Does Violent Protest Backfire?  
Testing a Theory of Public Reactions to Activist Violence

Brent Simpson¹, Robb Willer², and Matthew Feinberg³

Abstract
How do people respond to violent political protest? The authors present a theory proposing that the use of violence leads the general public to view a protest group as less reasonable, a perception that reduces identification with the group. This reduced identification in turn reduces public support for the violent group. Furthermore, the authors argue that violence also leads to more support for groups that are perceived as opposing the violent group. The authors test this theory using a large (n = 800) Internet-based survey experiment with a politically diverse sample. Participants responded to an experimental scenario based on recent violent confrontations between white nationalist protesters and antiracist counter-protesters, allowing the authors to study whether violent protest would reduce public support even when used against a widely reviled group. The authors found that the use of violence by an antiracist group against white nationalists led to decreased support for the antiracist group and increased support for the white nationalist group. Furthermore, the results were consistent with the theorized causal process: violence led to perceptions of unreasonableness, which reduced identification with and support for the protest group. Importantly, the results revealed a striking asymmetry: although acts of violence eroded support for an antiracist group, support for white nationalist groups was not reduced by the use of violence, perhaps because the public already perceives these groups as very unreasonable and identifies with them at low levels. Consistent with this interpretation, the authors found that self-identified Republicans, a subset of the sample that reported less extremely negative views of white nationalists, showed reduced support for white nationalists when they engaged in violence.

Keywords
violence, protest, social psychology, public perception

Over two days in August 2017, a group of white nationalists and white supremacists, many carrying automatic weapons, marched in Charlottesville, Virginia, in a “Unite the Right” rally protesting the removal of a confederate monument. They were met by a group of mostly peaceful antiracist counter-protesters, one of whom died when she was hit by a white nationalist who drove his car through a crowd of the antiracists. A few weeks later, a large group of antiracist counter-protesters met a small group of far-right protesters, including many white supremacists, in Berkeley, California. The events turned violent when a small group of the antiracist counter-protesters (AR counter-protesters), many associated with the Antifa movement, attacked the right-wing protesters.

These and other recent incidents have rekindled long-standing societal debates about the role of violence in public protest and civic resistance.¹ Although much of this discussion has centered on the righteousness of using violence as a tool of political advocacy (are violent protest actions morally acceptable?), debate has also swirled around the pragmatics of violent protest (are violent protest actions effective?). Here we focus on the latter debate, investigating a specific question regarding the practical

¹For example, while giving an interview following Donald Trump’s presidential inauguration in January 2017, Richard Spencer, an influential white supremacist and leader of the “alt-right,” was punched in the head by a masked man, an incident that led to wide discussion of whether it was pragmatic and/or morally acceptable to “punch a Nazi.”
effects of violent protest: how does the use of violence affect public perception of protest groups?

A large literature addresses the efficacy of different approaches to civil disobedience and public protest (Elsbach and Sutton 1992; Gamson 1990; Jenkins and Eckert 1986; Stephan and Chenoweth 2008; Tarrow 2011; Thomas and Louis 2014; Wang and Piazza 2016), much of it focused on the effectiveness of violence versus nonviolence in garnering attention and support for a given group or cause. Here we contribute to this literature by presenting and testing a new theory of popular responses to violent protest. The theory proposes that the use of violence leads protest groups to be perceived as less reasonable by the general public, a perception that reduces the public’s identification with and support for the groups. Furthermore, our theory proposes that violence typically increases support for groups perceived to be in opposition to the violent groups.

Importantly, violent protests may have other types of effects on public opinion that might make them more or less practical (see Haines 1984; Wang and Piazza 2016). For example, these tactics may attract, or deter, activists to join a given group or movement. Violence could also be effective in deterring the mobilization of rival activist groups. Violent protest could also conceivably be effective in promoting awareness of a protest group’s main positions. In the present research we focus on the immediate public opinion effects of violent protest and consider these other possible effects below in the general discussion.

We test the claim that the use of violence reduces support for protest groups, and the intervening process we propose, in a large \( n = 800 \) Internet-based survey experiment including a politically diverse sample. Our experiment is based on recent violent confrontations between white nationalist protesters (WN protesters) and AR counter-protesters, allowing us to ask whether violent protest undermines public support even when used against a widely reviled group. This test case allows a demonstration of the extremity of the negative effects of violent protest in that we find both lower public support for the group that uses violence and more positive views of the opposing group. Importantly, our results also show a striking asymmetry in the effects of violence: although public perceptions of antiracist groups are eroded by violence, support for white nationalist groups is impacted much less when they engage in violence. We argue that this is because identification with white nationalist groups is already so low that violence does minimal damage to their already negative reputations. Below, we briefly review prior work that informs our arguments and then present our theory and experiment.

**The Effects of Violent Protests on Public Support**

Existing research generally finds that violence by protesters reduces public support for the protesters and can even erode support for the causes they support. Drawing on a large data set of more than 300 resistance campaigns against political regimes and foreign occupiers, Stephan and Chenoweth (2008) found that nonviolent movements were more effective at winning domestic and international support and at winning loyalty shifts within a regime’s security forces. Other studies have shown that public support for a movement wanes when it uses violent tactics, and that violence damages the perceived legitimacy of groups (Wang and Piazza 2016). For instance, one recent analysis found that in the 1960s, regions featuring nonviolent protests by black civil rights activists saw increased Democratic voting by whites, whereas regions where violent civil rights protests occurred saw decreased Democratic vote share among whites (Wasow 2017). Indeed, Wasow’s analysis suggests that backlash to violent civil rights protests could have been sufficient to tip the 1968 election in Richard Nixon’s favor.

That said, prior work tends to use data from naturally occurring situations. It is difficult to know in such data whether violence versus nonviolence alters public perceptions of protesters, whether the use of violence versus nonviolence is a strategic response by protesters to differences in perceived public support, or whether organizations that use violence lose popularity for other reasons. Indeed, Wang and Piazza (2016) found that activists tend to avoid the use of violence when their claims or goals are supported by a broad swath of the public. Thus, we cannot tell from prior work whether, or to what extent, the use of violence by activists is a cause or consequence of relatively negative public perceptions. Here we extend recent work that uses carefully controlled experiments to assess the impact of extreme protest tactics on public support for protesters. For example, Feinberg, Willer, and Kovacheff (2017) found that protest behaviors that are disruptive, harmful, and/or highly counter-normative tend to reduce support for protest groups and the causes they advocate for. Another set of experiments (Thomas and Louis 2014) found that reports of highly disruptive protest activities—reports that included reference to violent acts, such as hurling projectiles—diminished support for protest groups compared with reports of peaceful protest activities.

Although past research suggests that violence will reduce public support for groups, it remains unclear why violence has these negative effects. We outline a theory asserting that violence leads protests groups to be perceived as less reasonable, identified with less, and ultimately supported less. By extension, the theory asserts that violence will tend to increase support for groups perceived to be opposed to ones that use violence.

Also unclear from prior work is whether there might be boundaries to the negative effects of violence on public opinion. Specifically, research has not examined whether the negative effects of violence on public perceptions might obtain when the violence is directed at groups that are viewed very negatively by most of the public. After all, the literature on morality finds that individuals more easily rationalize, and engage in, dehumanization and aggression against those
viewed as moral outcasts (Bandura et al. 1996; Bar-Tal 1990; Opotow 1990). From this perspective, we might expect that those who use violence against reviled groups will garner public support. Thus, if we observe negative effects of violence even when used against a widely despised group that is itself known for violence, it would be particularly strong evidence of the detrimental effects of violence on public support for protest groups.

**Theory and Hypotheses**

*Violence Decreases Public Support*

On the basis of insights from prior work, we argue that protesters who use violence will be viewed as less reasonable than those using nonviolent protest methods. Given that people tend to view themselves as reasonable and rational (Pronin, Gilovich, and Ross 2004) and that similarity is a key basis of identification (Byrne et al. 1971), we expect that observers will be less inclined to identify with protesters or activists they view as less reasonable. This is critical because identification with protests and protesters is central to movement support (Polletta and Jasper 2001; van Zomeren, Postmes, and Spears 2008). This argument is consistent with Feinberg et al.’s (2017) recent demonstration that protesters using more extreme tactics were identified with less and supported less than protesters engaged in more normative, less disruptive, forms of action.

The following propositions reflect the foregoing theoretical reasoning regarding how the use of violence by protest groups shapes public perceptions of them:

**Proposition 1**: People perceive protest groups that engage in violence to be less reasonable than protest groups perceived to be nonviolent.

**Proposition 2**: People identify less with protest groups they perceive to be unreasonable.

**Proposition 3**: The less people identify with a protest group, the less they will support that group.

We derive from these propositions the following specific hypotheses regarding the effects of protest group violence on public perceptions of the group.

**Hypothesis 1**: People will support a protest group less if the group engages in violence.

Second, the propositions suggest that the negative effects of violence by a group on public perceptions of it will be driven by perceptions that the group is unreasonable and by lower levels of identification with it.

**Hypotheses 2**: The negative effect of a protest group’s violence on popular support for the group will be mediated by (1) perceptions of the group’s reasonableness and (2) identification with the group.

*Violence Increases Public Support for Opposing Groups*

Groups are often viewed as making competing claims about values or goals. To the extent the public views two groups’ goals as competing, or in opposition to each other, support for the two groups will tend to be zero-sum, such that lower levels of support for one group imply higher levels of support for the competing group. Thus, we further propose that violence not only affects perception of the group using it but can also have downstream effects on perceptions of rival groups, or salient outgroups perceived as directly opposing the violent group.

**Proposition 4**: The less people support a given protest group, the more they will show support for rival protest groups.

Application of the propositions implies that a group that uses violence will be viewed as less reasonable (proposition 1), identified with less (proposition 2), and supported less (proposition 3). In turn, these lower levels of support for the violent group will lead to higher levels of support for rival advocacy groups (proposition 4).

**Hypothesis 3**: Use of violence by a protest group will lead to higher levels of public support for rival protest groups.

**Theoretical Scope**

We restrict the scope of these propositions and hypotheses to situations in which a decrease in public opinion toward the protest group is possible. For instance, if a group is already supported at very low levels by the public, violence is unlikely to further erode support for the group. Similarly, if a group is viewed as extremely unreasonable, violence is unlikely to further increase the public’s perception that the group is unreasonable. This implies an important scope condition on application of the foregoing propositions:

**Scope Condition**: The theory applies when decreases in public support and more negative perceptions of a group are possible.

This scope restriction implies that hypotheses 1 and 2 will apply only when the group using violence is not already very negatively perceived. For the same reasons, we should expect the process leading up to hypothesis 3 to apply only when public perceptions of the violent group are not already so negative and fixed that they cannot be reduced further, as explained in greater detail later. Whether this scope condition is satisfied can be assessed via measures of support for the group in a baseline condition in which it does not engage in violence.
Empirical Strategy

In the next section we introduce the experimental test of our hypotheses. As noted earlier, we test whether this theory and the derived hypotheses are valid even when violence is used by antiracist groups against a broadly reviled political group, specifically white nationalists. Besides being a setting with contemporary relevance, it is also a theoretically important one for several reasons. Despite increased prominence of white nationalists since the candidacy and election of Donald Trump (Swenson 2017), they remain a widely despised group that likely already elicits strong associations with violence. If there are any conditions under which violence might increase support for a protest group, this is likely one of them.

Second, this case allows us to more carefully assess the scope conditions of our arguments, namely, whether our hypotheses hold when antiracist groups are violent but not when white nationalist groups are. Specifically, even when not engaged in violence, white nationalists should be viewed as unreasonable, and most members of the public would be loath to state that they identify with or support them, as we demonstrate below. Thus, whereas we expect that antiracist protest groups using violence will lose public support, because white nationalists are a widely reviled group, we do not expect that the use of violence will further reduce public support for them, though we investigate the possibility in the following experiment. For the same reasons, whereas we expect that violence by antiracists will increase support for white nationalist groups (hypothesis 3), we do not expect that violence by white nationalist groups will increase support for antiracist groups. Thus, this case allows us to test an important asymmetry in the effects of violence by groups that have relatively broad support versus those that have much narrower support.

Although our main goal is to test our theory, we also measure a number of exploratory variables, as explained in the section to follow.

Method

Eight hundred participants (50.4 percent female) were recruited from a population of approximately 70,000 users of Amazon Mechanical Turk who had previously completed a short prescreen survey. This allowed us to sample relatively equal proportions of respondents across the political spectrum (28.7 percent identified mainly as Republicans, 39.1 percent as independents, and 32.3 percent as Democrats) and ethnic categories based roughly on their numbers in the general population (60.4 percent white, 20.0 percent Hispanic or Latino, and 13.9 percent black or African American).

Upon completing the consent process, participants read one of four newspaper articles, depending on condition. Although the newspaper articles were based on recent protests and counter-protests, including those that occurred in Charlottesville and Berkeley in August 2017, they were (unbeknownst to participants) fabricated to maximize experimental control. The full text of each is given in Appendix B.

The experiment featured a 2 (WN protesters violent or not) × 2 (AR counter-protesters violent or not) design. Participants in the “no violence,” or control, condition read about a group of white nationalists who staged a protest of the removal of confederate monuments. In the story, AR counter-protesters showed up at the event to stage a counter-demonstration. The newspaper report made clear that there was no physical violence on either side.

In the other three conditions, the article depicted either the WN protesters alone as violent, the AR counter-protesters alone as violent, or both sides as violent. For instance, in the condition in which AR counter-protesters were violent, the report read that “one white nationalist was knocked to the ground by counter-protesters carrying homemade shields,” that “a counter-protester was seen punching and kicking a white nationalist who lay on the ground shielding his face from the blows,” and that “the anti-racism counter-protesters . . . pepper-sprayed white nationalists.” Apart from whether or which side used violence, the newspaper reports were very similar.

After reading the newspaper articles, participants answered questions about their perceptions of the WN protesters and the antiracist counter-protestors. Most relevant to our hypotheses, the participants indicated on separate 7-point scales whether they supported the WN protesters and AR counter-protesters (1 = “not at all,” 7 = “a great deal”).3

We also measured identification with the WN protesters (WN identification) and AR counter-protesters (AR identification), as well as the extent to which the participants perceived the WN protesters and AR counter-protesters to be reasonable. To measure identification, participants indicated on 7-point scales (from 1 = “not at all” to 7 = “a great deal”) the extent to which they “felt similar to,” “identified with,” and “related to” the white nationalists (α = .96) and antiracists (α = .96) (Luhtanen and Crocker 1992). For the

---

3Prior work has shown that Amazon Mechanical Turk samples yield reliable, high-quality data (Berinsky, Huber, and Lenz 2012; Buhrmester, Kwang, and Gosling 2011; Paolacci and Chandler 2014; Weinberg, Freese, and McElhattan 2014).
measure of perceived reasonableness, participants indicated how “reasonable” and “senseless” (reverse coded) the white nationalist ($\alpha = .84$) and antiracist demonstrators ($\alpha = .88$) were. These latter items were measured using a 7-point scale that ranged from 1 (“strongly disagree”) to 7 (“strongly agree”).

In addition to these primary measures that allow tests of our key hypotheses, we included several more exploratory measures. First, we measured the extent to which participants perceived support for the WN protesters (WN support) and AR counter-protesters (AR support) to be normative. Specifically, each participant was asked to estimate the percentage of Americans he or she believed would support both the WN protesters and AR counter-protesters. We also included items designed to tap whether violence on either side influenced the perceived importance of fighting for racial justice more generally; the extent to which participants felt that whites, blacks, and Latinos were victims of discrimination in the United States; President Trump’s job approval ratings; perceptions of white nationalist groups in general; and whether participants felt that “both white nationalists and the anti-racism counter-protesters deserve blame for recent problems in the United States.” After these measures, participants completed a sociodemographic questionnaire and were paid for their participation. Average time to complete the study was 9 minutes, 27 seconds.

Results

The manipulations had the intended effects, with participants in the WN protester violence (WN violence) conditions reporting higher levels of WN violence ($M = 5.74$ on a scale ranging from 1 to 7) than participants in conditions in which WN protesters were not violent ($M = 2.99$, $t = 22.36$, $p < .001$). Likewise, participants in the AR counter-protester violence (AR violence) conditions reported that the AR counter-protesters were more violent ($M = 5.51$) than did participants in conditions in which the AR counter-protesters were not violent ($M = 2.50$, $t = 26.54$, $p < .001$).

Table 1 gives means, standard deviations, and analysis of variance (ANOVA) results for our primary dependent variable and hypothesized mediating variables. Although our experimental design is $2 \times 2$, our primary focus will be on main effects of our experimental manipulations, because they provide tests of our three hypotheses. Below we first review the main effects of AR violence on the dependent variables. We then turn to other analyses, including possible effects of WN violence and how WN violence might moderate the effects of AR violence.

Before turning to these analyses, we use the perceptions of WN protesters and AR counter-protesters in the control condition to assess whether each group falls within our theoretical scope. As shown in Table 1, AR counter-protesters are supported ($M = 5.09$ on a scale ranging from 1 to 7), viewed as reasonable ($M = 5.42$), and identified with ($M = 4.72$) at relatively high levels, thus satisfying our core scope condition. Our arguments therefore suggest that AR violence will result in less AR support (hypothesis 1) because they will be viewed as less reasonable and identified with less (hypothesis 2) and that AR violence will further lead to greater WN support (hypothesis 3). On the other hand, baseline support is very low for WN protesters ($M = 1.77$), as is the average level of identification with this group ($M = 1.64$). The only exception is a more reasonable view of WN protesters (WN reasonable), which is closer to the midpoint of the scale ($M = 3.64$). The very low levels of WR support and WR identification at the baseline imply that they lie outside the scope of our arguments. As a consequence, we do not expect WN support to be further eroded when they use violence, nor do we expect WN violence to lead to greater AR support.

Table 1. Effect of WN Violence and AR Violence on Perceptions of Each Group.

<table>
<thead>
<tr>
<th>Mean (SD) by Condition</th>
<th>p Value (ANOVA)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Violence</td>
</tr>
<tr>
<td>AR counter-protester support</td>
<td>5.09 (2.01)</td>
</tr>
<tr>
<td>AR counter-protester reasonable</td>
<td>5.42 (1.47)</td>
</tr>
<tr>
<td>AR counter-protester identification</td>
<td>4.72 (1.93)</td>
</tr>
<tr>
<td>WN protester support</td>
<td>1.77 (1.41)</td>
</tr>
<tr>
<td>WN protester reasonable</td>
<td>3.64 (1.73)</td>
</tr>
<tr>
<td>WN protester identification</td>
<td>1.64 (1.26)</td>
</tr>
</tbody>
</table>

Note: ANOVA = analysis of variance; AR = antiracist; WN = white nationalist.
Results from analyses of our other dependent measure, belief that AR violence is justified (AR justified), are identical (see Appendix A.)

Hypothesis 2 predicts that the effects of violence on AR support will be driven by a tendency for the public to view violent counter-protesters as less reasonable and to identify with them less than if they were not violent. As shown in Table 1, we found the predicted effects of AR violence on perceptions of AR counter-protesters as reasonable (AR reasonable) and also on levels of identification with them. Bootstrap mediation analyses (Preacher and Hayes 2008) showed that AR reasonable fully mediated the effects of AR violence on AR support (confidence interval [CI] [.79, -.92]; Figure 1a). A parallel analysis found that AR identification also partially mediated the effects of AR violence on AR support (CI [.89, .97]; Figure 1b). We therefore conducted a serial mediation analysis using Hayes’s (2012) PROCESS Model 6 to test the causal chain predicted by hypothesis 2, namely, that AR violence would lead first to perceptions of AR counter-protesters as less reasonable, next to diminished reported identification with them (CI [-1.22, -.87]), this result supports hypothesis 2, which predicts that violent AR counter-protesters will be viewed as less reasonable and thus identified with less, which will lead to lower levels of support for them. In addition to the predicted path, Figure 1c shows that zero was not in the 95 percent CI (CI [-.58, -.31]) for only one other path, the more direct path via AR reasonable. No other paths in the model were significant. Again, results from our other dependent measure, AR justified, yield identical results (see Appendix A).

**Figure 1.** (a) AR reasonable fully mediates the effect AR violence on AR support. (b) AR identification partially mediates the effect of AR violence on AR support. (c) Serial mediation model for AR support.

Note: AR reasonable = perceptions of antiracist counter-protesters as reasonable; AR support = support for antiracist counter-protesters; AR violence = violence by antiracist counter-protesters.

As shown in Figure 1c, these analyses supported the hypothesized serial mediation. Specifically, these results showed that the predicted causal impact of AR violence on AR support operates through a tendency to perceive the AR counter-protesters as less reasonable, leading to diminished reported identification with them (CI [-1.22, -.87]). This result supports hypothesis 2, which predicts that violent AR counter-protesters will be viewed as less reasonable and thus identified with less, which will lead to lower levels of support for them. In addition to the predicted path, Figure 1c shows that zero was not in the 95 percent CI (CI [-.58, -.31]) for only one other path, the more direct path via AR reasonable. No other paths in the model were significant. Again, results from our other dependent measure, AR justified, yield identical results (see Appendix A).

Does Violence by Antiracist Counter-protesters Increase Support for WN Protesters?

We now turn to hypothesis 3, which predicts that AR violence will have spillover effects on views of a salient rival protest
group, here increasing WN support. As shown in Table 1, we found that the effect of AR violence on WN support falls just shy of a conventional significance level ($p = .056$), providing suggestive evidence for hypothesis 3. (As shown in Appendix A, results from our other dependent measure, belief that WN violence is justified [WN justified], yield similar but clearer support for the hypothesis.) We explored whether this effect of AR violence on WN support is driven by decreased perceptions of AR counter-protesters as reasonable, decreased identification with them, or decreased support for them. Although the total effect of AR violence on WN support was marginally significant ($p = .056$), we conducted separate bootstrap mediation models, finding that this effect was fully mediated by AR reasonable (CI $[-.26, -.14]$; Figure 2a), lower AR identification (CI $[-.21, -.11]$; Figure 2b), and lower AR support (CI $[-.21, -.11]$; Figure 2c). We therefore tested a serial mediation path that included all three of these potential mediators. Only for the AR reasonable path was zero not in the 95 percent confidence interval (CI $[.03, .37]$), suggesting that violence by one group increased support for an opposing group via reduced tendency to view the violence as reasonable. No other paths in the model were significant.

**Does Violence by a Widely Reviled Group Further Decrease Public Support for Its Members?**

Because WN protesters are negatively perceived even in the baseline condition, we do not necessarily expect violence to

---

**Figure 2.** (a) AR reasonable fully mediates the effect of AR violence on WN support. (b) AR identification fully mediates the effect of AR violence on WN support. (c) AR reasonable fully mediates the effect of AR violence on WN support.

Note: AR identification = identification with antiracist counter-protesters; AR reasonable = perceptions of antiracist counter-protesters as reasonable; AR violence = violence by antiracist counter-protesters; WN support = support for white nationalist protesters.

*p = .058, **p = .058.*
further influence perceptions of them (hypotheses 1 and 2) or to alter perceptions of AR counter-protesters (hypothesis 3). These analyses should therefore be considered exploratory.

First, we find that WN violence led to lower levels of WN support ($p = .013$) and to perceptions that WN protesters were less reasonable ($p < .001$). We cannot test whether these effects are mediated by the process identified in hypothesis 2, because WN violence did not alter the already very low levels of WN identification ($p = .108$). We can, however, assess the logic underlying the hypothesis for WN violence by limiting our analysis to a subset of participants with sufficiently high baseline perceptions of WN protesters. Specifically, we analyzed results for participants who identified as Republican on a prescreen measure of party identification. Among this subset of our sample, perceptions of WN protesters in the control condition varied from moderately low (WN identification = 2.4 and WN support = 2.6 on 7-point scales) to moderate (WN reasonable = 4.4). Thus, they allow a within-scope test of the mediation process detailed in hypothesis 2.

Among this subset of participants, we found that WN violence led to lower levels of WN support ($p = .02$), consistent with hypothesis 1. Furthermore, we found that WN violence decreased WN reasonable ($p < .001$) and led to lower levels of WN identification ($p = .03$). We therefore assessed whether the impact of WN violence on WN support among these participants was mediated by changes in perceptions of WN protesters as reasonable and identification with them.

Bootstrap mediation analyses showed that WN reasonable fully mediated the effects of WN violence on WN support (CI [.25, .50]; Figure 3a). WN identification also fully mediated the effects of WN violence on WN support (CI [.87, 1.00]; Figure 3b). The results of the serial mediation, as given in Figure 3c, are highly consistent with hypothesis 2. The results show that the impact of WN violence on WN support was fully mediated by decreased perceptions of WN reasonable on WN identification, which in turn reduced WN support (CI [.−.78, −.30]). No other paths in the mediation model were significant. Thus, a focused analysis of participants who identify more strongly as Republicans found

---

**Figure 3.** (a) Among Republicans, WN reasonable fully mediates the effect of WN violence on WN support. (b) Among Republicans, WN reasonable fully mediates the effect of WN violence on WN support. (c) Serial mediation model for Republicans’ WN support.

Note: WN reasonable = perceptions of white nationalist protesters as reasonable; WN support = support for white nationalist protesters; WN violence = violence by white nationalist protesters.

*p < .05. ***p < .001.

---

4Furthermore, participants who identified as Republican perceived WN protesters in the control condition to be more reasonable and reported higher levels of WN identification and WN support ($p \leq .01$ for all) than all other participants.
clear evidence of the full mediation path predicted by hypothesis 2, even for WN violence.\textsuperscript{5}

**Does Violence by White Nationalist Groups Increase Support for Antiracist Groups?**

Earlier, we found that AR violence had carryover effects on perceptions of a rival protest group, leading to increased WN support (following hypothesis 3). As noted earlier, the low levels of WN support at the baseline imply that hypothesis 3 does not apply to violence by white nationalists. In exploratory analyses, we did not find evidence that WN violence led to increased AR support, as shown by the absence of a significant main effect of WN violence on AR support ($p = .539$) in Table 1.\textsuperscript{6} This finding stands in stark contrast to results for hypothesis 3 presented earlier, where AR violence led to more favorable perceptions of WN protesters. Here we saw that perceptions of AR counter-protesters if anything, reduced by violence perpetrated by WN protesters. We return to this asymmetry in the section on the consequences of violence in the discussion.

**Does Violence by White Nationalists Moderate the Effects of Antiracist Violence?**

We did not predict that violence by one group would moderate the effects of violence by the opposing protest group, though our design allows us to explore this question. As shown in Table 1, we did not observe any significant interaction effects for perceptions of WN protesters. However, we did find interaction effects for all three of our measures of perceptions of AR counter-protesters, suggesting that the effect of AR violence on perceptions of the AR counter-protester group depended to some extent on whether WN protesters were also violent.

---

\textsuperscript{5}On the other hand, among participants who identified as Democrats, WN violence influenced only WN reasonable ($p < .001$). It did not alter WN identification ($p = .58$) or WN support ($p = .65$). When coupled with the results for Republican participants, these analyses provide evidence for the appropriateness of our scope condition.

\textsuperscript{6}Indeed, the ANOVA results given in Table 1 show significant interaction effects on perceptions of AR counter-protesters, which suggests that, if anything, these counter-protesters were perceived more negatively when the WN protesters were violent, even when the AR counter-protesters were not themselves violent. To assess this, we compared the control condition (no violence by either group) with the condition in which WN protesters were violent but AR counter-protesters were not. These analyses show that nonviolent AR counter-protesters were viewed as significantly less reasonable ($p = .009$) when WN protesters were violent compared with when WN protesters were not violent. Although they do not reach statistical significance, the patterns for AR identification ($p = .106$) and AR support ($p = .055$) are in the same direction.

Simple comparisons between the “both sides violent” condition and the condition in which only the AR counter-protesters were violent showed that AR counter-protesters were viewed as more reasonable ($p = .017$) when both sides were violent. Although the patterns were in the same direction for the AR identification and AR support measures, differences between conditions were only marginally significant ($p = .08$ for both). Thus, although we found some evidence that the effects of AR violence depended on whether WN protesters were also violent, the main effects of AR violence mattered much more, as indicated in the analyses outlined above, as well as a comparison of means in Table 1.

**Results for Exploratory Measures**

As noted earlier, we also assessed a number of exploratory measures designed to assess whether our manipulations influenced participants' tendencies to view WN support and AR support as normative, as well as whether violence on either side influenced participants' positions on broader, race-related issues. We discuss these only briefly here; detailed results are available upon request.

First, echoing our key results, we found that AR violence led participants to believe fewer Americans would support the AR counter-protestors ($p < .001$) and that more Americans would support the WN protesters ($p < .005$). Although there was a marginal tendency for participants to assume that more Americans would support the AR counter-protesters when the WN protesters were violent ($p = .055$), WN violence did not significantly influence participants' estimates of the number of Americans who supported WN protesters ($p = .17$). Thus, as with our main results, the data on normative perceptions shows that AR violence mattered more for perceptions of both groups than did WN violence.

We also measured how our manipulations influenced perceptions that both white nationalists and AR counterparters “deserve blame for recent problems in the United States.” This is important in the context of President Trump’s highly criticized statement following the violence in Charlottesville, which left one antiracist demonstrator dead, that “there is blame on both sides.” We were interested in whether participants would be more likely to blame “both sides” for more general problems in the United States following exposure to AR violence. Although WN violence had no impact on participants’ views of both sides as blameworthy ($p = .90$), participants who were exposed to AR violence were substantially more likely to agree that both sides are to blame for recent problems in the United States ($p < .001$).

We did not find any evidence that our manipulations influenced participants’ general perceptions that whites, blacks, or Latinos are victims of discrimination in the United States. Nor did we find any impact of our independent
variables on a composite of two items designed to measure perceptions of racial justice (“It is critical that we fight for racial justice in this country” and “The struggle for racial equality is of utmost importance”), reported support for white nationalist groups in general, or perceptions that Donald Trump is doing a good job as president.

**Discussion**

The results of our experiment provide consistent support for hypotheses derived from our theory. In line with hypothesis 1, we found that AR violence undermined AR support. Although our scope condition suggested that hypothesis 1 should not apply to WN protesters, we found that WN violence also reduced WN support. These results suggest that the effect of violence by protesters on loss of support may be more general than we expected.

We also found consistent support for hypothesis 2, which details the intervening mechanisms through which AR violence erodes support for AR counter-protesters. Specifically, we found that violence led participants to view AR counter-protesters as less reasonable, which in turn reduced identification with and, ultimately, support for them. Following the scope condition, we did not expect to find evidence for hypothesis 2 for WN violence, given the low levels of identification with them in our general sample. However, as a check on the logic of our argument, we analyzed responses from a subsample of our participants who did fall within the scope of our theory, namely, those identifying as Republicans, who reported moderate levels of WN support. Among these participants, we found evidence of the serial mediation path predicted by hypothesis 2.

Although the first two hypotheses concern perceptions of groups that use violence, hypothesis 3 deals with how violence by one group spills over, influencing support for an opposing group. One of our most striking findings was that although AR violence led to increased WN support (following hypothesis 3), WN violence did not increase AR support, as expected. Why does this asymmetry occur? One possibility is suggested by our finding, described earlier, that AR violence increased WN support because it led to more negative perceptions of AR counter-protesters. But perceptions of WN protesters were already quite negative at the baseline and thus could not be harmed as much by violence. A related possibility is that AR violence likely runs counter to stereotypical expectations. If so, it would likely be more salient and lead to stronger reactions. By comparison, WN protesters already tend to bring to mind associations with violence, and knowledge that they are violent is less likely to have downstream effects. Whatever the specific reasons for this asymmetry, it has important practical implications, as discussed below.

It is important to note that in our study, the damaging effects of either WN violence or AR violence were limited primarily to perceptions of the protesters and counter-protesters. We did not find evidence that the negative effects of violence extended to the causes associated with the antiracists or white nationalists. This is likely because all these issues are related in some way to race, and positions on race-related issues are among the most strongly held in the United States (Tesler 2016). Thus, it is not surprising that our relatively minimal stimulus of a protest event was insufficient to affect these often deeply held issue positions.

One exception is that AR violence, but not WN violence, increased participants’ tendencies to assign blame to “both sides” for recent problems in the country. This is significant given Trump’s claim that “many sides” deserved blame for the violence in Charlottesville, where a white nationalist killed an antiracist demonstrator. It is not clear whether Trump’s claims, made three to seven weeks prior to our data collection, acted as a “frame” that influenced participants’ responses on this measure. Future research might investigate the effects of violence in the presence of reminders of different frames by political and movement leaders on each side.

**Limitations and Questions for Future Research**

The scenario we used for our empirical test was selected to be externally valid and of significant contemporary interest. However, this choice incurred methodological trade-offs in terms of internal validity and generalizability of our results. For instance, consistent with recent events where white nationalists’ demonstrations (Adone, Simon, and Sidner 2018) or protests of the planned removal of confederate monuments prompted counter-protests by antiracist groups (Spencer 2017), our experiment depicted white nationalists and antiracists in the roles of “protesters” and “counter-protesters,” respectively. But it may be that the label “counter-protester” has generally negative connotations, leading to participants’ more negative reactions to counter-protester violence in our study, and perhaps we would have observed different results had antiracists instead been in the role of protesters. If so, this would limit the generalizability of our research.

Two pieces of evidence suggest this is not an issue for our research. First, we find high levels of support for (nonviolent) AR counter-protesters in the control condition.
(M = 5.09 on a 7-point scale), suggesting that participants were not particularly put off by the counter-protester label. Second, our analyses of Republican respondents specifically showed they responded to WN violence as predicted by the theory and parallel to how the entire sample responded to violence by AR counter-protesters. That said, one can also imagine a setting in which roles were reversed (e.g., a protest by antiracists is disrupted by a group of white nationalist counter-protesters). For instance, in August 2018, protesters tore down a confederate memorial on the University of North Carolina campus, although there were no reports of white nationalists present at the event (Deconto and Blinder 2018). Future work should address the robustness of our results, in which no group is labeled counter-protesters or the protester and counter-protester roles are reversed.

Also in line with events such as the 2017 Charlottesville protests, the newspaper articles participants read contained an asymmetry between the relatively innocuous chants of antiracists (“Fight racism!”) and the highly charged statements made by white nationalists (e.g., “Jews will not replace us!”). Although future research should more explicitly explore whether this asymmetry matters, the fact that we observed similar effects for white nationalist violence in the subsample of Republicans and antiracist violence in the entire sample suggests that our results were driven by violence, rather than an asymmetry in the chants used by white nationalists and antiracists in the newspaper articles.

Relatedly, to construct a conservative test of the effects of violent protest by antiracists, we sought to portray AR violence as an instrumental effort to end the white nationalists’ protest, rather than retaliatory aggression. To signal this motivation, we included the chants of AR counter-protesters (“Nazis go home!” and “Fight racism!”) in those conditions in which AR counter-protesters were violent. We did not include these specific chants in the nonviolent AR counter-protester conditions, as we were concerned that doing so might create the impression that there was also violence in those conditions. But these choices also introduced a confound, in that it may have been the specific chants, rather than physical violence, that led to the effects we observed.

We can address this confound first by assessing whether participants’ perceptions of AR violence, measured via the manipulation check on violence, mediated the effects of the experimental treatment on our key outcomes. For all our main dependent measures, it did. For instance, the AR counter-protester manipulation, which significantly predicted AR support (p < .001), became nonsignificant (p = .14) when we introduced a term for the extent to which AR counter-protesters were perceived to be violent. These results suggest that the observed findings were driven by AR violence rather than the counter-protesters’ chants. Second, as indicated in the results section above, a within-scope analysis of Republicans in our sample showed similar results for WN violence, for which this confound was not present. These results further confirm that differences between our experimental conditions were driven by violence. But future research should more carefully distinguish physical violence from other actions, including other “extreme” protest tactics (Feinberg et al. 2017).

It is also important to ask whether our results might have been influenced by participants’ perceptions of Antifa. Although the newspaper articles participants read did not label the AR counter-protesters as members of Antifa, the experiment took place at a time when the social movement was becoming more widely known. And because Antifa is arguably the most prominent social movement associated with antiracist counter-protesting that sometimes uses violence, it could be that our manipulation is confounded with whether the counter-protesters are assumed to be Antifa members or not.

On the one hand, the tactics of Antifa are heterogeneous and do not always entail violence (Bray 2017). Thus, it would be reasonable to assume that at least some of the protesters are Antifa in both conditions, neutralizing any confound concerns. But it may be that those who are exposed to conservative media outlets, which can be expected to be much more critical of Antifa, may perceive Antifa to always use violence. If so, for consumers of conservative media, it could be that negative perceptions of anarchism, more than violence, drove the results we observed. To examine whether participants’ responses to AR violence depended on political orientation, we compared the responses of the more conservative (those responding 6 or 7 on a 7-point scale) and liberal (those responding 1 or 2 on the scale) participants in our sample to the AR violence manipulation. Results of the ANOVA showed, unsurprisingly, a main effect of the political ideology variable on support for the AR counter-protesters (p < .001) and, as expected on the basis of the theory, a main effect of violence by the AR counter-protesters (p < .001). Importantly, however, political ideology did not interact with AR violence (p = .97), meaning that the responses of the most conservative participants in our sample to AR violence were similar to those of the most liberal participants. Furthermore, in separate mediation models of each subset of participants, the effects of the AR violence, which significantly predicted AR support (p < .001), became nonsignificant (p > .90) when we introduced the manipulation check measuring the extent to which AR counter-protesters were perceived to be violent. This further suggests that across the political spectrum, it was the heightened perception of violence, rather than other characteristics of Antifa (e.g., anarchism, anticapitalism) that drove our findings.

To maximize our ability to draw causal inferences, our experiment necessarily limited the number of factors that we could investigate. As one example, although the

---

8 Results for other outcomes are similar and available upon request.
newspaper accounts participants read did not suggest that all AR counter-protesters were violent or not violent, our study did not attempt to capture the heterogeneity and fragmentation that often exists in demonstrations or movements, including among AR counter-protesters. In the real world, those who act in opposition to white supremacy range from those who are committed to nonviolence in all forms to groups, such as Antifa, that have supported the use of violence. And as noted above, even groups such as Antifa are themselves heterogeneous, using a range of strategies. Whether such fragmentation and heterogeneity affect the successes of movements has been studied by social movements theorists as “radical flank effects” (Haines 1984, 2013). This line of work has hypothesized that the radical flank of a larger movement, in terms of both advocacy positions and tactics, can have positive effects on movement success, by directing support to more moderate arms of the movement using more mainstream tactics. The alternative possibility, that radical flank groups erode support for the entire movement (the “negative radical flank effect”) has also been hypothesized (e.g., Ellefsen 2018).

Because our research design did not explicitly distinguish between these factions, it cannot directly speak to whether violence by a subset of a movement can alter support for a more moderate faction or the broader movement. However, we found that participants exposed to violence by antiracists were not more willing than others to state that they supported groups who protested antiracism nonviolently, offering suggestive evidence for the absence of either positive or negative radical flank effects. On the other hand, recent work on radical flanks (Ellefsen 2018) suggests that we cannot simply ask whether the impact of radical flanks is positive or negative, because they may be negative in one arena and positive in another. For instance, a radical flank may have negative effects on public perceptions of the movement as a whole but still prompt concessions from governmental agencies. Similarly, Ellefsen (2018) noted that whether radical flanks help or harm the broader movement may vary over time. Extreme tactics may turn off observers in the short run, but over time the public may come to view moderate factions as reasonable alternatives to both the radical flank and the status quo.

This raises another important limitation of our research, namely, that long-term public opinion effects could be different from the short-term effects we observed. Although we did not observe changes in support for a protest group’s cause after exposure to one instance of violence, it could be that longer term exposure to violent protest does erode support for the cause. This is consistent with the findings from a recent study of the public opinion effects of civil rights protests described above, in which repeated exposure to violent protests weakened public support for civil rights (Wasow 2017). Although longer term exposure is difficult to study in experiments, future research could extend the research design we used, in which a single encounter between white nationalists and antiracists is replaced with accounts of a series of encounters with similar qualities.

There are also potential practical benefits of violence that our research design did not attempt to capture. For instance, like other extreme or disruptive forms of protest (Amenta et al. 2009; Myers and Caniglia 2004; Oliver and Myers 1999), violence likely attracts more media attention, which can be useful in drawing public awareness to the group’s message and expanding participation in the movement (e.g., Gamson 1990). It is also important to note that we focused only on practical effects of violence (i.e., whether violence helps or harms public perceptions of demonstrators). In so doing, we left aside normative questions about whether or when violence is morally justified. Indeed, supporters of groups such as Antifa argue that antiracists are morally justified in using violence against demonstrably violent groups, further arguing that violence is often used as self-defense (Bray 2017). As Cornel West noted in the aftermath of Charlottesville, “If it hadn’t been for the antifascists protecting us from the neo-fascists, we would have been crushed like cockroaches.”

Yet the asymmetry in the effects of violence by each side that we observe in our research suggests that this is precisely the dilemma AR counter-protesters face when confronting white nationalists. A widely despised group known for violence has little to lose and much to gain if it can draw out violence from antiracists. Violence does little, if any, further damage to their reputations. They have much more to gain in that violence by antiracists can not only damage public support for antiracists; as our results show, it may also increase WN support.

But future work should test the theoretical predictions under conditions in which groups with competing demands have relatively similar levels of support from the public. That is, we chose to study white nationalists and antiracist groups because they provided a particularly timely test of our hypotheses, given the frequency of confrontations between these two groups since the presidential candidacy and election of Donald Trump. And as noted earlier, these particular groups also provide a conservative test of our hypotheses, because we might expect that public perceptions of those who use violence against reviled groups will suffer less than those who use violence against less reviled groups. Future work should move beyond the empirical case we looked at here to test the predictions in contexts in which neither group is so widely reviled, such as confrontations between pro-choice and prolife activists. Again, this could be done with the same basic procedures used in the current research.

**Conclusion**

The real test of nonviolence lies in its being brought in contact with those who have contempt for it.

—Mahatma Gandhi
Summing up, the present research makes a number of different types of contributions. First, we introduce a theory explaining when and why violence influences public support for protest groups. We propose that protest groups that use violence are supported less and that this operates through reduced perceptions that group members are reasonable and, in turn, diminished levels of identification with them. Furthermore, the theory predicts that violence also leads to increased support for groups perceived as opposing the violent group. Finally, the theory predicts an asymmetry in the effects of violence on public support. Whereas groups that are otherwise supported by much of the public suffer from the use of violence, groups that suffer from low levels of public support do not. We found consistent support for the theoretical predictions about both why and when violence reduces public support for groups. This theory can be extended in future research to address a range of other questions, including potential ways that violence can influence groups described earlier (e.g., by attracting media attention or affecting support for more moderate factions).

At an empirical level, the research offers a number of related contributions. Foremost, it offers a clear test of the theory, including the proposed mechanisms through which support is reduced via violence. It also offers a demonstration of the extremity of the negative effects of violent protest. As noted earlier, although previous research strongly suggests that the use of violence reduces support for groups, we do not know whether these effects occur even when used against widely reviled groups that have histories of violence. Prior work shows that people tend to view violence against those viewed as moral outcasts to be more justifiable (Bandura et al. 1996; Bar-Tal 1990; Opotow 1990). Thus, it is plausible that the predicted effects of violence on loss of support would not apply when used against groups widely perceived to be objectionable. But consistent with our hypotheses, we found that violence cost group members support even when it targeted white nationalists, a widely despised group, pointing to the robustness of the effects.

By testing the theory in a carefully controlled Web-based experiment, we also offer a demonstration of the effects of violent protest under controlled conditions that limit the range of alternative explanations for the results of past research on the effects of protest violence. Although the experiment necessarily reduces some of the complexity that characterizes accounts of protest and intergroup conflict, as discussed earlier, it provides a base on which future studies can add nuance without sacrificing the ability to test causal claims.

Finally, our research has practical importance. Violent confrontations between white nationalist right-wing protesters and AR counter-protesters have become more common since Donald Trump was elected president in November 2016 (Swenson 2017) and continue to make headlines (Friess 2018). These clashes have precipitated widespread debate in the popular media, among activists, and among ordinary citizens about whether protesters should oppose white supremacists and other members of the far right with violence. Although our research obviously cannot speak to whether violence toward hate groups is morally justified, it does speak to some of the practical implications. Specifically, in terms of the effects of violence on public support of anti-racist groups, our research suggests that violent protest tactics reduce support for the protest groups using those tactics and can even increase support for rival groups.

References


Author Biographies

Brent Simpson is a professor of sociology and the graduate program director at the University of South Carolina and currently coeditis Social Psychology Quarterly. His current research focuses on altruism, cooperation, collective action, and inequality. His recent research has appeared in the American Sociological Review, the Proceedings of the National Academy of Sciences, the Annual Review of Sociology, and Social Networks.

Robb Willer is a professor of sociology and the director of the Morality and Hierarchy Laboratory at Stanford University. He studies a variety of topics, including the social psychology of contemporary American politics. His research has appeared in the American Sociological Review, the American Journal of Sociology, the Journal of Personality and Social Psychology, the Proceedings of the National Academy of Sciences, and elsewhere.

Matthew Feinberg is an assistant professor of organizational behavior and human resource management in the Rotman School of Management at the University of Toronto. He earned his PhD in social and personality psychology in 2012 from the University of California, Berkeley. His research explores the interpersonal and psychological processes that unite people together to form cohesive and productive groups, organizations, and societies, as well as the underlying processes that divide people into opposing and polarized groups. His focus is particularly on the role morality, emotions, and status concerns play in these processes.