

The Dangers of False News: How Sensational Content and Outgroup Cues Strengthen Support for Violence and Anti-Muslim Policies*

Jeffrey Javed[†] Blake Miller[‡]

August 9, 2019

Abstract

Does reading sensationalized false news stories make people more violent and discriminatory? Research on false news has focused more on its diffusion and less on its effects. This study tested the effects of sensationalization, outgroup cues, and public opinion perception on support for violence and anti-Muslim policies. We used an online survey experiment with a realistic, interactive website treatment detailing a homicide story in small-town America, written in the style of a false news article. We found that sensationalist language increased individuals' support for violence by provoking feelings of anger and fear, while identifying the suspect in the homicide as a Muslim refugee, versus specifying no outgroup affiliation, increased support for anti-Muslim policies. Lastly, perceived public support for violence increased the likelihood of upvoting or writing violent comments. This study contributes to our understanding of the ill effects of false news and the public debate on online content moderation.

*We would like to thank the Lieberthal-Rogel Center for Chinese Studies and the Rackham Graduate School at the University of Michigan for their generous funding. We would also like to thank our research assistants, Grace Phillips and Lyla Yu, for their help with text annotation. We would also like to thank Pablo Barbera, Iza Ding, Mary Gallagher, Mai Hassan, Hakeem Jefferson, Tom O'Mealia, Walter Mebane, Steven Moore, Dan Slater, Mike Thompson-Brusstar, Nick Valentino, Nicole Wu, and members of the Center for Political Studies Workshop and the Comparative Politics Workshop at the University of Michigan for their feedback on previous drafts of this paper.

[†]Research Fellow, Weiser Center for Emerging Democracies, University of Michigan (Email: jeffrey.javed@gmail.com)

[‡]Assistant Professor, London School of Economics (Email: b.a.miller@lse.ac.uk)

1 Introduction

Social media companies have identified false news as a significant threat to the integrity of their online communities across the globe (Hern 2018; Lyons 2018). Media coverage has frequently linked the dissemination of false news stories—particularly those detailing sensationalized accounts of criminality or moral wrongdoing—to politically-motivated attempts to mobilize violence against vulnerable minority groups. In India, the largest market for Facebook’s WhatsApp platform, political parties have used WhatsApp to spread false stories of murder and religious sacrilege by Muslims to inflame Muslim-Hindu tensions (Parth and Bengali 2018). A recent United Nations report claimed that the military in Myanmar had disseminated false news on Facebook that may have driven genocidal violence against the Rohingya (Mozur 2018). Non-state actors similarly use false news toward violent ends. The Sri Lankan government blocked Facebook after witnessing an increase in mob violence in response to fake, sensationalized content about the alleged misdeeds of Muslims spread by extremist groups (Goel, Kumar and Frenkel 2018). In the US, the proliferation of anti-immigrant false news from the alt-right has coincided with a dramatic rise in the number of reported hate crimes, as recorded by the FBI and the Southern Poverty Law Center (Barrett 2018). While systematic data on the intent of American alt-right content creators are unavailable, there is anecdotal evidence that some of these producers intentionally seek to promote or normalize violence. Notably, a leaked “style manual” from the neo-Nazi website *The Daily Stormer* instructed its writers that: “It’s illegal to promote violence on the Internet. At the same time, it’s totally important to normalize the acceptance of violence as an eventuality/inevitability” (Feinberg 2017).

Do these kinds of false news stories featuring sensationalized claims of wrongdoings by targeted minority groups affect mass attitudes toward the acceptability of violence and

anti-minority policies? Despite growing concerns over the global proliferation of false news through social media, we know more about its patterns of diffusion than its effects (Brady et al. 2017; Lazer et al. 2018; Vosoughi, Roy and Aral 2018); and we know nearly nothing about its effects on violent attitudes and behavior. Social media provides strong economic incentives to produce and spread sensationalized, outrage-inducing content (Crockett 2017), yet we do not know if fabricated news content increases the acceptability of violence, nor do we understand the cognitive mechanisms by which such an effect would occur. Research on related phenomena of hate speech and propaganda provide scant guidance: research has focused mainly on the legal and ethical dimensions of hate speech moderation and less on its real-world effects (Gates Jr et al. 1996; Waldron 2012), and observational research on genocidal propaganda has produced mixed results (Fujii 2004; Hagan and Rymond-Richmond 2008; Straus 2007; Yanagizawa-Drott 2014). While psychologists have found links between the consumption of violent media and aggressive behavior, this research has been restricted to children (Anderson and Sobel 2003; Drabman and Thomas 1974; Huesmann and Miller 1994).

Our study examines the effects of sensationalization, outgroup cues, and public opinion perception on support for violence and anti-Muslim policies. We used an online 2^3 factorial experiment with a realistic, interactive website treatment detailing a homicide story in small-town America, written in the style of a false news article. Through an analysis of survey, text, and online behavioral data, we find that sensational content—text, visual memes, and videos that employ moral-emotional language to present claims of moral violation or wrongdoing—increases individuals’ endorsement of violence by provoking anger and fear. In addition, we find that identifying the suspect in the homicide as a Muslim refugee, versus specifying no outgroup affiliation, increased support for anti-Muslim policies, while perceived public support for violence increased the likelihood of upvoting or writing violent comments.

This argument builds on research in social psychology that suggests that exposure to moral violations lowers an individual’s threshold for using violence against perceived norm

violators by provoking outrage and providing justifications for their punishment (Baumeister 1999; Beck 1999; Crockett 2017; Fiske and Rai 2014; Goldberg, Lerner and Tetlock 1999) and explores the mechanisms by which political and non-state actors can mobilize a critical mass of support for violence among the public. The rest of this article proceeds as follows. The next section briefly reviews the literature on political violence, political mobilization, and emotion. We then detail our experimental design and variable measurement strategies and present our results. The final section concludes with a discussion of how these findings apply to the false news phenomenon and violent political mobilization more generally.

2 Literature Review

The scholarship on collective violence, genocide, and war has long noted the outsized role of fabricated, sensational content in promoting violence against a targeted outgroup (Charny 2019; Cohn 1967; Dower 1986; Fein 1979; Herf 2006; Hill 1995; Goldhagen and Wohlgeleitner 1997; Goldhagen 2009; Tsesis 2002). For example, the circulation of Cotton Mather’s *Memorable Providences* detailing “real” accounts of bewitchings of innocent children is thought to have fueled the persecution of witches in the Massachusetts Bay Colony (Hill 1995); the *Protocols of the Elders of Zion*, which described a plan for Jewish world domination, was an important precursor to anti-Semitic Nazi mobilization (Cohn 1967; Herf 2006); and sensationalized accounts of Japanese atrocities against American POWs featured heavily in American anti-Japanese propaganda designed to recruit soldiers and bolster their fighting spirit (Dower 1986).

Few scholars, however, have directly tested the causal effects of content on violent attitudes, and the results of these studies have been mixed. Looking at political attitudes, Kalmoe (2014) finds that violent rhetoric interacts with trait aggression to increase support for political violence. In the context of genocidal violence, Hagan and Raymond-Richmond (2008) find that the use of dehumanizing racial epithets during attacks on black African pop-

ulations in Darfur correlated with more extreme violence. Other scholars have challenged the causal link between hateful content and violence. Looking at the Rwandan genocide, Straus (2007) argues that the spatial coverage of Hutu-controlled “hate radio” cannot explain the geographic variation in the onset of violence, while others have argued that anti-Tutsi propaganda were critical to driving violence or, at the very least, normalizing it (Fujii 2004; Yanagizawa-Drott 2014).

Moreover, it is unclear what kinds of content promote violence. Though scholars have emphasized the primacy of dehumanizing content in genocidal violence (Fein 1979; Charny et al. 1982), there is little empirical evidence for its efficacy. A major underexplored alternative is moral frameworks that justify the righteousness of violence (Viterna 2014). Social psychologists have long emphasized the significance of morality—conceptions of right and wrong traits and behaviors—in understanding participation in violence, private and political (Bandura, Underwood and Fromson 1975; Baumeister 1999; Beck 1999; Fiske and Rai 2014). Because using violence requires first overcoming formidable moral-emotional reservations, moral narratives that condone and justify violence can erode these otherwise powerful restraints on violent behavior (Baumeister 1999; Beck 1999; Fincher and Tetlock 2016). Moralistic content may also mobilize participation in violent causes by drawing upon individuals latent moral convictions. Wood (2003) emphasizes the pleasure of agency that participants derive from acting on their moral convictions in a movement, which Viterna (2014) extends to the willingness of citizens to accept violence in the name of righteous causes, movements where “interested publics believe that the enactors of political violence are defending society’s most vulnerable and protecting a morally legitimate social order.” Indeed, Kirkpatrick (2008) documents a long history of extrajudicial and vigilante violence in America that invokes revolutionary values of freedom, justice, and democracy to frame itself as righteous.

Understanding the effects of highly moralized content on violent behavior requires attention to the emotional responses this content provoke, in particular anger and fear. An

abundant literature has analyzed the role of emotion in voter mobilization (Ansolabehere and Iyengar 1997; Banks 2014; Brader 2005, 2006; Freedman and Goldstein 1999; Huber et al. 2015; Marcus, Neuman and MacKuen 2000; Mendelberg 2001), and, more recently, in the context of misinformation (Vosoughi, Roy and Aral 2018); however, the emotional microfoundations of violent mobilization remain underexplored (Viterna 2013). Anger, or outrage, appears to be the most relevant mobilizing emotion in the context of violence. Unlike fear, which tends to demobilize, social psychologists have found that anger and outrage have mobilizing effects (Banks 2014; Lerner and Keltner 2001; Valentino et al. 2011). Significantly, content that conveys or sensationalizes moral violations may lower an individual's threshold for using violence through provoking moral outrage against violators, which motivates people to shame and punish wrongdoers (Crockett 2017; Goldberg, Lerner and Tetlock 1999). Under some circumstances, however, uncontextualized, or generalized, anger may be sufficient to mobilize. Banks (2014) argues that non-group specific appeals that stimulate generalized anger are sufficient to mobilize voters because they trigger latent racist feelings that, ultimately, are rooted in experiences of anger. These studies, however, test emotions acontextually and do not consider their possible mediating role. Although research has shown that messages that provoke anger are more likely to mobilize political participation (Ryan 2012; Valentino, Hutchings and White 2002; Valentino et al. 2011) and moral-emotional content is more likely to be shared (Brady et al. 2017), it is unclear if such content can influence violent attitudes or engagement and whether anger, fear—or both—mediate this relationship.

3 Theory and Hypotheses

We argue that false news content makes individuals more punitive by sensationalizing transgressive behavior with highly moralized and emotional language. We hypothesize that the mechanism linking sensationalized content and increased support for violence is outrage—

i.e. anger that arises from perceived violations of cultural norms of right or wrong behavior (Crockett 2017; Stets 2012; Turner and Stets 2006). If there is a cue that explicitly links the norm violator to an outgroup, we predict that individuals will be more likely to support negative sanctions on that entire outgroup. We expect an interaction between content type and the outgroup cue: we predict that sensationalized content and the presence of an outgroup cue will greatly increase support for sanctions on the outgroup. In addition, we consider that individual support for violence will be partly a function of perceptions of peer attitudes. Perceived public opinion shapes attitudes and behavior (Neubaum and Kramer 2017; Noelle-Neumann 1993). Public opinion perception may be particularly relevant for punitiveness, since a perceived audience to one’s decision-making increases an individual’s willingness to punish moral transgressions (Kurzban, DeScioli and O’Brien 2007) because proposing punishment of moral transgressors signals to ones peers that one is of high quality and potentially a good cooperator (Fessler and Haley 2003; Gintis, Smith and Bowles 2001). The perceived audience consensus on treating a violation may possibly influence individual attitudes: if one’s peers are uniformly signaling support for punishment, we expect that an individual will want to conform to the peer consensus to avoid reputational loss. We do not hypothesize any interaction effects between the peer effects treatment and the other treatments.

4 Experimental Design and Variable Measurement

Observed correlations between exposure to sensational content and the endorsement of violence could be due to selection effects, since people prone to violence might be attracted to such news in the first place. For this reason, we used a survey experiment embedded in a realistic, fully interactive online news article about a local homicide in the US to examine how content sensationalization, outgroup cues, and violent comments by other online users may mobilize or inhibit violent attitudes and online engagement. We chose a homicide story

for several reasons. Homicide stories have long been a mainstay of national and local news media in America and a possible source of anti-outgroup bias (Gilliam Jr and Iyengar 2000). Moreover, as mentioned earlier, there is plentiful anecdotal evidence that misleading crime stories feature heavily in false news content used to mobilize violence and outrage on social media. In addition, psychologists have frequently used crime vignettes to test the effects of perceived violations on punitiveness, though has been done mainly outside of a realistic news context (Fincher and Tetlock 2016; Tetlock et al. 2007).

To reduce the artificiality of our treatment, we constructed a news site that mimicked the Associated Press (AP) website and included a functioning comment section (see pp. 1-2 in the SI for a description of the website).¹ The use of the AP template is significant not only for its perceived objectivity but also because some extremist websites refer to it explicitly as a desirable format for presenting their own stories (Feinberg 2017). Additionally, the concise style of AP newswire articles makes the short form of a standard survey experiment vignette seem more realistic and nonpartisan. We had professional journalists at Politico and the New York Times vet our vignettes and our constructed website for plausibility and proper news formatting.

We used a two-by-two-by-two factorial design that randomly assigned individuals to a combination of three two-level factor treatments, for a total of eight possible combinations. The treatment vector was a news article that described an alleged homicide, the identity of the suspect, and retaliatory violence against the suspect by a group of locals in a small American town (see Figures S2-S5, pp. 4-6). The three treatment factors were: 1) the description of the homicide, sensationalized (“gruesome slaying of a local child”) or objective

¹The interactive survey web app was custom coded using the Python Django web framework and deployed to the web using the Amazon Elastic Beanstalk service. It was designed to be responsive, working on any screen size or device. Click data, comments, and browser metadata were stored in a PostgreSQL database hosted on Amazon Relational Database Services (RDS).

(“homicide of a local child”); 2) the presence of an outgroup cue, identifying the perpetrator as a “middle-aged male Muslim refugee” or not (“middle-aged male”); and 3) the content of the most upvoted comment (violent or conciliatory). Below each vignette were three comments based on real user comments, web scraped from Breitbart News, that express: 1) support for violence against the alleged perpetrator (“If he killed my child: I’d have nothing to live for. I’d rain fiery retribution down on anyone who killed my child. His end would be brutal.”); 2) neutral information-seeking about the event (“What could have driven this man to kill a child?”); and 3) conciliatory (“I hope that this conflict can be resolved peacefully. My heart goes out to the victims”). Regardless of the treatment condition, the top comment had 53 likes, the middle comment had eight, and the bottom comment had two. This distribution of 63 likes follows a power-law distribution that typically characterizes the distribution of likes on social media. In order to reduce experimenter demand, we included distractor questions before and after the treatment article that asked respondents about their online shopping habits.

The pre-treatment survey contained a battery of background questions concerning respondents’ demographic characteristics, party identification, and individual dispositions towards authoritarianism, symbolic racism, and ethnocentrism. We use Feldman and Stenner (1997)’s four-item Child Rearing Values (CRV) scale to measure authoritarian personality. Our symbolic racism and ethnocentrism scales are taken directly from ANES. We calculate ethnocentrism by averaging outgroup feeling thermometers towards Asians, Blacks, Hispanics, Muslims, and refugees. We calculate symbolic racism by averaging the four items from the ANES symbolic racism scale, reversing two items to make the scale range from low to high racism.

The post-treatment survey asked a series of questions regarding respondents’ emotional responses to the homicide and the attack on the alleged perpetrator, separately. Each emotion barometer asks respondents to indicate their felt intensity of eight emotions using five-point scales. To avoid biasing respondents’ emotional reactions, we included a list of eight

emotions, positive and negative, based loosely on Robert Plutchik’s typology of basic discrete emotions. We measured violent attitudes directly and indirectly using survey scale and open-ended responses. To measure punitiveness toward the alleged perpetrator in the story, we asked respondents to rate their support for nine punishments on seven-point (-3 to 3) scales. To test for support for extrajudicial violence, we ask if the perpetrator should be tortured and if he should be jailed without trial. Because direct questions about violence are subject to social desirability bias, we included indirect measures of punitiveness. We asked respondents how they think the locals who engaged in street violence against the suspect should be punished for beating the homicide suspect, and we ask if they would be willing to donate to a legal defense fund for the attackers. To measure attitudes toward outgroups, we presented respondents with four seven-point (-3 to 3) scales to indicate their support for or opposition to policies on refugee immigration, Muslim immigration, and monitoring and registering Muslim communities.

In addition to these survey scale responses, we collected behavioral data on respondents’ interaction with the news website and text data from respondents’ written reactions to the article. Our survey instructions encouraged respondents both to read the article and to interact with the website by liking, reporting, or posting comments in the comments section below the text of the news story, though doing so was not mandatory. Respondents’ comments and any clicks on “report” or “like” buttons were stored in our database as behavioral measures of support or disapproval of violence. We asked respondents to provide twenty-word responses to the homicide and the mob attack on the suspect in two separate open-ended questions. We devised a novel typology to classify violent content (see Figure S14, p. 23), which a team of four researchers, blind to treatment condition, used to hand annotate all text data collected in open-ended responses and comments. Agreement between coders was nearly perfect according to Cohen’s kappa and F1 measures (see Table S15, p. 24).² After exiting the survey, we debrief respondents to ensure that they understood that

²To measure inter-coder reliability, two coders coded a random sample of open-ended

the news article we presented was fictitious and why the study required this use of deception. The IRB reviewed this study and deemed it exempt; the research design, text coding scheme, and experimental vignettes for this study were pre-registered prior to conducting the study.

Our sample of 1655 respondents, recruited through Amazon’s Mechanical Turk in early December 2018, roughly approximated the demographics of the American population. Within our sample, 49 percent identified as female and 77 percent as white, in line with the US Census Bureau’s 2018 population estimates of 51 percent and 77 percent, respectively. The median age of our respondents was 36, versus 38 in the population. Looking at Gallup’s 2017 estimates of party affiliation, our sample was over-represented in self-identified Democrats (41.3 percent versus 29 percent) but proportionally representative of Republicans (28 percent). To ensure proportional geographic representativeness, we stratified our sampling according to the population size of each of the nine regional divisions designated by the US Census Bureau. This sample includes non-compliers who failed an attention check and respondents who received the treatment but did not finish the survey. Both noncompliance and attrition were extremely low: less than 1 percent failed the attention check and 2.54 percent did not complete the entire survey post treatment. Here we conduct an intent-to-treat analysis and do not exclude subjects on any post-treatment variables. The Supporting Information (pp. 2-4) provides an in-depth description of the sampling and randomization process.

responses. Cohen’s kappa is an inter-coder reliability measure for nominal scales and two coders. Though there is no agreed-upon interpretation of the kappa coefficient, it has been suggested that 0.61-0.80 indicates substantial agreement, and 0.81-1.0 indicates near-perfect agreement. F1 is a common performance measure in machine learning ranging from 0 to 1. It is defined as the harmonic mean of precision and recall. F1 is a good measure of coder agreement when choices are imbalanced. With imbalance in choices, percent agreement (accuracy) will overestimate coder agreement.

5 Results

Sensational content significantly increased support for violence when measured indirectly through questions about the mob attack on the alleged perpetrator and open-ended responses on the events in the article.³ Respondents presented with sensational content were more likely to donate to a legal defense fund for people who engaged in a mob attack on the alleged perpetrator ($P < 0.05$), and were more likely to believe that not punishing them was justified, though this finding is not statistically different from zero at conventional levels of significance (Figure 1). This effect for sensational content further holds when looking at the text data from the open-ended responses. We find that sensational content significantly increased the probability that respondents will express support for violence in their written responses, including extrajudicial, mob-style violence ($P < 0.05$); for both measures this effect represents a roughly fifteen percent increase in probability (Figure 2).⁴

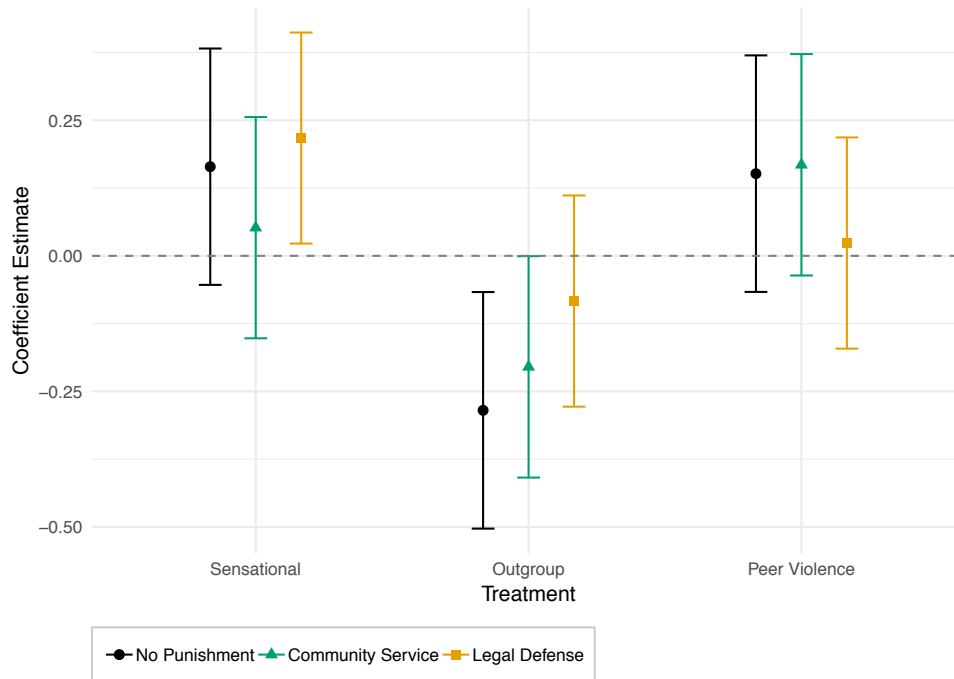
We used a structural topic model (STM) to systematically summarize the responses of our subjects to the open ended questions (see Figure S11 and Table S8, p. 14-15). A cursory comparison of representative responses per topic illustrate differences in punitiveness across treatment conditions. One respondent given the sensationalized content treatment wrote the following, which was coded as supportive of violence: “I am absolutely disgusted, appalled,

³With the exception of support for detainment without trial, there are no significant average treatment effects for sensational content on punitiveness when looking at survey scale questions that directly asked respondents what kinds of punishments they would support for the alleged perpetrator in the article (see Figure S7, p. 7). We expected these direct measures to be vulnerable to social desirability bias.

⁴The survey scale and text measures of support were highly correlated. Expression of support for violence in the open-ended responses significantly and positively correlates with indirect survey scale responses as well as support for the death penalty for the alleged perpetrator (all $P < 0.001$) (see Figure S6 and Table S1, pp. 6-7).

and at a loss for words about this attack. They should kill him in the street. Let everyone who wants a piece of him have a piece of him.” Contrast this with a non-violent response in the control condition: “I want to know what happened. I want to know if the person that was beat up is the one accused of the murder. I want to know the details.” See the Supporting Information (pp. 17-18) for example open-ended responses and comments by coding category and treatment condition.

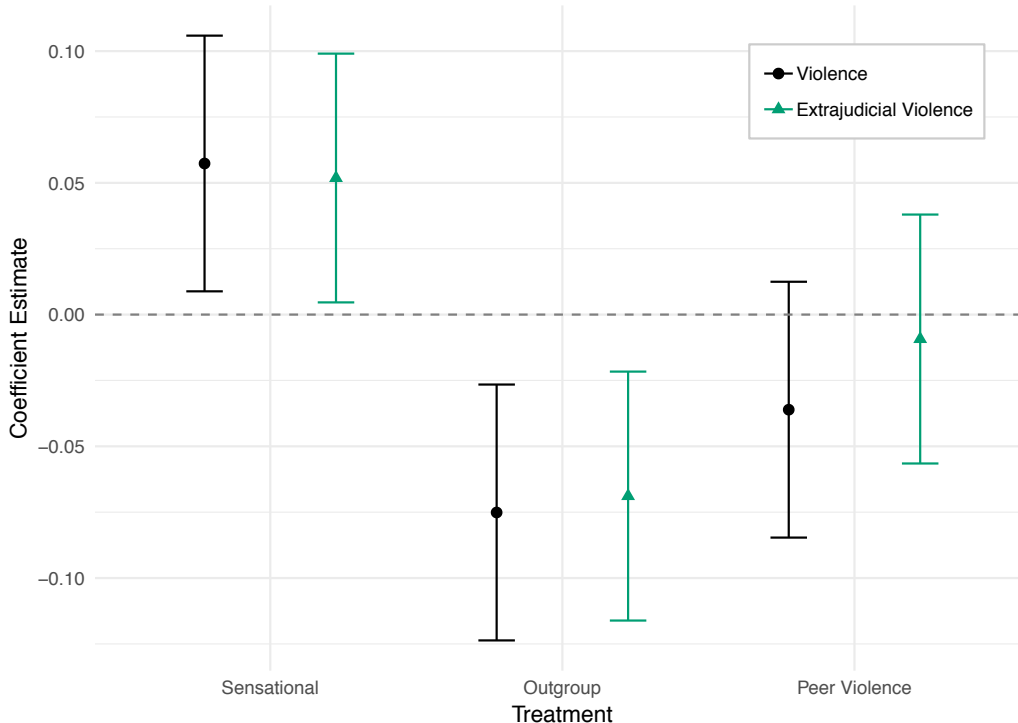
Figure 1: Support for the Mob Attack by Treatment Condition



Unexpectedly, we find that outgroup cues decreased support for violence against the perpetrator in the story when measured indirectly and in the open-ended responses ($P < 0.05$) (Figures 1 and 2). We suspect this negative effect was driven by perceptions that the event or even the article itself was racist. Respondents were significantly more likely to believe that the mob attack or the article itself was racially-motivated when given the outgroup treatment (see Figures S12-13, p. 22).

We find that sensational content and violent peer influence affected how people interact with online content. The sensational content treatment increased the probability that

Figure 2: Support for Violence in the Open-ended Responses by Treatment Condition



respondents would leave a comment on the website by roughly sixteen percent ($P < 0.05$) (Figure 3). Since the act of leaving a comment is vulnerable to significant selection effects, we run Heckman two-step selection models to detect and correct for selection bias (Heckman 1979). Older respondents and respondents scoring high on symbolic racism were far more likely to leave a comment. After taking this into account, we find that the sensational treatment increased the probability of writing a violent comment by roughly twenty-four percent, though this effect is not statistically significant at conventional levels ($P = 0.14$ level) (Figure 4).

While peer signals mattered little for attitudinal outcomes, they consistently affected how respondents interacted with the news article, creating a “cascade effect” (Bikhchandani, Hirshleifer and Welch 1992) for endorsement of violence. That is, respondents bandwagoned onto violent comments if they were already highly-liked. If the top, most-liked comment was violent, respondents were more likely to like a violent comment ($P < 0.05$) and to leave a violent comment of their own ($P < 0.05$); and far less likely to like a non-violent comment

($P < 0.001$) (Figures 3 and 4). In addition, respondents were less likely to report the violent comment when it was highly liked, though this effect was not statistically significant ($P = 0.12$). Because the most liked comment was also positioned at the top of the three comments—as is normally the case in a comments section on a news website—this effect may possibly be driven by the position of the comment rather than the fact that it was highly liked; that is, respondents may have been merely satisficing by interacting with the upvoted violent comment because it was at the top of comment list. We believe the violent comment’s upvoting was more important than its positioning for two reasons. First, we found that the upvoted violent comment treatment had a *negative* effect on reporting the violent comment and a *positive* treatment effect on expressing violence in a written comment, which suggests the effect is not due to simply interacting with whatever comment was positioned at the top of the comments list. Second, each comment was only one or two lines long, so users would have seen all three comments at once when they scrolled down; it is unlikely that the bottom two comments would have been less visible to respondents.

Figure 3: Behavioral Outcomes by Treatment Condition

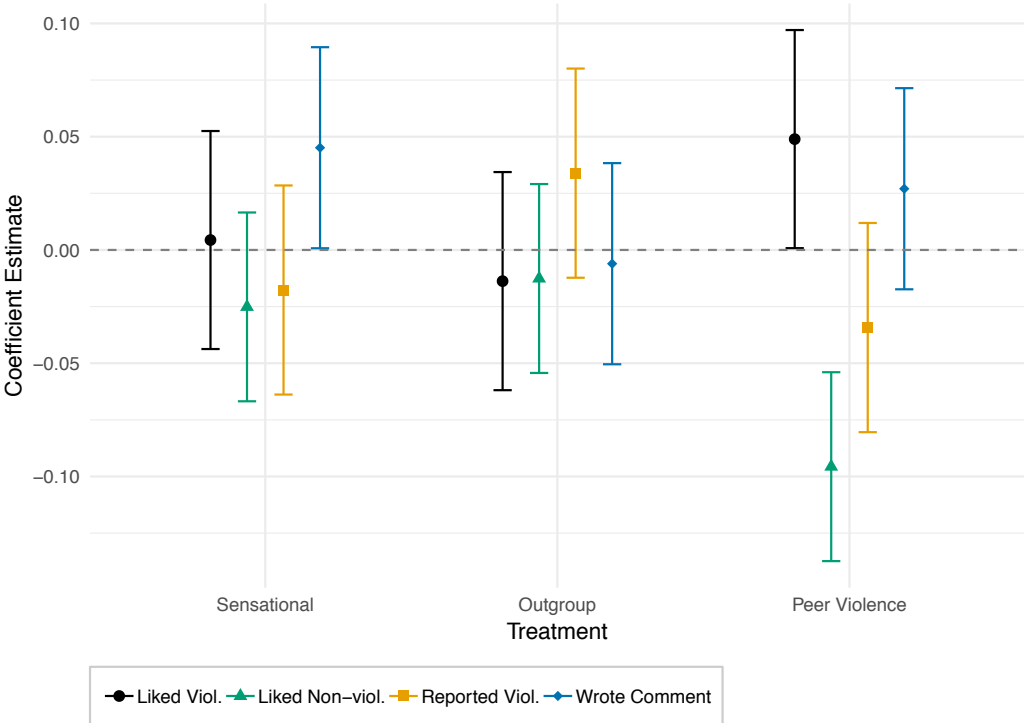
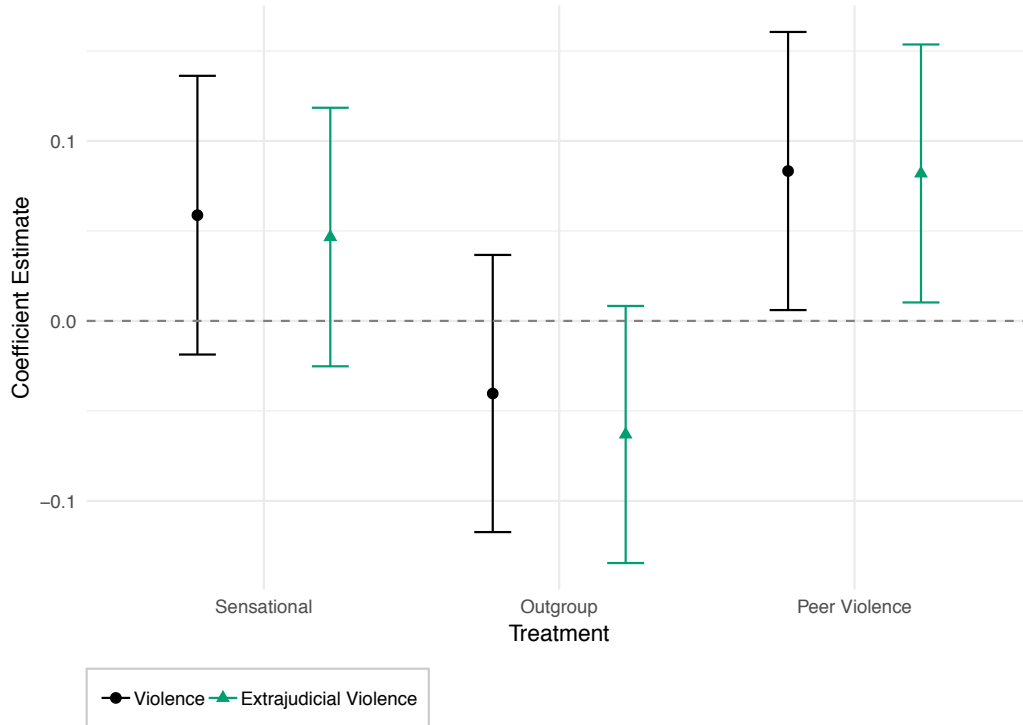
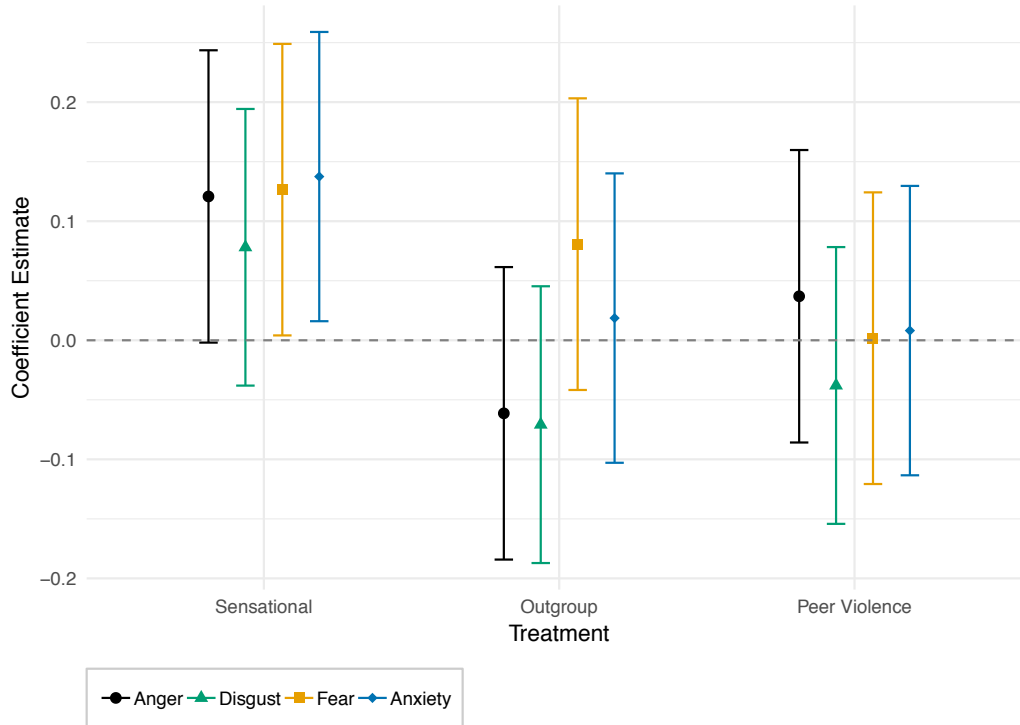


Figure 4: Comment Outcomes by Treatment Condition



Sensational content elicited strong emotional responses of anger ($P < 0.1$), fear ($P < 0.05$), and anxiety ($P < 0.05$) from respondents, while outgroup cues and peer support did not (Figure 5). To gauge whether emotional mechanisms mediated the treatment effect of sensational content on violent attitudes and behavior, we use causal mediation analysis (Imai, Jo and Stuart 2011). Because emotion was not randomly assigned, violating the assumption of sequential ignorability, we include possible confounding covariates to the models—i.e. measures of authoritarianism (Stenner 2005), ethnocentrism (Kinder and Kam 2010), symbolic racism (Banks 2014), and whether the respondent resided in the American South (Nisbett 1996). Causal mediation analysis revealed that anger, fear, and anxiety mediated the effect of sensational content on support for the mob attackers’ legal defense fund ($ACME = 0.06$, 0.04 , and 0.03 , respectively; all $P < 0.05$) and the likelihood of expressing support for violence in the open-ended responses ($ACME = 0.007$, 0.003 , and 0.002 , respectively; all $P < 0.1$) (see Figures S8-10, pp. 13-14). We do not find emotional mediating effects for the other two treatments or for behavioral and anti-outgroup outcomes.

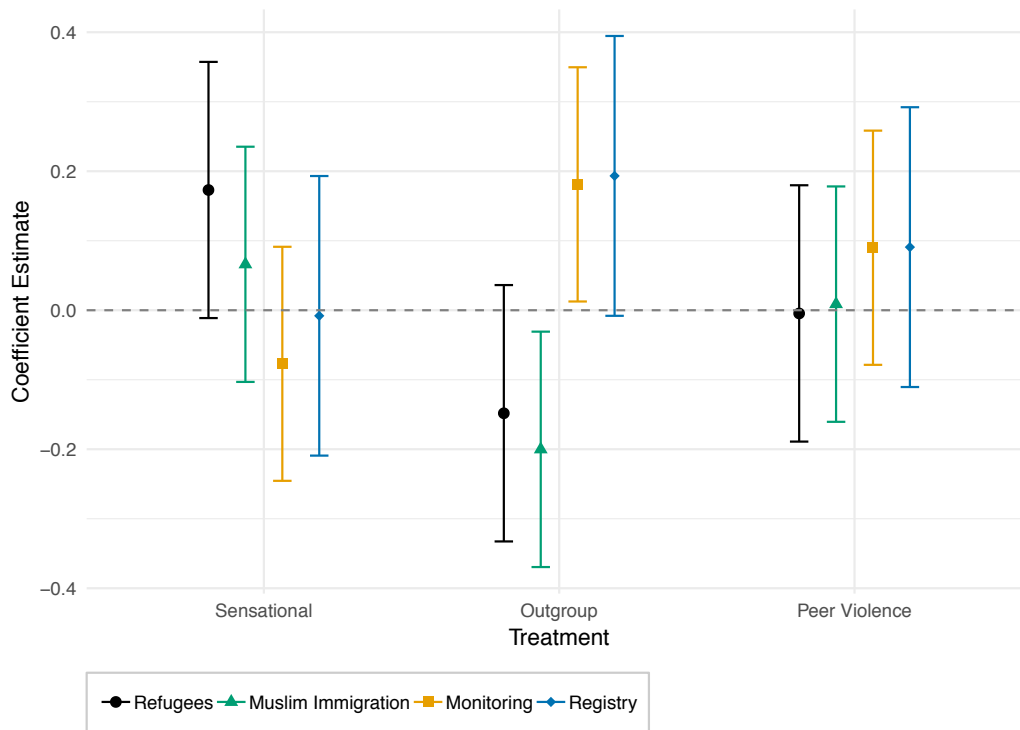
Figure 5: Emotional Response by Treatment Condition



Did this increased punitiveness translate into support for sanctions on Muslims and refugees? We find sensational content and peer support for violence did not increase support for anti-Muslim and anti-refugee policies; however, the mere inclusion of an outgroup cue on average increased support for anti-Muslim and anti-refugee policies (Figure 6). Identifying the alleged perpetrator as a Muslim refugee caused respondents to support tightening refugee quotas ($P < 0.1$) and restricting Muslim immigration ($P < 0.05$), and positive for support for increased monitoring of Muslims ($P < 0.05$) and approval of a Muslim “registry” ($P < 0.1$). This finding contrasts with the earlier finding that including an outgroup cue increased individuals’ suspicion that the attack described in the article, or the article itself, was racist. Indeed, this contradiction appeared in some of the open-ended responses in the outgroup cue treatment condition. One respondent in the outgroup cue treatment condition wrote: “I was thinking it was a fake story to invoke Islamophobia...I also do not want them moving next door to me or in our country, they can keep their backwards traditions in their own countries.” Perceived racial bias or untrustworthiness of the content’s source did not temper

respondents' attitudes towards Muslims and refugees.

Figure 6: Support for Anti-Muslim Policies by Treatment Condition



This increased punitiveness towards refugees and Muslims may be a function of a general increase in punitiveness. Outrage arising from morally transgressive behavior may trigger an “intuitive prosecutor” mindset that heightens one’s desire to punish all subsequent transgressors, regardless of their connection to the original transgressor (Crockett 2017; Goldberg, Lerner and Tetlock 1999; Tetlock et al. 2007). To this end, we looked at the effects of the treatments on preferences regarding punishment of homicide, the crime featured in the article. We did not find a significant effect between sensational content, outgroup cues, or peer support on respondents’ attitudes towards the punishment of homicide.

Contrary to our expectations, we did not find a significant interaction effect between outgroup cues and moral-emotional content on anti-immigrant and anti-Muslim policy preferences, and the effect is not in the expected direction. The average treatment effect for peer support is insignificant, as is its interaction between outgroup cues, though the signs of these effects are in their expected direction (see Table S5, p. 11).

6 Discussion

This study contributes to our understanding of false news and mass opinion in three ways. First, this is one of the first studies, to our knowledge, to demonstrate the causal effects of false news consumption on attitudes towards violence and online behavior; in contrast, the burgeoning literature on misinformation has focused mainly on its patterns of diffusion. Second, it identifies the importance of both anger and fear as emotional mediators of the effects of false news consumption on punitiveness. Third, we show that the threshold by which false news increases anti-outgroup sentiment is rather low: simply attributing a crime to a member of a certain outgroup is sufficient to increase discriminatory attitudes. Contrary to research that explicit outgroup cues do not increase support for anti-outgroup policies (Huber and Lapinski 2006; Mendelberg 2001) or are unnecessary (Banks 2016), we find that outgroup cues increase punitiveness towards an outgroup even when these cues raise concerns about racial bias.

Understanding the importance of online content in political life has perhaps never been more important than in the present. While these findings are a far cry from evidence that false news mobilizes on-the-ground violence, they underscore the importance of content and discourse in inflaming violent attitudes and carry significant policy implications for governments and social media companies working to stem the harmful effects of false news.

While outside the scope of this study, we believe that these findings are not restricted to the contemporary false news phenomenon or even online content: the use of sensational content to mobilize violence has a long genealogy. Why leaders choose these strategies for mobilizing popular support for violence and when they translate into on-the-ground violence remain crucial and promising areas for future scholarly inquiry.

References

- Anderson, Adam K and Noam Sobel. 2003. “Dissociating intensity from valence as sensory inputs to emotion.” *Neuron* 39(4):581–583.
- Ansolabehere, Stephen and Shanto Iyengar. 1997. *Going negative: How political advertisements shrink and polarize the electorate*. The Free Press,.
- Bandura, Albert, Bill Underwood and Michael E Fromson. 1975. “Disinhibition of aggression through diffusion of responsibility and dehumanization of victims.” *Journal of research in personality* 9(4):253–269.
- Banks, Antoine J. 2014. *Anger and racial politics: The emotional foundation of racial attitudes in America*. Cambridge University Press.
- Banks, Antoine J. 2016. “Are group cues necessary? How anger makes ethnocentrism among whites a stronger predictor of racial and immigration policy opinions.” *Political Behavior* 38(3):635–657.
- Barrett, Devlin. 2018. “Hate crimes rose 17 percent last year, according to new FBI data.” *Washington Post* .
- URL:** https://www.washingtonpost.com/world/national-security/hate-crimes-rose-17-percent-last-year-according-to-new-fbi-data/2018/11/13/e0dcf13e-e754-11e8-b8dc-66cca409c180_story.html?utm_term=.fdc68819676c
- Baumeister, Roy F. 1999. *Evil: Inside human violence and cruelty*. Macmillan.
- Beck, Aaron T. 1999. *Prisoners of hate: The cognitive basis of anger, hostility, and violence*. HarperCollins Publishers.
- Bikhchandani, Sushil, David Hirshleifer and Ivo Welch. 1992. “A theory of fads, fashion, custom, and cultural change as informational cascades.” *Journal of political Economy* 100(5):992–1026.

- Brader, Ted. 2005. "Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions." *American Journal of Political Science* 49(2):388–405.
- Brader, Ted. 2006. *Campaigning for hearts and minds: How emotional appeals in political ads work*. University of Chicago Press.
- Brady, William J, Julian A Wills, John T Jost, Joshua A Tucker and Jay J Van Bavel. 2017. "Emotion shapes the diffusion of moralized content in social networks." *Proceedings of the National Academy of Sciences* 114(28):7313–7318.
- Charny, Israel W. 2019. *How Can We Commit the Unthinkable?: Genocide: the Human Cancer*. Routledge.
- Charny, Israel W et al. 1982. "How can we commit the unthinkable." *Genocide: The Human Cancer (Boulder, CO: Westview Press, 1982)* 207.
- Cohn, Norman. 1967. *Warrant for Genocide: The Myth of the Jewish World Conspiracy and the*. New York: Harper and Row.
- Crockett, MJ. 2017. "Moral outrage in the digital age." *Nature Human Behaviour* 1(11):769.
- Dower, John. 1986. *War without mercy: Race and power in the Pacific War*. Pantheon.
- Drabman, Ronald S and Margaret H Thomas. 1974. "Does media violence increase children's toleration of real-life aggression?" *Developmental Psychology* 10(3):418.
- Fein, Helen. 1979. *Accounting for genocide: National responses and Jewish victimization during the Holocaust*. New York: Free Press.
- Feinberg, Ashley. 2017. "This is the daily stormers playbook." *Huffington Post* .
- Feldman, Stanley and Karen Stenner. 1997. "Perceived threat and authoritarianism." *Political Psychology* 18(4):741–770.

- Fessler, DM and Kevin J Haley. 2003. “The strategy of affect: Emotions in human cooperation 12.” *The Genetic and Cultural Evolution of Cooperation*, P. Hammerstein, ed pp. 7–36.
- Fincher, Katrina M and Philip E Tetlock. 2016. “Perceptual dehumanization of faces is activated by norm violations and facilitates norm enforcement.” *Journal of Experimental Psychology: General* 145(2):131.
- Fiske, Alan Page and Taze Shakti Rai. 2014. *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge University Press.
- Freedman, Paul and Ken Goldstein. 1999. “Measuring media exposure and the effects of negative campaign ads.” *American journal of political Science* pp. 1189–1208.
- Fujii, Lee Ann. 2004. “Transforming the moral landscape: the diffusion of a genocidal norm in Rwanda.” *Journal of Genocide Research* 6(1):99–114.
- Gates Jr, Henry Louis, Anthony P Griffin, Donald E Lively and Nadine Strossen. 1996. *Speaking of race, speaking of sex: Hate speech, civil rights, and civil liberties*. NYU Press.
- Gilliam Jr, Franklin D and Shanto Iyengar. 2000. “Prime suspects: The influence of local television news on the viewing public.” *American Journal of Political Science* pp. 560–573.
- Gintis, Herbert, Eric Alden Smith and Samuel Bowles. 2001. “Costly signaling and cooperation.” *Journal of theoretical biology* 213(1):103–119.
- Goel, Vindu, Hari Kumar and Sheera Frenkel. 2018. “In Sri Lanka, Facebook Contends With Shutdown After Mob Violence.” *New York Times* .
- URL:** <https://www.nytimes.com/2018/03/08/technology/sri-lanka-facebook-shutdown.html>

- Goldberg, Julie H, Jennifer S Lerner and Philip E Tetlock. 1999. "Rage and reason: The psychology of the intuitive prosecutor." *European Journal of Social Psychology* 29(5-6):781–795.
- Goldhagen, Daniel Jonah. 2009. *Worse than war: Genocide, eliminationism, and the ongoing assault on humanity*. Hachette UK.
- Goldhagen, Daniel Jonah and Maurice Wohlgelemer. 1997. "Hitlers willing executioners." *Society* 34(2):32–37.
- Hagan, John and Wenona Rymond-Richmond. 2008. *Darfur and the Crime of Genocide*. Cambridge University Press.
- Heckman, James J. 1979. "Sample selection bias as a specification error." *Econometrica: Journal of the econometric society* pp. 153–161.
- Herf, Jeffrey. 2006. "The jewish enemy." *Nazi propaganda during World War II and the holocaust* .
- Hern, Alex. 2018. "YouTube to Crack Down on Fake News, Backing 'Authoritative' Sources." *The Guardian* .
URL: <https://www.theguardian.com/technology/2018/jul/09/youtube-fake-news-changes>
- Hill, Frances. 1995. *A delusion of Satan: The full story of the Salem witch trials*. Doubleday.
- Huber, Gregory A and John S Lapinski. 2006. "The race card revisited: Assessing racial priming in policy contests." *American Journal of Political Science* 50(2):421–440.
- Huber, Michaela, Leaf Van Boven, Bernadette Park and William T. Pizzi. 2015. "Seeing Red: Anger Increases How Much Republican Identification Predicts Partisan Attitudes and Perceived Polarization." *PLoS ONE* 10(9).
- Huesmann, L Rowell and Laurie S Miller. 1994. Long-term effects of repeated exposure to media violence in childhood. In *Aggressive behavior*. Springer pp. 153–186.

- Imai, Kosuke, Booil Jo and Elizabeth A Stuart. 2011. "Commentary: Using potential outcomes to understand causal mediation analysis." *Multivariate Behavioral Research* 46(5):861–873.
- Kalmoe, Nathan P. 2014. "Fueling the fire: Violent metaphors, trait aggression, and support for political violence." *Political Communication* 31(4):545–563.
- Kinder, Donald R and Cindy D Kam. 2010. *Us against them: Ethnocentric foundations of American opinion*. University of Chicago Press.
- Kirkpatrick, Jenet. 2008. *Uncivil disobedience: Studies in violence and democratic politics*. Princeton University Press.
- Kurzban, Robert, Peter DeScioli and Erin O'Brien. 2007. "Audience effects on moralistic punishment." *Evolution and Human behavior* 28(2):75–84.
- Lazer, David MJ, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A Sloman, Cass R Sunstein, Emily A Thorson, Duncan J Watts and Jonathan L Zittrain. 2018. "The Science of Fake News." *Science* 359(6380):1094–1096.
- Lerner, Jennifer S and Dacher Keltner. 2001. "Fear, anger, and risk." *Journal of Personality and Social Psychology* 81(1):146.
- Lyons, Tessa. 2018. "Hard Questions: What's Facebook's Strategy for Stopping False News?" *Facebook Newsroom* .
URL: <https://newsroom.fb.com/news/2018/05/hard-questions-false-news/>
- Marcus, George E, W Russell Neuman and Michael MacKuen. 2000. *Affective intelligence and political judgment*. University of Chicago Press.

- Mendelberg, Tali. 2001. *The race card: Campaign strategy, implicit messages, and the norm of equality*. Princeton University Press.
- Mozur, Paul. 2018. “A Genocide Incited on Facebook, With Posts From Myanmar’s Military.” *New York Times* .
URL: <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>
- Neubaum, G. and Nicole Kramer. 2017. *Monitoring the Opinion of the Crowd: Psychological Mechanisms Underlying Public Opinion Perceptions on Social Media*. Vol. 20.
- Nisbett, Richard E. 1996. *Culture of honor: The psychology of violence in the South*. Routledge.
- Noelle-Neumann, E. 1993. *The Spiral of Silence: Public Opinion, Our Social Skin*. University of Chicago Press.
- Parth, M.N. and Shashank Bengali. 2018. “Rumors of Child-kidnapping Gangs and Other WhatsApp Hoaxes are Getting People Killed in India.” *Los Angeles Times* .
URL: <https://www.latimes.com/world/asia/la-fg-india-whatsapp-2018-story.html>
- Ryan, Timothy J. 2012. “What Makes Us Click? Demonstrating Incentives for Angry Discourse with Digital-Age Field Experiments.” *Journal of Politics* 74(4):1138–1152.
- Stenner, Karen. 2005. *The authoritarian dynamic*. Cambridge University Press.
- Stets, Jan E. 2012. “Current emotion research in sociology: Advances in the discipline.” *Emotion Review* 4(3):326–334.
- Straus, Scott. 2007. “What is the relationship between hate radio and violence? Rethinking Rwanda’s radio machete.” *Politics & Society* 35(4):609–637.

- Tetlock, Philip E., Penny S. Visser, Ramadhar Singh, Mark Polifroni, Amanda Scott, Sara Beth Elson, Philip Mazzocco and Phillip Rescober. 2007. "People as intuitive prosecutors: The impact of social-control goals on attributions of responsibility." *Journal of Experimental Social Psychology* 43(2):195–209.
- Tsesis, Alexander. 2002. *Destructive messages: How hate speech paves the way for harmful social movements*. Vol. 778 NYU Press.
- Turner, Jonathan H and Jan E Stets. 2006. "Sociological theories of human emotions." *Annu. Rev. Sociol.* 32:25–52.
- Valentino, Nicholas A, Ted Brader, Eric W Groenendyk, Krysha Gregorowicz and Vincent L Hutchings. 2011. "Election nights alright for fighting: The role of emotions in political participation." *The Journal of Politics* 73(1):156–170.
- Valentino, Nicholas A, Vincent L Hutchings and Ismail K White. 2002. "Cues that matter: How political ads prime racial attitudes during campaigns." *American Political Science Review* 96(1):75–90.
- Viterna, Jocelyn. 2013. *Women in war: The micro-processes of mobilization in El Salvador*. Oxford University Press.
- Viterna, Jocelyn. 2014. "Radical or Righteous? Using Gender to Shape Public Perceptions of Political Violence." *Dynamics of Political Violence: A Process-Oriented Perspective on Radicalization and the Escalation of Political Conflict* pp. 189–216.
- Vosoughi, Soroush, Deb Roy and Sinan Aral. 2018. "The spread of true and false news online." *Science* 359(6380):1146–1151.
- Waldron, Jeremy. 2012. *The harm in hate speech*. Harvard University Press.
- Wood, Elisabeth Jean. 2003. *Insurgent collective action and civil war in El Salvador*. Cambridge University Press.

Yanagizawa-Drott, David. 2014. "Propaganda and conflict: Evidence from the Rwandan genocide." *The Quarterly Journal of Economics* 129(4):1947–1994.

Supporting Information for “The Dangers of False News”

Contents

1	Sampling Procedures	1
1.1	Website Design and Randomization	1
1.2	Exclusion Criteria	2
2	Survey Experiment Vignettes	3
2.1	Article Vignettes	3
3	Measurement Validity Checks	6
4	Attitudes Towards Violence	7
4.1	Judicial Violence (Direct)	8
4.2	Extrajudicial Violence (Direct)	9
4.3	Extrajudicial Violence (Indirect)	10
5	Attitudes Towards Outgroups	11
6	Behavioral Indicators	11
7	Emotional Mediators	12
8	Causal Mediation Analysis	13
9	Open-Ended Responses	14
9.1	Structural Topic Model of Open-Ended Responses	14
9.2	Open-Ended Outcomes by Treatment Group	16
9.3	Example Open-ended Responses by Condition	17
9.4	Example Comments by Condition	18
10	Comments	19
10.1	Comment Text Models	19
10.2	Coding Diagrams	23
11	Intercoder Reliability	24

1 Sampling Procedures

We recruited our respondents through Amazon Mechanical Turk in early December 2018. We advertised the Human Intelligence Task (HIT) as a study on media consumption behavior and compensated respondents 2.11 USD for taking our 15-minute survey, which corresponded to the hourly rate of the federal minimum wage in 2018. After consenting to the take survey, respondents were directed to an external website that hosted an embedded pre-treatment and post-treatment survey (hosted through Qualtrics) as well as the AP-styled news website and article.

1.1 Website Design and Randomization

The interactive survey web app was custom coded using the Python Django web framework and deployed to the web using the Amazon Elastic Beanstalk service. It was designed to be responsive, working on any screen size or device. Click data, comments, and browser metadata were stored in a PostgreSQL database hosted on Amazon Relational Database Services (RDS).

In the web app view, we displayed the informed consent script. When respondents consented and continued to the next view, they were asked to provide their MTurk Worker ID to ensure that they did not fill out the survey twice. They were then asked to share their location to ensure that they were located within the United States. After these first two views, respondents were taken to the pre-treatment survey, followed by the article vignette.

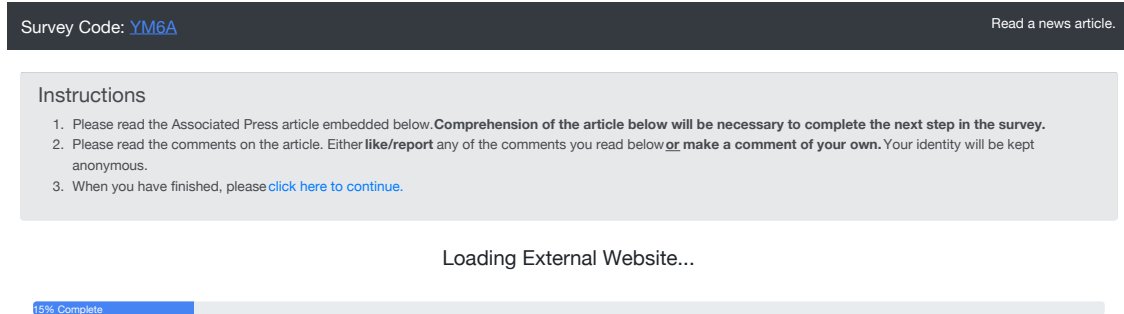
We reverse-engineered the AP website article view for our vignette to maximize its credibility and, thereby, the external validity of our inferences. When the respondent loaded the article vignette, the web app randomly assigned them to one of eight possible treatment combinations which were dynamically displayed in a Django template view. After an interstitial loading page was displayed with a loading bar and a message reading, "Loading External Website...", the article vignette view loaded in an iframe below instructions for this portion of the survey.

Treatments were randomly assigned through the dynamic web-app which was hosted on Amazon Web Services (AWS). The web app assigned each respondent a randomly generated 4-character alphanumeric code. This code was used to match click, comment, and treatment assignment data with the respondent's Qualtrics survey responses. We manually checked survey codes for typographical errors to ensure proper matching.

The comment section was interactive, allowing users to like, report, and comment on the article. The web app tracked comments, likes, reports, and link clicks using AJAX to communicate between the web client

and the PostgreSQL database. Each interaction was saved with the respondent's survey ID to facilitate matching these data with Qualtrics data.

Figure S1: Loading Interstitial for the Website



1.2 Exclusion Criteria

We drew an initial sample of 1879 respondents, of which we excluded 224 according to the following criteria.

Bots and Fraudulent Accounts. Many accounts on the Amazon Mechanical Turk platform are fraudulent and needed to be screened out of our data. We anticipated this and used four methods to screen them out of our data. First, we required users to enter a survey completion code on the Amazon Mechanical Turk platform. This code was displayed to them after completion of the survey. They were also required to provide their MTurk ID within the survey to ensure that their response could be matched to their MTurk account. This limited each MTurk ID to one survey response. Second, because many fraudulent users simultaneously take surveys from multiple MTurk accounts, we collected data on the user's device (screen size, battery level, browser, etc.), location, and IP address to identify users who are completing the survey multiple times from the same machine. Third, to identify automated survey responses from bots, we read every open-ended response and excluded obvious machine-generated text or text scraped from the internet using reference words from the open-ended questions. Finally, we asked a very simple attention check question at the end of the survey. After reviewing failed attention checks, most were from accounts that appeared to be automated. Those who failed the attention check but did not appear to be bots were kept in the sample.

Duplicates. Many subjects took the survey several times due to an error in our screening procedure that was fixed early on. For these subjects we dropped all but the first response from our sample using their MTurk ID, as described above.

Extreme Non-compliers. Respondents who finished the survey in under 7 minutes were excluded, since

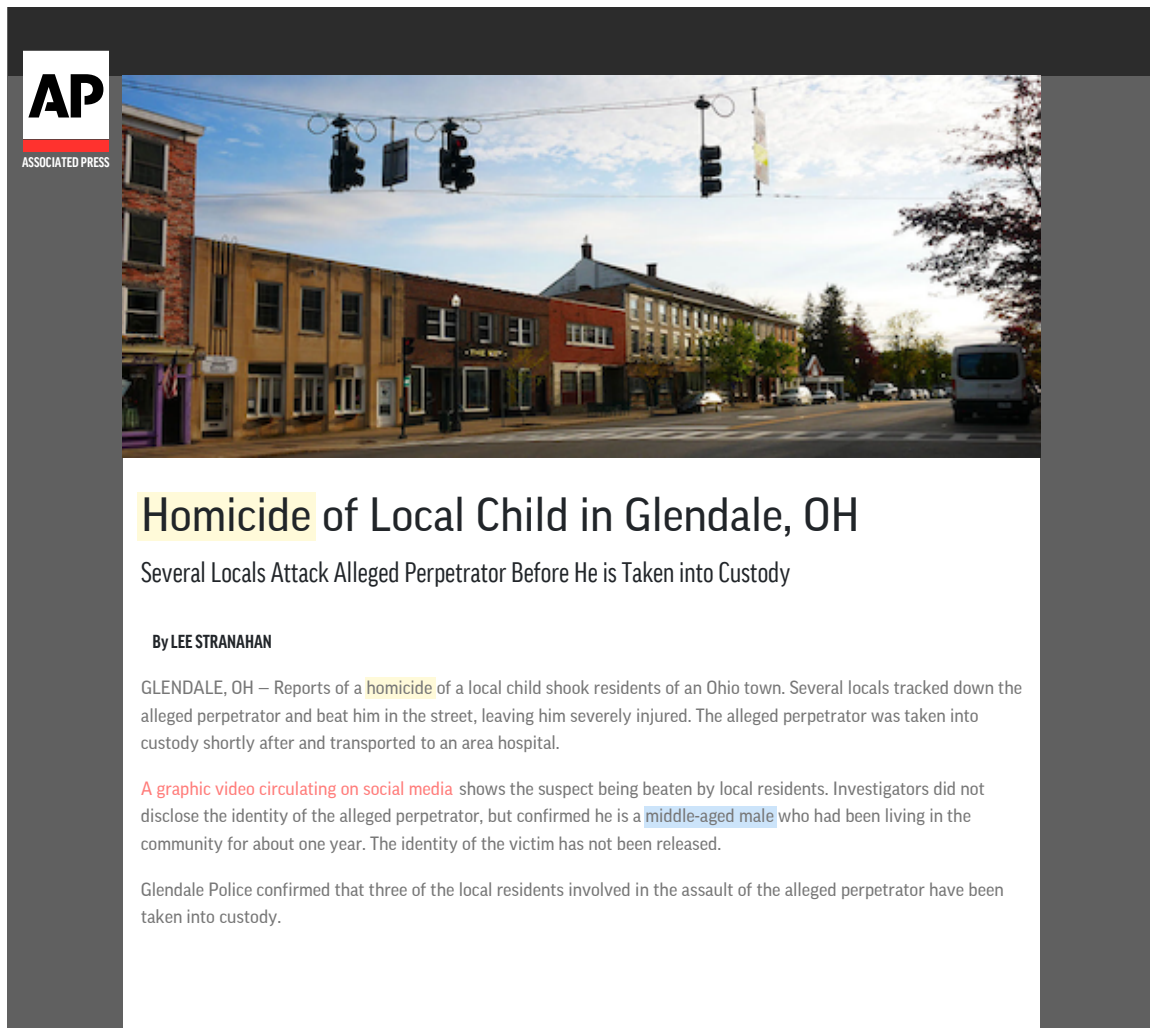
it was not within reasonable expectations that they could have completed the survey within that amount of time.

Non-US Residents. Since our relevant population was US residents, we required respondents to reside in the United States. However, because MTurk's screening criteria is not perfect, several respondents who were from outside the United States took the survey. We screened these respondents by using geolocation and IP address filters and identifying non-English open-ended responses.

2 Survey Experiment Vignettes



2.1 Article Vignettes

Figure S2: Vignette with Objective Content and Unspecified Perpetrator



The image is a screenshot of a news article from the Associated Press. At the top left is the AP logo with 'ASSOCIATED PRESS' written below it. The main image shows a street scene in Glendale, Ohio, with brick buildings, traffic lights, and a white van. Below the image is the article title 'Homicide of Local Child in Glendale, OH' and a subtitle 'Several Locals Attack Alleged Perpetrator Before He is Taken into Custody'. The author is identified as 'By LEE STRANAHAN'. The article text describes a homicide in Glendale, Ohio, where several locals tracked down and beat an alleged perpetrator. A graphic video circulating on social media shows the suspect being beaten. Investigators confirmed the suspect is a middle-aged male who had been living in the community for about one year. The identity of the victim has not been released. Glendale Police confirmed that three of the local residents involved in the assault have been taken into custody.

Figure S3: Vignette with Moral-Emotional Content and Muslim Refugee Perpetrator



Gruesome Slaying of Local Child in Glendale, OH

Several Locals Attack Alleged Perpetrator Before He is Taken into Custody

By **LEE STRANAHAN**

GLENDALE, OH – Reports of a **gruesome slaying** of a local child shook residents of an Ohio town. Several locals tracked down the alleged perpetrator and beat him in the street, leaving him severely injured. The alleged perpetrator was taken into custody shortly after and transported to an area hospital.

A **graphic video circulating on social media** shows the suspect being beaten by local residents. Investigators did not disclose the identity of the alleged perpetrator, but confirmed he is a **middle-aged male muslim refugee** who had been living in the community for about one year. The identity of the victim has not been released.

Glendale Police confirmed that three of the local residents involved in the assault of the alleged perpetrator have been taken into custody.

Figure S4: Comment List with Most Upvoted Violent Comment

3 COMMENTS

Login disabled. Comments will be posted anonymously.

Share your thoughts...

Post

Donut227 48 minutes ago
If he killed my child: I'd have nothing to live for. I'd rain fiery retribution on anyone that killed my child. His end would be brutal.
Like (53) Report

FilmFanic6 43 minutes ago
What could have driven this man to kill a child?
Like (8) Report

Stargazer14 41 minutes ago
I hope that this tragedy can be resolved peacefully. My heart goes out to the victims.
Like (2) Report

Figure S5: Comment List with Most Upvoted Conciliatory Comment

3 COMMENTS

Login disabled. Comments will be posted anonymously.

Share your thoughts...

Post

Stargazer14 41 minutes ago
I hope that this tragedy can be resolved peacefully. My heart goes out to the victims.
Like (53) Report

FilmFanic6 43 minutes ago
What could have driven this man to kill a child?
Like (8) Report

Donut227 48 minutes ago
If he killed my child: I'd have nothing to live for. I'd rain fiery retribution on anyone that killed my child. His end would be brutal.
Like (2) Report

3 Measurement Validity Checks

Figure S6: Correlation of Measures of Support for Violence

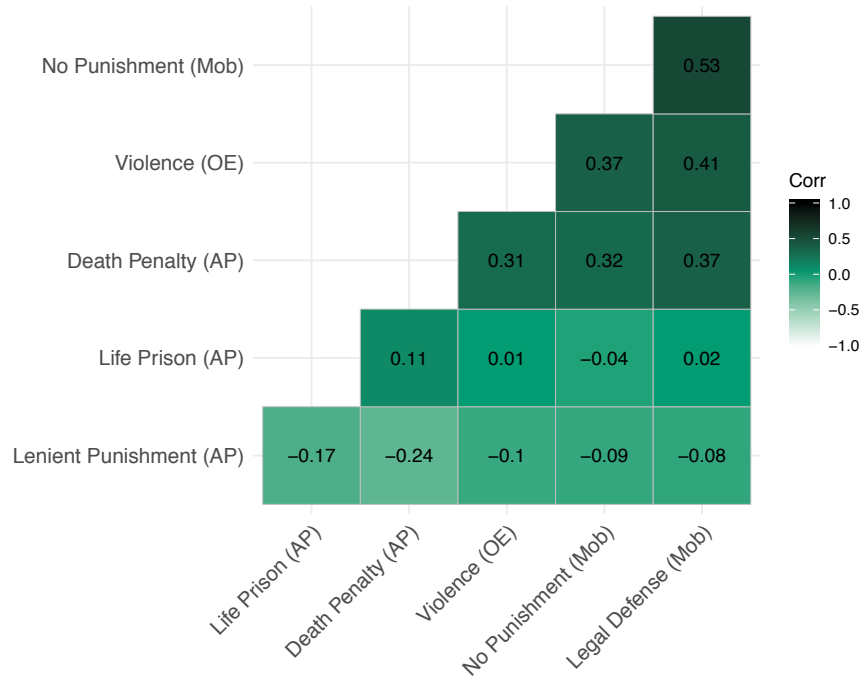


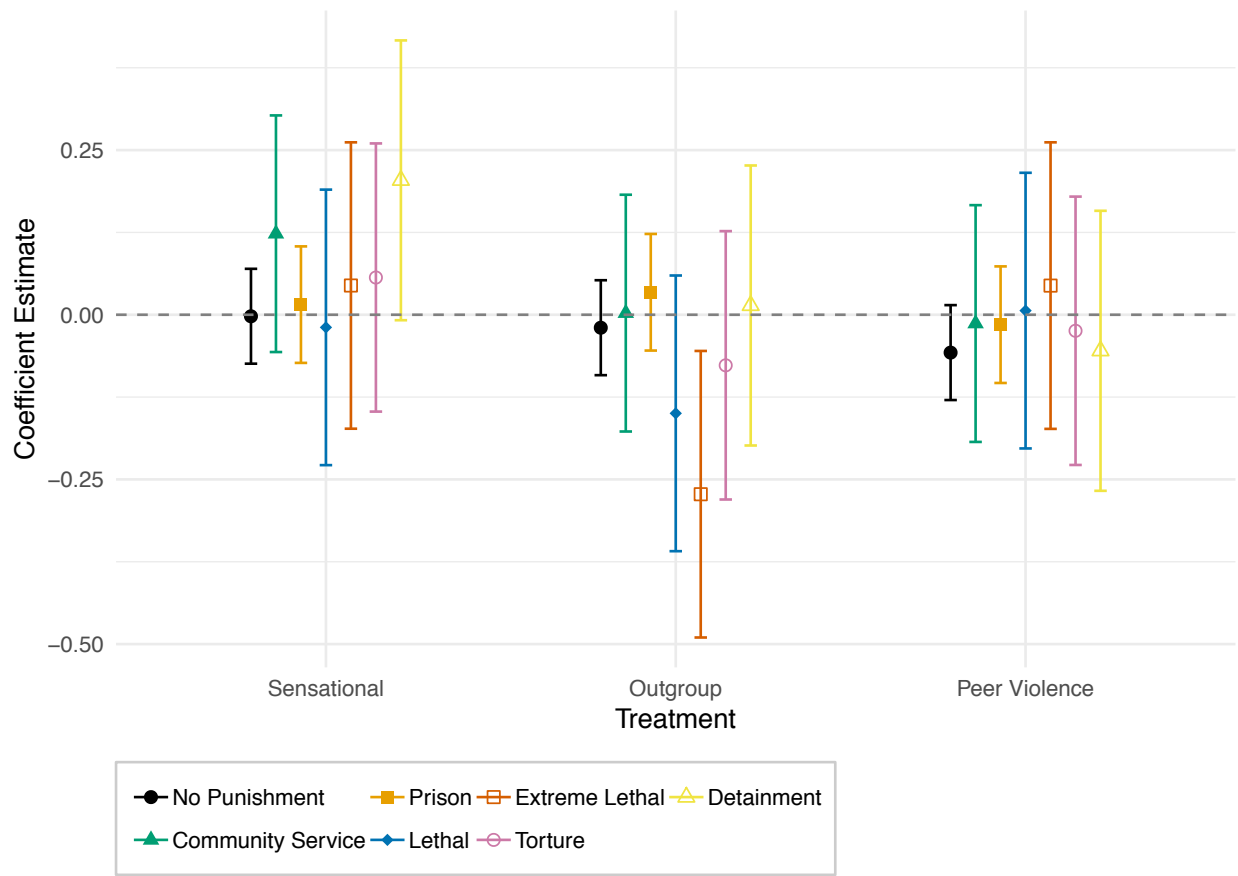
Table S1: Willingness to Donate to Legal Defense Fund for Mob Attackers and Expressed Support for Extrajudicial Violence in Open-ended Responses

	<i>Dependent variable:</i>
	Extrajudicial
Legal Defense Fund	0.118*** (0.005)
Observations	1,602

Note: *p<0.1; **p<0.05; ***p<0.01

4 Attitudes Towards Violence

Figure S7: Support for Judicial and Extrajudicial Punishment by Treatment Condition (Direct Measures)



4.1 Judicial Violence (Direct)

Table S2: Attitudes Toward Judicial Punishment of the Alleged Perpetrator

	<i>Dependent variable:</i>							
	Lenient	Lenient	Prison	Prison	Lethal	Lethal	Extreme Lethal	Extreme Lethal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sensational	0.060 (0.053)	0.058 (0.053)	0.015 (0.045)	0.006 (0.045)	-0.019 (0.107)	0.019 (0.099)	0.044 (0.111)	0.099 (0.102)
Outgroup	-0.009 (0.053)	0.233 (0.179)	0.034 (0.045)	0.092 (0.161)	-0.150 (0.107)	-0.222 (0.378)	-0.272** (0.111)	-0.487 (0.370)
Peer Violence	-0.035 (0.053)	-0.039 (0.053)	-0.015 (0.045)	-0.013 (0.045)	0.006 (0.107)	0.019 (0.098)	0.044 (0.111)	0.044 (0.102)
Auth. (Log)		-0.043 (0.063)		-0.076 (0.053)		0.513*** (0.122)		0.466*** (0.127)
Ethno. (Log)		0.003 (0.003)		-0.003 (0.002)		0.007 (0.004)		0.013*** (0.004)
Symb. Racism (Log)		-0.045 (0.119)		-0.117 (0.094)		1.053*** (0.218)		1.143*** (0.228)
South		0.088 (0.056)		-0.017 (0.046)		0.110 (0.102)		0.220** (0.107)
Republican		-0.192*** (0.067)		-0.077 (0.059)		0.675*** (0.111)		0.577*** (0.131)
Outgroup x Auth. (Log)		-0.003 (0.086)		0.125 (0.077)		0.025 (0.171)		-0.079 (0.177)
Outgroup x Ethno (Log)		0.002 (0.003)		0.001 (0.003)		-0.006 (0.006)		0.003 (0.006)
Outgroup x Symb. Racism (Log)		-0.238 (0.153)		-0.147 (0.140)		0.175 (0.302)		0.133 (0.315)
Observations	1,629		1,629		1,629		1,629	

Note:

*p<0.1; **p<0.05; ***p<0.01

4.2 Extrajudicial Violence (Direct)

Table S3: Direct Questions About Extrajudicial Punishment of the Alleged Perpetrator

	<i>Dependent variable:</i>			
	Torture	Torture	Detainment	Detainment
	(1)	(2)	(3)	(4)
Sensational	0.057 (0.104)	0.083 (0.097)	0.204* (0.108)	0.224** (0.104)
Outgroup	-0.077 (0.104)	0.031 (0.331)	0.014 (0.108)	-0.314 (0.357)
Peer Violence	-0.024 (0.104)	-0.018 (0.097)	-0.055 (0.108)	-0.041 (0.104)
Auth. (Log)		0.659*** (0.121)		0.680*** (0.127)
Ethno. (Log)		0.016*** (0.004)		0.006 (0.004)
Symb. Racism (Log)		0.833*** (0.225)		0.381 (0.232)
South		0.268*** (0.101)		0.230** (0.108)
Republican		0.100 (0.130)		0.241* (0.137)
Outgroup x Auth. (Log)		0.127 (0.170)		0.043 (0.180)
Outgroup x Ethno (Log)		-0.0001 (0.005)		0.005 (0.006)
Outgroup x Symb. Racism (Log)		-0.169 (0.307)		0.093 (0.321)
Observations	1,629		1,629	

Note:

*p<0.1; **p<0.05; ***p<0.01

4.3 Extrajudicial Violence (Indirect)

Table S4: Indirect Questions About Extrajudicial Punishment of the Alleged Perpetrator

	<i>Dependent variable:</i>					
	No Punish.	No Punish.	Comm. Service	Comm. Service	Leg. Defense	Leg. Defense
	(1)	(2)	(3)	(4)	(5)	(6)
Sensational	0.164 (0.111)	0.203* (0.108)	0.052 (0.104)	0.054 (0.104)	0.217** (0.099)	0.250*** (0.096)
Outgroup	-0.285** (0.111)	-1.005*** (0.378)	-0.205** (0.104)	0.067 (0.379)	-0.083 (0.099)	-0.842** (0.337)
Peer Violence	0.152 (0.111)	0.145 (0.107)	0.168 (0.104)	0.166 (0.104)	0.024 (0.099)	0.027 (0.095)
Auth. (Log)		0.314** (0.137)		-0.188 (0.134)		0.434*** (0.126)
Ethno. (Log)		0.011** (0.004)		-0.003 (0.004)		0.005 (0.004)
Symb. Racism (Log)		0.349 (0.247)		0.034 (0.239)		0.280 (0.221)
South		0.204* (0.112)		-0.040 (0.109)		0.234** (0.099)
Republican		0.472*** (0.142)		0.224* (0.132)		0.372*** (0.123)
Outgroup x Auth. (Log)		-0.079 (0.188)		-0.253 (0.185)		0.042 (0.169)
Outgroup x Ethno (Log)		0.003 (0.006)		-0.003 (0.006)		0.003 (0.006)
Outgroup x Symb. Racism (Log)		0.535 (0.339)		-0.003 (0.335)		0.495* (0.301)
Observations	1,615		1,615		1,615	

Note:

*p<0.1; **p<0.05; ***p<0.01

5 Attitudes Towards Outgroups

Table S5: OLS Models for Attitudes Toward Outgroups

	<i>Dependent variable:</i>							
	Refugees (1)	Refugees (2)	Muslim Immigration (3)	Muslim Immigration (4)	Monitoring (5)	Monitoring (6)	Registry (7)	Registry (8)
Sensational	0.173* (0.094)	0.103 (0.069)	0.066 (0.086)	0.004 (0.067)	-0.077 (0.086)	-0.032 (0.073)	-0.008 (0.103)	0.035 (0.084)
Outgroup	-0.148 (0.094)	-0.444* (0.238)	-0.200** (0.086)	-0.577** (0.229)	0.181** (0.086)	0.065 (0.257)	0.193* (0.103)	0.328 (0.269)
Peer Violence	-0.005 (0.094)	0.0003 (0.069)	0.009 (0.086)	0.021 (0.066)	0.090 (0.086)	0.084 (0.072)	0.091 (0.103)	0.098 (0.083)
Auth. (Log)		-0.408*** (0.087)		-0.377*** (0.082)		0.472*** (0.090)		0.864*** (0.107)
Ethno. (Log)		-0.016*** (0.003)		-0.019*** (0.003)		0.016*** (0.003)		0.020*** (0.004)
Symb. Racism (Log)		-2.051*** (0.169)		-1.791*** (0.159)		1.289*** (0.173)		1.371*** (0.193)
South		-0.135* (0.072)		-0.155** (0.069)		0.122 (0.076)		0.292*** (0.089)
Republican		-1.089*** (0.088)		-0.792*** (0.085)		0.432*** (0.094)		0.529*** (0.116)
Outgroup x Auth. (Log)		-0.011 (0.123)		0.063 (0.114)		0.161 (0.126)		0.250* (0.151)
Outgroup x Ethno (Log)		-0.001 (0.004)		-0.004 (0.004)		0.002 (0.005)		-0.001 (0.005)
Outgroup x Symb. Racism (Log)		0.314 (0.226)		0.408* (0.212)		-0.081 (0.238)		-0.249 (0.271)
Observations	1,613		1,612		1,612		1,612	

Note:

*p<0.1; **p<0.05; ***p<0.01

6 Behavioral Indicators

Table S6: OLS Models for Behavioral Indicators

	<i>Dependent variable:</i>							
	Liked Viol. (1)	Liked Viol. (2)	Liked Non-viol. (3)	Liked Non-viol. (4)	Reported Viol. (5)	Reported Viol. (6)	Wrote Comment (7)	Wrote Comment (8)
Sensational	0.004 (0.025)	0.011 (0.024)	-0.025 (0.021)	-0.028 (0.021)	-0.018 (0.024)	-0.016 (0.024)	0.045** (0.023)	0.050** (0.023)
Outgroup	-0.014 (0.025)	-0.044 (0.084)	-0.013 (0.021)	0.054 (0.072)	0.034 (0.024)	0.068 (0.082)	-0.006 (0.023)	-0.085 (0.080)
Peer Violence	0.049** (0.025)	0.050** (0.024)	-0.096*** (0.021)	-0.095*** (0.021)	-0.034 (0.024)	-0.035 (0.024)	0.027 (0.023)	0.026 (0.023)
Auth. (Log)		0.042 (0.029)		-0.015 (0.025)		-0.022 (0.029)		-0.002 (0.028)
Ethno. (Log)		-0.00002 (0.001)		-0.002** (0.001)		-0.001 (0.001)		-0.001 (0.001)
Symb. Racism (Log)		0.213*** (0.051)		-0.012 (0.043)		0.100* (0.051)		0.059 (0.049)
South		0.021 (0.025)		-0.013 (0.022)		-0.014 (0.024)		0.007 (0.024)
Republican		0.039 (0.030)		-0.082*** (0.027)		-0.026 (0.030)		0.006 (0.029)
Outgroup x Auth. (Log)		0.079* (0.041)		0.001 (0.036)		0.038 (0.040)		-0.005 (0.038)
Outgroup x Ethno (Log)		-0.0003 (0.001)		-0.001 (0.001)		-0.0002 (0.001)		0.001 (0.001)
Outgroup x Symb. Racism (Log)		-0.014 (0.071)		-0.031 (0.063)		-0.045 (0.072)		0.027 (0.068)
Observations	1,652		1,652		1,652		1,652	

Note:

*p<0.1; **p<0.05; ***p<0.01

7 Emotional Mediators

Table S7: Emotional Responses by Treatment Condition

	<i>Dependent variable:</i>							
	Anger (1)	Anger (2)	Disgust (3)	Disgust (4)	Fear (5)	Fear (6)	Anxiety (7)	Anxiety (8)
Sensational	0.121* (0.063)	0.121* (0.062)	0.078 (0.059)	0.083 (0.059)	0.126** (0.062)	0.115* (0.062)	0.137** (0.062)	0.134** (0.062)
Outgroup	-0.061 (0.063)	-0.155 (0.222)	-0.071 (0.059)	-0.346 (0.211)	0.081 (0.062)	0.016 (0.232)	0.019 (0.062)	-0.145 (0.234)
Peer Violence	0.037 (0.063)	0.051 (0.062)	-0.038 (0.059)	-0.029 (0.059)	0.002 (0.062)	0.008 (0.062)	0.008 (0.062)	0.013 (0.062)
Auth. (Log)		0.311*** (0.074)		0.166** (0.071)		0.214*** (0.075)		0.166** (0.075)
Ethno. (Log)		-0.009*** (0.003)		-0.010*** (0.002)		-0.001 (0.002)		-0.004 (0.003)
Symb. Racism (Log)		-0.078 (0.136)		0.007 (0.132)		-0.165 (0.138)		-0.149 (0.136)
South		0.095 (0.064)		0.069 (0.061)		-0.058 (0.065)		0.003 (0.065)
Republican		0.278*** (0.076)		0.153** (0.072)		0.074 (0.080)		0.063 (0.079)
Outgroup x Auth. (Log)		-0.102 (0.105)		0.037 (0.099)		0.0003 (0.107)		-0.061 (0.107)
Outgroup x Ethno (Log)		0.009** (0.004)		0.008** (0.004)		0.004 (0.004)		0.006 (0.004)
Outgroup x Symb. Racism (Log)		-0.114 (0.192)		-0.027 (0.180)		-0.063 (0.191)		0.001 (0.194)
Observations	1,647		1,647		1,647		1,647	

Note:

*p<0.1; **p<0.05; ***p<0.01

8 Causal Mediation Analysis

Figure S8: Causal Mediation Analysis for Anger

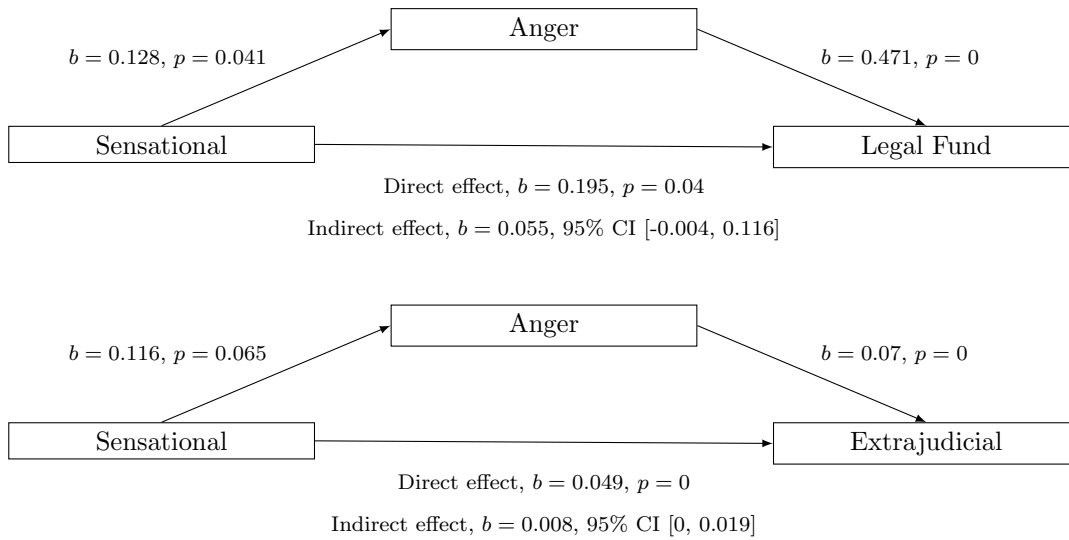


Figure S9: Causal Mediation Analysis for Anxiety

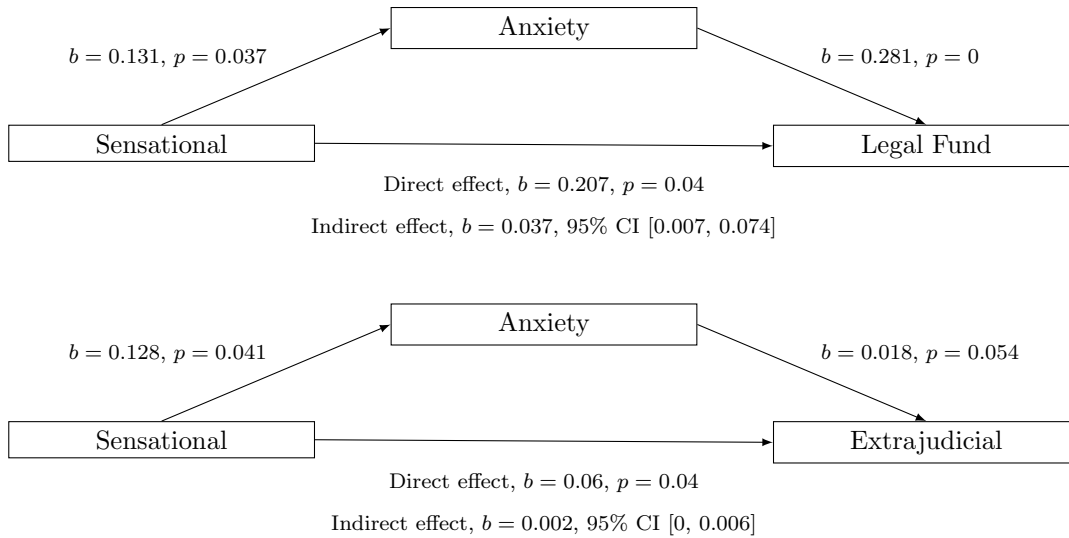
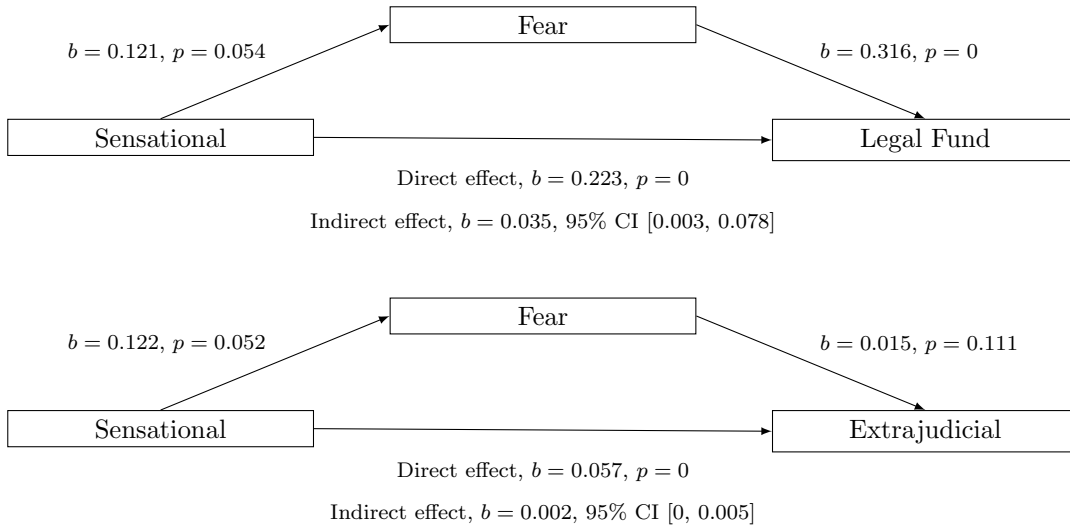


Figure S10: Causal Mediation Analysis for Fear



9 Open-Ended Responses

9.1 Structural Topic Model of Open-Ended Responses

Figure S11: Topic Probabilities

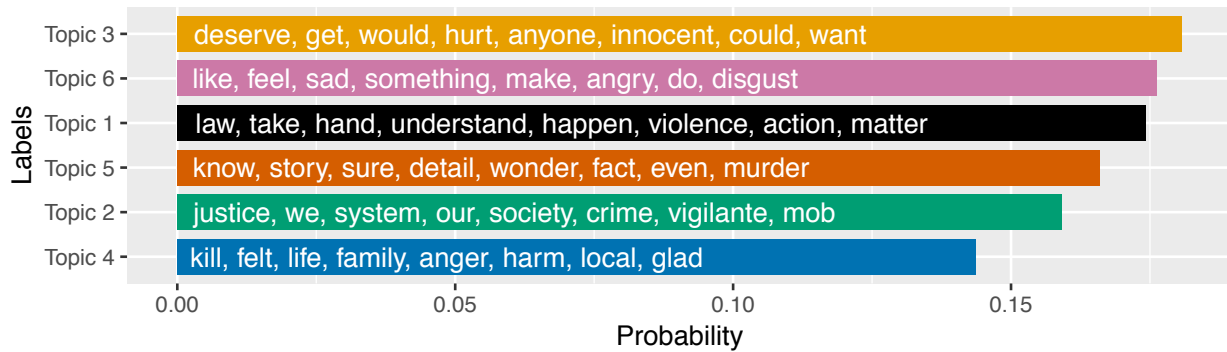


Table S8: Most Representative Responses by Topic

Topic	Text	Peer	Sens.	Out.
Topic 1	The loss of a child, especially by violence, is unthinkable. I understand the frustration of the locals but vigilante justice is equally unthinkable. taking the law into one's own hands is despicable. Vigilante justice is never warranted and is horrible. It is outlawed for a reason.	Yes	No	No
Topic 2	Child murderers are some of the worst scum this world can produce. Murderes of any kind are terrible but people that prey on the weakest and most vulnerable aspects of our society are at the bottom of the sludge pit. This man needs to be tried and convicted. A frothing mob cannot be expected to make rational choices and lynching someone in the street is garbage we used to do hundreds of years ago and garbage that third world hellholes continue to do today. I understand being enraged but let the police do their work.	No	No	No
Topic 3	the man got what he deserved. he should get more than that though. anyone who hurts a child should get the strictest punishment he got what he deserved. no one should hurt an innocent defenseless child	Yes	No	Yes
Topic 4	I felt anger and fear. I felt extreme sadness for the victim and the victim's family. I feel confusion about something like this could happen. I felt the attack of the perpetrator were justified. It was only a small step of vindication that the victim and the victim's family deserve.	No	Yes	Yes
Topic 5	As noted in my comment regarding the article provided, there is far too little detail to draw any conclusions regarding the incident other than it is profoundly horrific and sad. I have many questions to be answered to even begin to form a cohesive assessment of this tragedy. Again, there is not enough information to draw an intelligent conclusion. One can only presume that the emotional reaction of the bystanders was reactionary and vengeful due to the immediacy of the event.	Yes	Yes	No
Topic 6	I feel like people should not be driven to these extremes. It is a little ridiculous that humans act like this. I hope everyone begins to chill more I feel like he is someone to be punished but if you are to do the same thing as he doing you are no better. To become what others are negatively will ultimately be the same fate.	No	No	Yes

9.2 Open-Ended Outcomes by Treatment Group

Table S9: OLS Model for Expressed Support for Violence, Extrajudicial Violence, and Lethal Violence in Open-ended Responses

	<i>Dependent variable:</i>			
	Violence (1)	Violence (2)	Extrajudicial Violence (3)	Extrajudicial Violence (4)
Sensational	0.057** (0.025)	0.063*** (0.024)	0.052** (0.024)	0.057** (0.023)
Outgroup	-0.075*** (0.025)	-0.192** (0.084)	-0.069*** (0.024)	-0.148* (0.080)
Peer Violence	-0.036 (0.025)	-0.035 (0.024)	-0.009 (0.024)	-0.009 (0.023)
Auth. (Log)		0.048 (0.029)		0.038 (0.030)
Ethno. (Log)		0.002** (0.001)		0.003*** (0.001)
Symb. Racism (Log)		0.053 (0.052)		0.061 (0.053)
South		0.051** (0.025)		0.062** (0.025)
Republican		0.073** (0.030)		0.088*** (0.031)
Outgroup x Auth. (Log)		0.015 (0.041)		0.047 (0.041)
Outgroup x Ethno (Log)		-0.001 (0.001)		-0.0005 (0.001)
Outgroup x Symb. Racism (Log)		0.110 (0.072)		0.050 (0.071)
Observations	1,613		1,613	

Note:

*p<0.1; **p<0.05; ***p<0.01

9.3 Example Open-ended Responses by Condition

1. **Extrajudicial Violence and Lethal Violence (Outgroup Baseline, Sensational Treatment, Peer-Violence Treatment):** “I am absolutely disgusted, appalled, and at a loss for words about this attack. They should kill him in the street. Let everyone who wants a piece of him have a piece of him.”
2. **Lethal Violence (Outgroup Treatment, Sensational Treatment, Peer-Violence Baseline):**
“I think the perpetrator got what he deserved. It is a shame that they didn’t kill him the same way he killed the child. The attackers shouldn’t be punished for attacking the perpetrator.”
3. **Judicial Violence (Outgroup Treatment, Sensational Treatment, Peer-Violence Baseline):**
“Horrible that anyone would want to harm or kill a child. Worst of the worst in society. They should be punished to the fullest extent by the laws.”
4. **Information-Seeking (Outgroup Treatment, Sensational Baseline, Peer Violence Baseline):** “I want to know what happened. I want to know if the person that was beat up is the one accused of the murder. I want to know the details”
5. **Information-Seeking (Outgroup Baseline, Sensational Baseline, Peer-Violence Treatment):**
“I wanted to know more behind it. My reaction was what happened. The article only stated that a homicide had occurred to a child and that the victim’s identity as well as the perpetrator’s identity hadn’t been released, but there were not circumstances stated around the incident.”
6. **Cognitive Dissonance (Outgroup Treatment, Sensational Treatment, Peer-Violence Treatment):** “It’s weird because just not knowing the details, just gut reaction I feel good that the mofo got what he deserved but then I start thinking more and I really don’t know any details and the alleged perpetrator has not been convicted. I’m conflicted”
7. **Racism (Outgroup Treatment, Sensational Control, Peer-Violence Treatment):** “I wonder if, because he was Muslim, he was unjustly targeted. There didn’t seem to be any proof yet of who was responsible. I never agree with using violence to address a violent act.”
8. **Cognitive Dissonance and Racism (Outgroup Treatment, Sensational Treatment, Peer-Violence Treatment):** “I was originally thinking ”good for them”. Then I thought ”what if that dude didn’t do it and they angry mob just beat up a muslim. Then I felt remorse for the whole situation. ”

9. **Surprise at Lack of Emotion in Article (Outgroup Treatment, Sensational Baseline, Peer-Violence Baseline):** “That’s one of the most unemotionally written articles I’ve come across in quiet a while. Most articles are written in a way to elicit feelings of emotion from it’s audience. The behavior of the accused is deplorable but the author didn’t present it that way. It seemed to be more focused on the attack against the accused.”

9.4 Example Comments by Condition

1. **Judicial Violence (Outgroup Treatment, Sensational Baseline, Peer-Violence Treatment):**
“This middle-aged male Muslim refugee who had been living in the community for about one year should be sent back to where he came from and let them deal with him. This person is not going to last very long in prison.”
2. **Judicial and Lethal Violence (Outgroup Baseline, Sensational Baseline, Peer-Violence Treatment):** “I think this person should get death penalty for killing a child.”
3. **Lethal Violence (Outgroup Baseline, Sensational Treatment, Peer-Violence Baseline):**
“They should have killed him right then and there. Hopefully they put him in general population and do not waste tax money protecting him from other inmates.”
4. **Racism (Outgroup Treatment, Sensational Treatment, Peer-Violence Treatment):** “This is clearly a hate crime. There’s no mention of evidence, since of course they didn’t need any to savagely beat a Muslim refugee. This is just more right-wing paranoia, racism, and violence.”
5. **Concerns with Vigilantism (Outgroup Baseline, Sensational Baseline, Peer-Violence Treatment)** “We should be wary about rushing to judgment. The murder of a child is a most heinous crime, but unless someone actually witnessed it happen or has incontrovertible proof it is possible that the suspect is actually innocent. If that is the case then it would turn out that the local residents beat up an innocent man, while the real murderer is out there.”
6. **Demonizing (Outgroup Baseline, Sensational Treatment, Peer-Violence Treatment)** “What is wrong with him? Who can be so cruel!”
7. **Dehumanizing and Extrajudicial Violence (Outgroup Baseline, Sensational Baseline, Peer-Violence Treatment)** “Vigilante justice is the best kind of justice for monsters like this perpetrator!”

10 Comments

10.1 Comment Text Models

Table S10: OLS Model for Propensity to Post a Comment

	<i>Dependent variable:</i>	
	Posted a Comment	Posted a Comment
	(1)	(2)
Sensational	0.045** (0.023)	0.050** (0.023)
Outgroup	-0.006 (0.023)	-0.085 (0.080)
Peer Violence	0.027 (0.023)	0.026 (0.023)
Auth. (Log)		-0.002 (0.028)
Ethno. (Log)		-0.001 (0.001)
Symb. Racism (Log)		0.059 (0.049)
South		0.007 (0.024)
Republican		0.006 (0.029)
Outgroup x Auth. (Log)		-0.005 (0.038)
Outgroup x Ethno (Log)		0.001 (0.001)
Outgroup x Symb. Racism (Log)		0.027 (0.068)
Observations	1,652	
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	

Table S11: Logit Model for Propensity to Post a Comment

	<i>Dependent variable:</i> Posted a Comment
Gender	0.122 (0.108)
Age	0.009* (0.005)
Education	0.042 (0.041)
Authoritarian (Log)	0.014 (0.094)
Ethnocentrism (Log)	-0.001 (0.003)
Symbolic Racism (Log)	0.299* (0.175)
South	0.022 (0.111)
Republican	-0.002 (0.135)
Observations	1,655
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Table S12: Heckman 2-Step Selection Model OLS Model for Expressed Support for Violence, Extrajudicial Violence, and Lethal Violence in Comments

	<i>Dependent variable:</i> Violence
Sensational	0.047 (0.037)
Outgroup	-0.063* (0.036)
Peer Violence	0.082** (0.037)
Observations	1,650
ρ	-0.011
Inverse Mills Ratio	-0.004 (0.238)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Table S13: Additional Text Responses (Comments)

	<i>Dependent variable:</i>							
	Cognitive Dissonance (1)	Cognitive Dissonance (2)	Info. Seeking (3)	Info. Seeking (4)	Fake News (5)	Fake News (6)	Racism (7)	Racism (8)
Sensational	0.022 (0.019)	0.022 (0.021)	-0.042 (0.026)	-0.044* (0.026)	-0.009 (0.006)	-0.009 (0.006)	-0.006 (0.016)	-0.012 (0.015)
Outgroup	-0.007 (0.019)	-0.005 (0.072)	0.009 (0.026)	-0.050 (0.096)	-0.007 (0.006)	-0.008 (0.026)	0.056*** (0.016)	0.212*** (0.066)
Peer Violence	0.032* (0.019)	0.033* (0.020)	-0.034 (0.026)	-0.034 (0.026)	0.007 (0.006)	0.007 (0.005)	-0.023 (0.016)	-0.019 (0.015)
Auth. (Log)		-0.011 (0.026)		-0.028 (0.028)		0.005 (0.004)		0.011 (0.010)
Ethno. (Log)		0.0004 (0.0005)		0.001 (0.001)		0.0003 (0.001)		-0.0001 (0.0001)
Symb. Racism (Log)		-0.008 (0.032)		-0.086 (0.057)		-0.019 (0.014)		0.018 (0.014)
South		0.031 (0.021)		-0.055** (0.025)		-0.005 (0.004)		-0.002 (0.016)
Republican		-0.006 (0.024)		0.035 (0.032)		0.007 (0.007)		-0.021 (0.016)
Outgroup x Auth. (Log)		0.006 (0.034)		0.038 (0.043)		-0.005 (0.004)		-0.018 (0.024)
Outgroup x Ethno (Log)		-0.001 (0.001)		0.0002 (0.001)		-0.0003 (0.001)		-0.0003 (0.001)
Outgroup x Symb. Racism (Log)		0.020 (0.061)		0.024 (0.077)		0.013 (0.010)		-0.103*** (0.039)
Observations	498		498		498		498	

Note: *p<0.1; **p<0.05; ***p<0.01

Table S14: Additional Text Responses (Open Ended)

	<i>Dependent variable:</i>									
	Cognitive Dissonance (1)	Cognitive Dissonance (2)	Info. Seeking (3)	Info. Seeking (4)	Fake News (5)	Fake News (6)	Racism (7)	Racism (8)	Social Desirability (9)	Social Desirability (10)
Sensational	0.046** (0.019)	0.046** (0.019)	0.011 (0.015)	0.008 (0.015)	0.006 (0.006)	0.006 (0.006)	-0.007 (0.011)	-0.010 (0.011)	-0.001 (0.006)	-0.001 (0.006)
Outgroup	-0.036* (0.019)	-0.066 (0.069)	0.020 (0.015)	0.011 (0.055)	-0.003 (0.006)	-0.032** (0.016)	0.080*** (0.011)	0.169*** (0.044)	0.031*** (0.006)	0.028 (0.018)
Peer Violence	0.004 (0.019)	0.003 (0.019)	0.002 (0.015)	0.001 (0.015)	0.003 (0.006)	0.003 (0.006)	-0.025** (0.011)	-0.025** (0.011)	-0.003 (0.006)	-0.002 (0.006)
Auth. (Log)		-0.047* (0.024)		-0.049*** (0.017)		-0.017** (0.007)		-0.004 (0.007)		-0.001 (0.001)
Ethno. (Log)		0.001 (0.001)		0.001* (0.001)		0.0004* (0.0002)		-0.0002 (0.0003)		0.0001 (0.0001)
Symb. Racism (Log)		-0.037 (0.043)		-0.057* (0.031)		-0.009 (0.010)		-0.017 (0.016)		-0.002 (0.004)
South		0.007 (0.020)		-0.028* (0.015)		-0.004 (0.006)		-0.020* (0.011)		0.002 (0.006)
Republican		-0.0004 (0.024)		0.004 (0.018)		-0.001 (0.007)		0.002 (0.012)		0.005 (0.009)
Outgroup x Auth. (Log)		0.042 (0.032)		0.022 (0.025)		0.016* (0.008)		-0.046** (0.019)		0.011 (0.011)
Outgroup x Ethno (Log)		-0.001 (0.001)		-0.001 (0.001)		-0.0003 (0.0003)		0.0002 (0.001)		-0.0001 (0.0003)
Outgroup x Symb. Racism (Log)		0.016 (0.056)		0.032 (0.045)		0.022 (0.015)		-0.052 (0.034)		-0.001 (0.015)
Observations	1,623		1,623		1,623		1,623		1,623	

Note: *p<0.1; **p<0.05; ***p<0.01

Figure S12: Additional Comment Outcomes by Treatment Condition

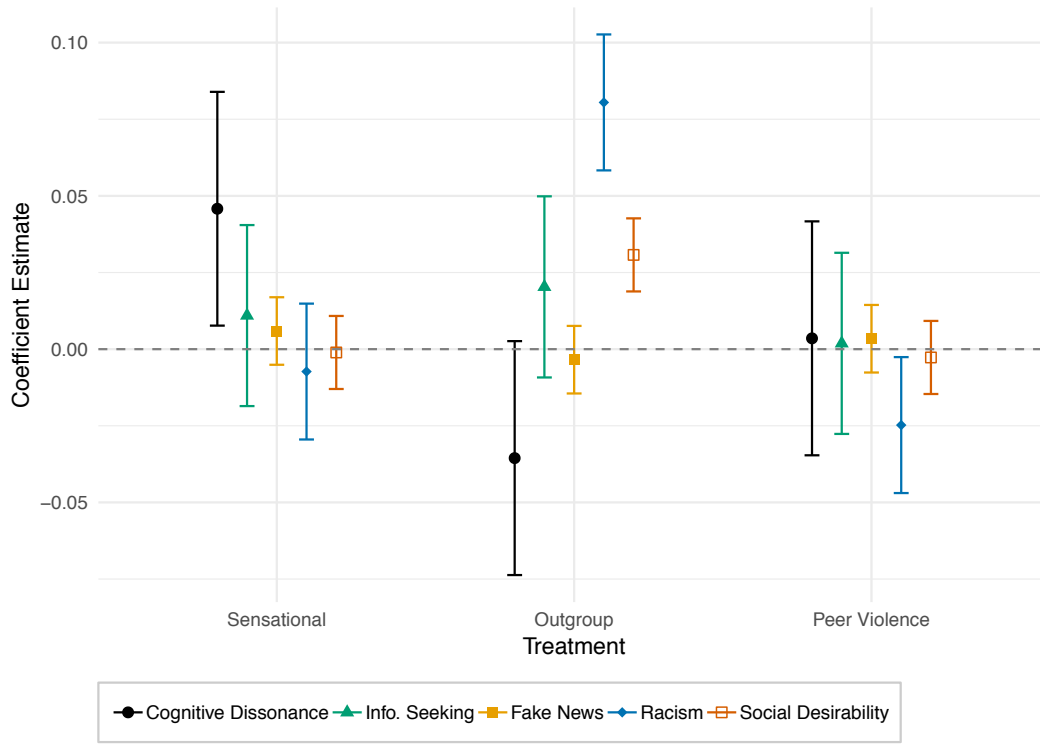
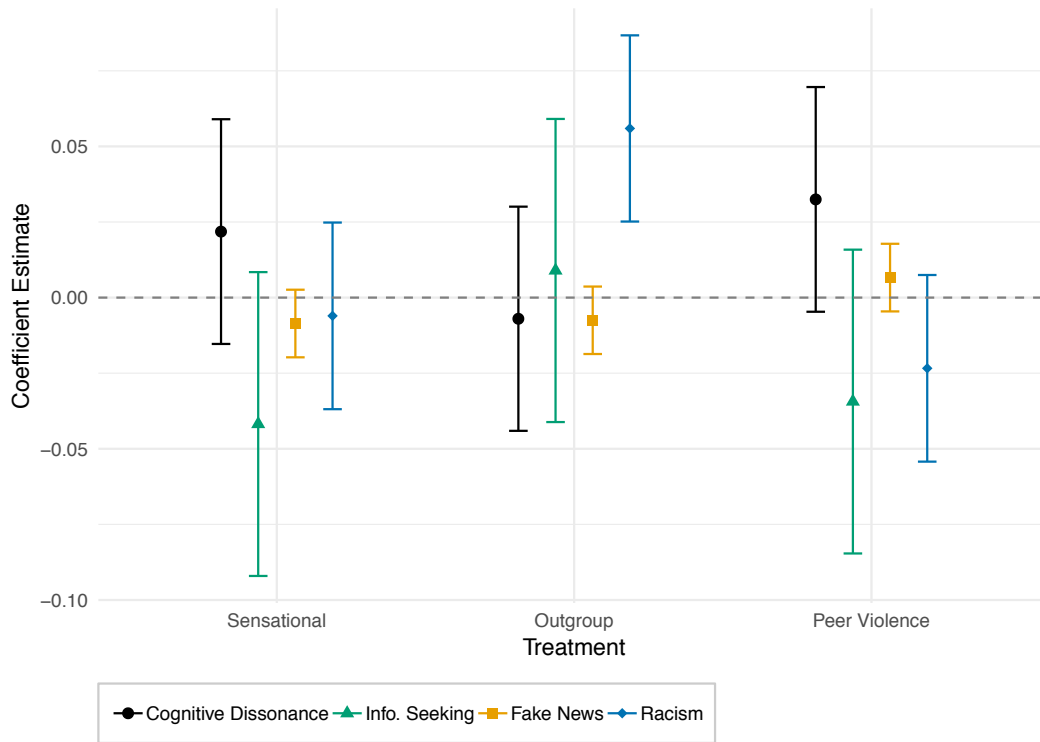


Figure S13: Additional Open-Ended Outcomes by Treatment Condition



10.2 Coding Diagrams

Figure S14: Violence Coding Diagram

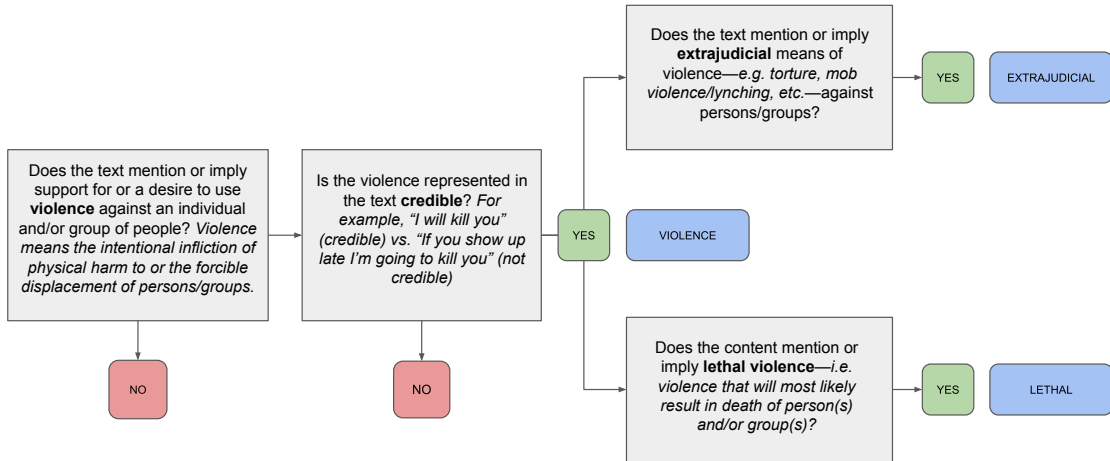


Figure S15: Racism and Social Desirability Coding Diagrams

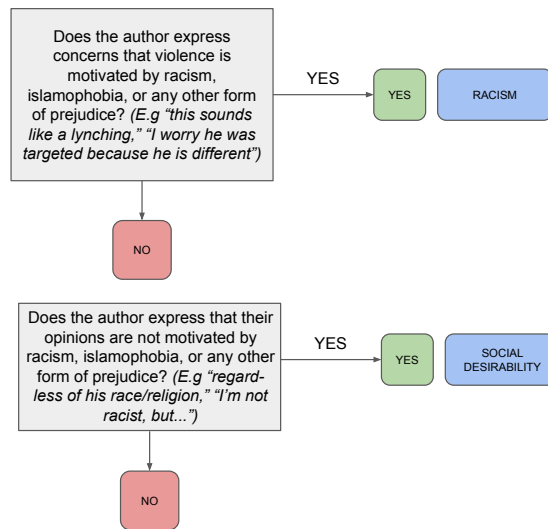


Figure S16: Information-Seeking and Fake News Coding Diagrams

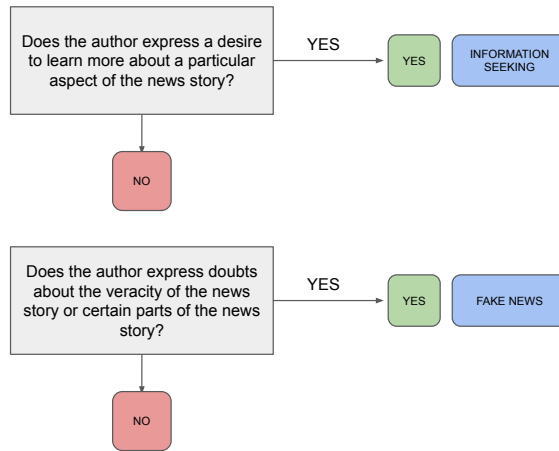
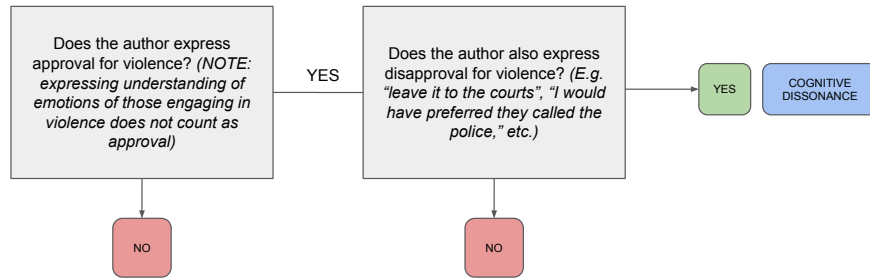


Figure S17: Cognitive Dissonance Coding Diagrams



11 Intercoder Reliability

Table S15: Intercoder Reliability Measures

Category	Kappa	F1
Cognitive Dissonance	0.80	0.96
Fake News	1.00	1.00
Info. Seeking	0.71	0.93
Racism/Hate	0.78	0.97
Social Desirability	0.74	0.99
Violence	0.81	0.90
Lethal Violence	0.98	1.00
Extrajudicial Violence	0.89	0.95