https://doi.org/10.1177/0093650218803276>
- Zhirkov, K., & Valentino, N. A. (2022). The origins and consequences of racialized schemas about U.S. parties. *Journal of Race, Ethnicity, and Politics, 7*(3), 484–504.
<https://doi.org/10.1017/rep.2022.4>

Supplementary Material

Group Empathy

“We have a couple of questions about your thoughts and feelings in a variety of situations. For each statement, indicate well it describes you.”

1. When I’m upset at someone from another racial or ethnic group, I usually try to “put myself in their shoes” for a while.
2. Before criticizing somebody from another racial or ethnic group, I try to imagine how I would feel if I were in their place.
3. I sometimes find it difficult to see things from the “other person’s” point of view, particularly someone from another race or ethnicity.
4. If I’m sure I’m right about something, I don’t waste much time listening to the arguments of people, particularly those of other racial or ethnic groups.
5. I often have tender, concerned feelings for people from another racial or ethnic group who are less fortunate than me.
6. I would describe myself as a pretty soft-hearted person towards people of other racial or ethnic groups.
7. The misfortunes of people from other racial or ethnic groups do not usually disturb me a great deal.
8. Sometimes I don’t feel very sorry for people of other racial or ethnic groups when they are having problems.

Respondents were randomly presented with four statements, one from each following pair: (1) and (2), (3) and (4), (5) and (6), (7) and (8). Statement order randomized. Answers given on a 5-point Likert-type scale from 1 = *Does not describe me* to 7 = *Describes me extremely well*

Democratic tweets, Study 1



Alexandria Ocasio-Cortez
@AOC

White supremacy has quickly turned into a domestic terror crisis in the US.

They rely on you thinking "it's not a big deal." It is a big deal.

White supremacy now makes up the majority of domestic terrorism in the United States. They radicalize online.



Alexandria Ocasio-Cortez
@AOC

& again, this time asking the FBI.

It seems easier to label domestic, self-radicalized shooters who are Muslim as terrorists.

Yet white supremacists, who are responsible for far more violence in the US, often escape these charges.

Now they're the #1 source of domestic terror ↓



Alexandria Ocasio-Cortez
@AOC

Video games aren't causing mass shootings, white supremacy is.

Sadly the GOP refuse to acknowledge that, bc their strategy relies on rallying a white supremacist base.

That's why the President hosts stadiums of people chanting "send her back" & targets Congress-members of color.



Alexandria Ocasio-Cortez
@AOC

White supremacists were responsible for *ALL* race-based domestic terrorism in 2018. 100%.

Trump's DoJ & Barr then worked **to hide that report from Congress**--all while defunding federal programs to combat white supremacist violence.

This is a white supremacist administration.



Alexandria Ocasio-Cortez
@AOC

When a mass shooting is committed by a Muslim, the crime is almost automatically labeled as Domestic Terrorism.

But when White Supremacists like Dylann Roof shot up a black church, or attack the Tree of Life Synagogue, the FBI declined to charge them w/ Domestic Terrorism.

Why?



Alexandria Ocasio-Cortez
@AOC

White supremacy is a fundamentally violent ideology that's been growing at an alarming rate, fueling one mass shooting after another.

It's about time we recognized white nationalism for the major terrorist threat that it is.

Republican tweets, Study 1



Louie Gohmert
@replouiegohmert

Could Someone Please Inform the Federal Judiciary That Jihadists Are Muslim? - bit.ly/2nvvlpF via @NRO



Steve King
@SteveKingIA

The #Muslim Invasion of Europe: a clear concise description of transformation of Western European Culture by hijrah.



Steve King
@SteveKingIA

Our hearts and prayers go out for the French victims of yet another terrorist attack. We must defeat the ideology of Islamic Jihad.



Louie Gohmert
@replouiegohmert

Coming soon: More immigrants from Muslim nations than population of D.C. -- 680,000 washex.am/1QlwUB5 via @DCExaminer



Steve King
@SteveKingIA

Sweden: What America is becoming. 12 terrorist bombings in 24 days. Sweden, highest per capita immigration in world. pgj.cc/dKgZJP



Steve King
@SteveKingIA

Are shop manuals for Boeings written in Arabic? A 30 year Arabic\only mechanic, whose brother is ISIS, sabotages a plane & NO terrorist consequence? "I do admit the guilt," Abdul-Majeed Marouf Ahmed Alani, 60, said through an Arabic interpreter...

Democratic tweets, Study 2



Alexandria Ocasio-Cortez
@AOC

& again, this time asking the FBI.

It seems easier to label domestic, self-radicalized shooters who are Muslim as terrorists.

Yet white supremacists, who are responsible for far more violence in the US, often escape these charges.

Now they're the #1 source of domestic terror 📉



Alexandria Ocasio-Cortez
@AOC

White supremacy is a fundamentally violent ideology that's been growing at an alarming rate, fueling one mass shooting after another.

It's about time we recognized white nationalism for the major terrorist threat that it is.



Hakeem Jeffries
@RepJeffries

Mother Emanuel AME Church. Tree of Life Synagogue. Cielo Vista Massacre.

White Supremacy = Domestic Terrorism.

It must be crushed.

Not coddled.

There are not very fine people on both sides.



Alexandria Ocasio-Cortez
@AOC

When a mass shooting is committed by a Muslim, the crime is almost automatically labeled as Domestic Terrorism.

But when White Supremacists like Dylann Roof shot up a black church, or attack the Tree of Life Synagogue, the FBI declined to charge them w/ Domestic Terrorism.

Why?



Congresswoman Bass
@RepKarenBass

Today the House Judiciary Committee held a hearing on the rise of hate crimes and white nationalism in this country.

This country DOES have a problem with domestic terrorism. We DO have a problem with White Supremacy. Congress should work to address it.



Rep. Don Beyer
@RepDonBeyer

Check out my latest, in [@businessinsider](#):

"America has a serious white-supremacy problem, which is helping fuel a rise in hate crimes, violence, and domestic terrorism, and it's time for Congress to do something to combat this growing threat to our nation."

Republican tweets, Study 2



Our hearts and prayers go out for the French victims of yet another terrorist attack. We must defeat the ideology of Islamic Jihad.



The [#Muslim](#) Invasion of Europe: a clear concise description of transformation of Western European Culture by hijrah.



Islamic terrorism must be stopped, and we must make changes to defend our nation & destroy terrorism at the root. majorityleader.gov/2016/06/13/the...



Are shop manuals for Boeings written in Arabic? A 30 year Arabic\only mechanic, whose brother is ISIS, sabotages a plane & NO terrorist consequence? "I do admit the guilt," Abdul-Majeed Marouf Ahmed Alani, 60, said through an Arabic interpreter...



Could Someone Please Inform the Federal Judiciary That Jihadists Are Muslim? - bit.ly/2nvwlpF via [@NRO](#)



Koran: "We shall certainly strike terror into the hearts of those who have disbelieved" (3:151) & "Verily, the disbelievers are your open enemies." (4:101).

On 9/11/01, fundamentalist Muslims attacked America, killing 2,996.

We must remember & learn. History must not repeat. Mo

Table S1. Interactive model coefficients for Democrats and Republicans, Study 1

	Democrats	Republicans
Anti-prejudice condition	-6.13*** (1.48)	-4.84** (1.67)
Group empathy	-1.20*** (0.26)	-0.81* (0.31)
Anti-prejudice × Empathy	1.64*** (0.35)	1.17** (0.45)
<i>N</i>	284	147

Note. Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S2. Mean levels of perceived appropriateness of Democratic tweets by condition and estimated differences by respondents' partisanship, Study 1

	Seen first	Seen second	Difference	<i>N</i>
Democrats	5.68	5.65	-0.03 [-0.34, 0.29]	284
Republicans	3.46	3.83	0.37 [-0.19, 0.93]	147

Note. 95% confidence intervals in brackets (difference estimates only)

Table S3. Interactive model coefficients for Democrats and Republicans, Study 2

	Democrats	Republicans
Anti-prejudice condition	-1.48 (1.18)	0.55 (0.94)
Group empathy	-1.04*** (0.22)	-0.39* (0.19)
Anti-prejudice × Empathy	0.55 (0.30)	-0.15 (0.27)
<i>N</i>	231	299

Note. Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S4. Mean levels of perceived appropriateness of Democratic tweets by condition and estimated differences by respondents' partisanship, Study 2

	Seen first	Seen second	Difference	<i>N</i>
Democrats	5.34	4.79	-0.53 [-1.01, -0.06]	231
Republicans	3.92	3.99	0.07 [-0.43, 0.57]	299

Note. 95% confidence intervals in brackets (difference estimates only)

Table S5. Interactive model coefficients for white and nonwhite Democrats, Studies 1 and 2

	Study 1		Study 2	
	White Dems	Nonwhite Dems	White Dems	Nonwhite Dems
Anti-prejudice condition	-6.13*** (1.48)	-1.40 (1.80)	-1.48 (1.18)	-2.70 (1.52)
Group empathy	-1.20*** (0.26)	-0.35 (0.30)	-1.04*** (0.22)	-0.88** (0.27)
Anti-prejudice × Empathy	1.64*** (0.35)	0.46 (0.43)	0.55 (0.30)	0.81 (0.42)
<i>N</i>	284	182	231	126

Note. Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$