Group Processes & Intergroup Relations
GPIR

Article

Reducing Islamophobia: An assessment of psychological mechanisms that underlie anti-Islamophobia media interventions

Group Processes & Intergroup Relations
1–24

DOI: 10.1177/13684302221085832
journals.sagepub.com/home/gpi
SAGE

Samantha L. Moore-Berg,[1,2] Boaz Hameiri,[3] Emily Falk[1] and Emile Bruneau[1,2]

Abstract

Western countries have witnessed increased hostility towards Muslims among individuals, and structurally in the ways that the media covers stories related to Islam/Muslims and in policies that infringe on the rights of Muslim communities. In response, practitioners have created media interventions that aim to reduce Islamophobia. However, it is unclear what causal effects these interventions have on reducing Islamophobia. Here, we test the effects of 11 media interventions developed by practitioners with an intervention tournament among U.S. samples. In Study 1, we identified three videos that most effectively reduced Islamophobia both immediately after watching and 1 month later. In Studies 2–4, we examined the psychological mechanisms of these successful videos and found an indirect effect of the interventions on reduced support for anti-Muslim policies through recognition of media bias against Muslims. This research highlights that drawing attention to structural biases, including biased media coverage of Muslims, is one potential strategy for ameliorating Islamophobia.

Keywords

identity overlap, intervention tournament, Islamophobia, media bias

Paper received 3 August 2020; revised version accepted 20 February 2022.

On March 15, 2019, a White supremacist entered two mosques in New Zealand and killed 51 worshippers (Graham-McLay, 2019). This attack was one of many attacks committed against innocent Muslims in Western countries over the previous years and highlights the recent increase in hate crimes (Federal Bureau of Investigation [FBI], 2019) and hostility (Kishi, 2017) toward Muslims in the West. Concurrent with anti-Muslim speech

[1]University of Pennsylvania, USA
[2]Beyond Conflict, USA
[3]Tel Aviv University, Israel

Corresponding author:
Samantha L. Moore-Berg, Annenberg School for Communication, University of Pennsylvania, 3620 Walnut St., Philadelphia, PA 19104, USA.
Email: Samantha.mooreberg@asc.upenn.edu

and violence, legislation has been introduced in the U.S. at both state and federal levels to ban Sharia law (Pilkington, 2017) and prevent Muslims from entering the country (American Civil Liberties Union, 2018). Further, non-Muslim U.S. Americans reliably report liking Muslims (and people from Muslim majority countries) less than non-Muslims (Bruneau et al., 2018; Sides & Gross, 2013), dehumanize Muslims more than non-Muslims (Kteily & Bruneau, 2017), reserve more empathy for non-Muslim Americans than for Muslims (Bruneau et al., 2017), and collectively blame Muslims more and collectively praise Muslims less than Christians or broader categories of White people for violence committed by extremists from each group (Bruneau et al., 2018; Gallardo et al., 2021). In Western Europe, bias against Muslims is similarly high and has similarly escalated in recent years (Bayrakli & Hafez, 2017).

Civil society organizations have been actively working to develop messages to combat Islamophobia. In the present research, we examine the efficacy of these existing anti-Islamophobia messages developed by activist organizations and then isolate the mechanism(s) through which the successful videos work. Using an intervention tournament, we investigated the effectiveness of several videos that could plausibly reduce Islamophobia, operationalized as support for anti-Muslim policies. These videos were developed based on the intuition of the people who created them, but the ability of these videos to causally reduce Islamophobia has not previously been determined. We do so here, and then conduct several follow-up studies to determine the mechanism(s) through which the "winning" interventions act.

Intervention Tournaments as a Method to Reduce Prejudice

Recently, scholars have developed interventions aimed at reducing prejudice towards marginalized groups (for reviews, see Bar-Tal & Hameiri, 2020; Paluck & Green, 2009). A standard framework for assessing the impact of a psychological intervention is to begin with a specific theoretical approach (e.g., the contact hypothesis), create an intervention that tests a specific aspect of the theoretical approach (e.g., an online platform that allows members of two groups to interact virtually), and then test the intervention in a controlled experimental setting (e.g., groups of students randomly assigned to interact virtually under the optimal conditions set out by Allport, 1954; [i.e., experimental condition] or to an equally motivated group of students who do not interact [i.e., a control condition]). Although this approach effectively tests theory, the constraints of a theory may limit how engaging the narrative of the intervention is—particularly if the narrative is constructed by scientists rather than professional storytellers. On the other hand, many real-world storytellers develop engaging content but do not always test the effectiveness of the content in achieving the desired outcomes (Davidson, 2017). Understanding what content is effective in real-world contexts is essential for determining where to invest resources to have the greatest impact. An approach that brings the key advantages of these two strategies together is an intervention tournament, or an experiment that tests and evaluates the causal effects of different approaches to real-world messaging (Bruneau et al., 2018; Lai et al., 2014, 2016; for a review, see Hameiri & Moore-Berg, in press).

For example, Bruneau et al. (2018) conducted an intervention tournament to identify videos that reduced collective blame hypocrisy towards Muslims. Here, we utilized a similar approach to identify videos that successfully reduce a broader set of anti-Muslim attitudes, including support for punitive community actions and anti-Muslim policies. We chose videos for the initial tournament that mapped onto at least one identifiable potential mechanism. However, following an initial exploratory study, two mechanisms became of particular interest in the intervention tournament: recognition of media bias and perceived identity overlap with Muslims. Thus, this research complements and extends the work conducted by Bruneau et al. (2018) by using a similar intervention tournament design to reduce Islamophobia. Specifically, Bruneau and colleagues highlighted the hypocrisy

of non-Muslims in blaming all Muslims for the acts of violence committed by extreme group members given that most non-Muslim White people do not hold all White people accountable for the acts of violence committed by extreme members of their ingroup. Here, we extend this work conceptually to show that highlighting sources of that bias (i.e., unfair media representations) and further highlighting the identity overlap between the ingroup and outgroup can have beneficial effects on reducing Islamophobia.

Recognition of Anti-Muslim Media Bias

Media bias against minoritized or marginalized groups in news coverage has been extensively documented by communication scholars (e.g., Figueroa-Caballero & Mastro, 2019). For example, the media tends to overrepresent Black Americans as criminals (Dixon & Linz, 2000; Gilliam & Iyengar, 2000). In turn, biased news coverage increases fear and prejudice against non-White Americans (Dixon, 2006; Gilliam & Iyengar, 2000; Gilliam et al., 2002; Mastro et al., 2009), and has important implications for policy support and broader political decision making (Gilliam et al., 2002; Valentino, 1999, 2001).

At least since 9/11, media bias against Muslims has also been extensively documented. For example, between 2008 and 2015, terror attacks carried out by Muslims received over 350% more coverage in the U.S. media than terror attacks committed by non-Muslims (Kearns et al., 2019), even though attacks by non-Muslims (e.g., White supremacists) were more prevalent over this time and represent an increasing threat (Miller, 2017). Negative news coverage about Muslims is also not limited to terrorism: even nonterrorism newspaper coverage about Muslims and Islam is overwhelmingly negative (Bleich & van der Veen, 2021; Media Tenor, 2013). Since most U.S. Americans do not personally know a Muslim person (Lipka, 2014), this media coverage is likely to strongly influence Americans' perceptions of Muslims (Shaver et al., 2017) and support for anti-Muslim policies (Saleem et al., 2017).

However, drawing attention to the harmful effects of biased media on behavior can have a positive impact on various outcomes, including reducing risky health (Bergsma & Carney, 2008), sexual (Allen et al., 1996), violent (Cantor & Wilson, 2003), and prejudiced behaviors (for a review, see Scharrer & Ramasubramanian, 2015). For example, informing people about relevant aspects of biased media can increase knowledge of that bias, which can have downstream effects of reducing related negative behaviors (for a meta-analytic review, see Jeong et al., 2012). Specifically, media literacy interventions (a) help people gain knowledge of the media, (b) increase people's awareness of how they are influenced by the media, and (c) enhance people's ability to interpret the veracity of the media they are exposed to (Jeong et al., 2012). This, in turn, can lead people to question whether relying on the media to form their intergroup attitudes is appropriate, which then can lead to prejudice reduction (Scharrer & Ramasubramanian, 2015). Thus, we extend the current media literacy intervention literature to test the possibility that media messages prompting people to recognize media bias against Muslims might help to reduce Islamophobia.

Identity Overlap

Another possible way that the media might help reduce Islamophobia is through stories that highlight the overlap between Muslim and American identities. Identity overlap refers to the extent that two or more identities are perceived to share similar traits/features. Recognizing identity overlap can serve as an effective way to reduce prejudice against outgroup members (e.g., Crisp & Beck, 2005; Riek et al., 2010; Wirtz & Doosje, 2013). Specifically, if people identify with a broader group, they are more likely to de-emphasize their negative evaluations of subgroup members of that group (Transue, 2007). For example, majority group members who recognize a common ingroup identity (see Gaertner & Dovidio, 2000) with immigrants report lower levels of racism towards immigrants, which leads to an increased willingness to support the integration of immigrants into their communities and to engage in pro-immigrant behaviors (e.g., donating money to immigrant causes; Kunst et al.,

2015). Likewise, simply bringing attention to a shared American identity between White Americans and racially minoritized groups increases support of pro-minority policies (e.g., progressive taxation to benefit minorities' education; Transue, 2007) and reduces prejudice more broadly (Dovidio et al., 2017). Given that 56% of Americans view Islam as incompatible with American values (New America, 2018), we also examine the role of increasing perceptions of identity overlap between Muslim and American identities to help reduce Islamophobia. Specifically, we examined the efficacy of videos that provide personal narratives of Muslim Americans who share a strong connection between their Muslim and American identities for reducing Islamophobia.

Overview of Research

Across six studies, we sought to identify media interventions that successfully reduce Islamophobia and then determine the mechanisms through which these interventions work. In Study 1, we conducted an intervention tournament to examine the immediate and lasting effects (1 month later) of 11 different videos aimed at reducing Islamophobia. Three videos emerged as most effective; thus, in Study 2, we focused on identifying measurable mechanisms through which the successful videos from Study 1 worked. Given that we did not control the content in these videos (as they were developed independently by practitioners), we identified possible mechanisms that could underlie the potential success of the videos prior to the start of these studies. We hypothesized that recognition of media bias could be the mechanism for the first successful video, and identity overlap could be the mechanism for the other two. Therefore, in Study 2, we tested whether these videos causally affect recognition of media bias and identity overlap, and whether these mechanisms are related to Islamophobia in this context. Finally, in Studies 3 and 4a–c, we replicated the mechanistic effects found for the successful videos in Study 2. Here, we focused on a sample of U.S. Americans because of the existing strong anti-Muslim prejudice in the U.S., and because the media we focused on are U.S. media.

Across all studies, we focused on two different outcome measures to assess whether the videos reduced specific consequences of Islamophobia: support for punitive efforts in Muslim American communities and support for anti-Muslim policies. Both measures have been used as outcome measures in previous assessments of interventions aimed at reducing Islamophobia (e.g., Bruneau et al., 2018; Kteily & Bruneau, 2017) and reflect real-life policies that continue to be proposed by the U.S. Congress.

Study 1

In Study 1, we conducted a two-wave longitudinal study to evaluate the efficacy of 11 videos aimed at reducing Islamophobia with an intervention tournament (i.e., randomized experiment) to compare the strength of the videos against a no-video control. Practitioners from social justice and Muslim advocacy nonprofit organizations provided us with videos that they had created or encountered aimed at reducing Islamophobia. Thus, the videos had intuitive appeal as anti-Islamophobia interventions (see Table 1 for video descriptions and links).

Before the tournament, we identified possible mechanisms through which these videos might work and named the videos based on these a priori hypothesized mechanisms. Four of the videos challenged common negative stereotypes about Muslims (see e.g., Prati et al., 2015) through a range of approaches, including Muslim Americans highlighting their American patriotism (Video 1: Counterstereotyping 1), Muslim American women describing their choice to wear a hijab (Video 2: Counterstereotyping 2), a Muslim American woman and a priest explaining Sharia law (Video 3: Counterstereotyping 3), and Americans pointing out how erroneous anti-Muslim stereotypes can be (Video 4: Counterstereotyping 4). Another video (Video 5: Media Bias) highlighted the bias against Muslims in the media (see e.g., Matthes et al., 2019) with an edited TED talk by a Muslim American woman

Table 1. Summaries and potential mechanisms for the videos used in the intervention tournament in Study 1.

Condition	Summary of video	Link to video	Length
No-video control condition	-	-	-
Video 1: Counterstereotyping 1	Interview with two Muslim New York police officers who were both responders during the 9/11 attacks.	https://player.vimeo.com/video/271933161	2:36
Video 2: Counterstereotyping 2	Several Muslim American women describe their personal motivations for choosing to wear a hijab.	https://player.vimeo.com/video/273366791	3:32
Video 3: Counterstereotyping 3	A priest and a Muslim woman describe Sharia law and break common stereotypes about it.	https://player.vimeo.com/video/273939730	2:57
Video 4: Counterstereotyping 4	A comedic clip that highlights stereotypes about Muslims in the United States.	https://player.vimeo.com/video/273939720	1:04
Video 5: Media Bias	TED talk by a Muslim American woman whose family members were recently murdered in a hate crime (same TED talk as used for Video 8). The video was edited to focus on her perceptions about media bias against Arabs and Muslims in the United States.	https://player.vimeo.com/video/273365939	1:59
Video 6: Identity Overlap 1	An informational video about the ways in which the American society benefits economically and socially from Muslim Americans.	https://player.vimeo.com/video/273366301	1:42
Video 7: Identity Overlap 2	A Muslim American comedian describes the challenges of being Muslim in America, and his persistent patriotism following the 9/11 attacks.	https://player.vimeo.com/video/273367313	4:02
Video 8: Identity Overlap 3	TED talk by a Muslim American woman whose family members were recently murdered in a hate crime (same TED talk as used for Video 5). The video was edited to focus on the pain she felt from losing her brother and sister-in-law.	https://player.vimeo.com/video/273364270	4:12
Video 9: Identity Overlap 4	Ex-White supremacists describe the similarities between White extremists and Muslim extremists.	https://player.vimeo.com/video/273368560	3:30
Video 10: Empathy 1	An interview with several Muslim refugee children describing their lives as refugees.	https://player.vimeo.com/video/273371322	1:49
Video 11: Modeling Contact	Two news clips of White men describing the transformation of their attitudes towards Muslims, from hatred to respect, after meeting Muslims.	https://player.vimeo.com/video/273368001	3:01

who asked viewers to consider how murders of Muslims by White extremists and murders of White people by Muslim extremists are covered differently by the media. This edited video was taken from a different segment of the TED talk used for Video 8: Identity Overlap 3 (see what follows). Four other videos highlighted overlap between identities (see e.g., Gaertner & Dovidio, 2000), such as between Americans and Muslims (Video 6: Identity Overlap 1; Video 7: Identity

Overlap 2; and Video 8: Identity Overlap 3) and between Muslim extremists and White extremists (Video 9: Identity Overlap 4). One of the videos (Video 10: Empathy 1) described suffering experienced by Muslim refugees and therefore seemed likely to induce empathy (see e.g., Batson & Ahmad, 2009). Finally, one video demonstrated the impact of intergroup contact (see e.g., Pettigrew & Tropp, 2006) by following two White men who experienced a change of heart toward Muslims after interacting with them (Video 11: Modeling Contact). Although we preidentified the potential mechanisms for these videos, they are not mutually exclusive, and it is likely that some mechanisms overlap between the videos; therefore, we included a variety of both confirmatory and exploratory measures to track a range of possible mechanisms in each study.

To determine the effects of the videos on Islamophobia, we randomly assigned non-Muslim American participants at Time 1 to either view one of the 11 intervention videos or not, in the empty control condition. Given that we were also interested in the mechanisms that make the videos effective, we included, for exploratory purposes, additional measures strongly associated with anti-Muslim policy support and anti-Muslim behaviors that we thought could be plausibly altered by one or more of the videos (i.e., collective blame of Muslims, identity overlap;[1] see supplemental material for details). At Time 2, participants completed the same outcome measures as in Time 1 but without rewatching the videos.

Methods

Participants. We recruited participants for all studies from Amazon's Mechanical Turk (MTurk).[2] At Time 1, we recruited 2,385 non-Muslim U.S. participants; however, 109 participants failed the attention check question (~5%; "This is a check question. Please answer 5 for this question"), resulting in a final sample of 2,276 participants (M_{age} = 36.10 years, SD_{age} = 11.79; 54% female; 72% White, 9% African American, 8% Asian, 5% Hispanic, 6% other; 49% Christian, 42% atheist/agnostic, 9% other), with at least 184 per condition. At Time 2, 2,049 participants from Time 1 completed the survey (86% retention rate); however, 173 participants failed the attention check questions at Time 1 and Time 2 (~8%; "This is a check question. Please answer 5 for this question"), resulting in 1,876 participants (54% female; M_{age} = 36.76, SD_{age} = 11.94; 73% White, 9% African American, 8% Asian, 5% Hispanic, 5% other; 49% Christian, 41% atheist/agnostic, 10% other), including at least 149 per condition (see Table S1 for sample size and demographics for each condition; see supplemental material for power and sensitivity analyses for all studies).[3] Participants were compensated US$1.50 per wave (US$3.00 total), lasting each ~10 minutes.

Measures and procedure. Participants were randomly assigned to either one of 11 video conditions or to the no-video control. Participants in the video conditions first watched a 1- to 4-minute video (only at Time 1); all participants then completed a survey which included key outcome measures of support for punitive community action and anti-Muslim policies at Times 1 and 2 (see supplemental material for additional methods).

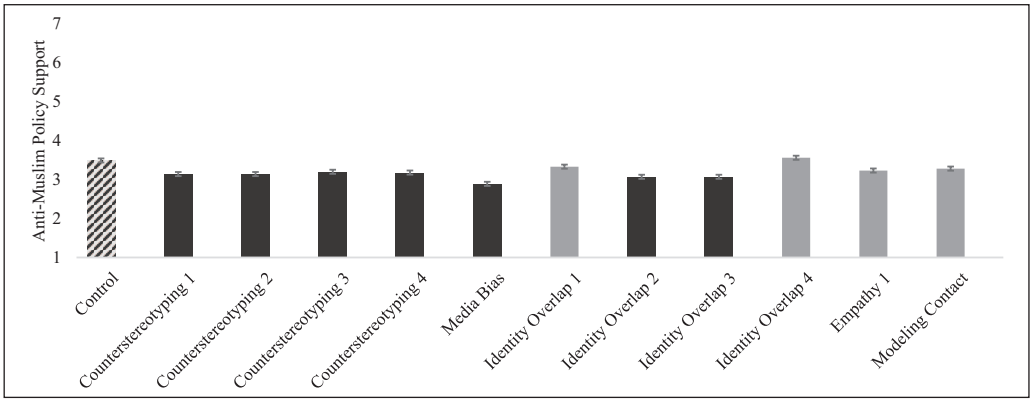
Support for punitive community action was assessed by providing participants with the following prompt:

> In an effort to give back to some of the communities that are targets of our studies, we have received a small grant that allows us to distribute some money to anti-terrorism efforts. We're giving each of our participants the opportunity to decide where this money should be distributed. Please indicate below what percentage of the money you would like to be distributed to each of the projects in the U.S.—we will then base our contributions on participants' recommendations.

We considered proactive community action as building libraries or investing in schools to increase educational opportunities, and punitive community action as increasing policing and surveillance. We used the percentage, out of 100%, allocated to policing and surveillance as our measure of punitive community action (see Bruneau et al., 2018).

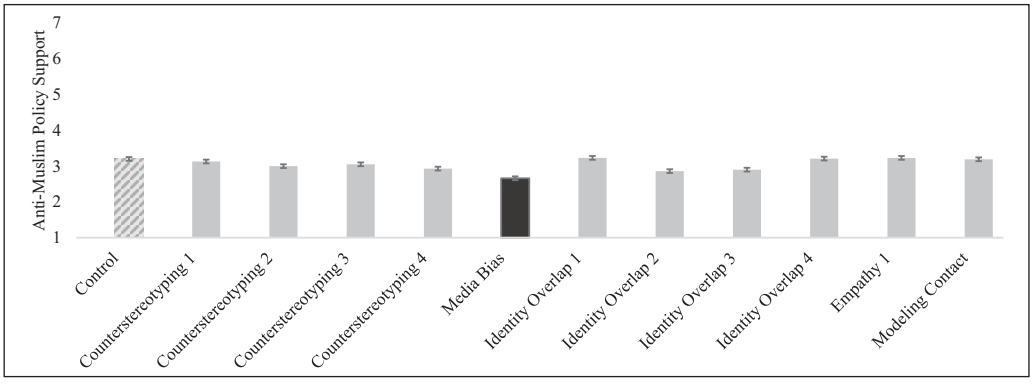
Anti-Muslim policy items were adapted from Kteily and Bruneau (2017) and were assessed using a 7-point scale (1 = *strongly disagree*, 7 = *strongly*

Figure 1a. Differences between condition and the control for support for anti-Muslim policies at Time 1: Study 1.



Note. Black bars = difference between the condition mean and control mean is significantly different, $p < .05$.

Figure 1b. Differences between condition and the control for support for anti-Muslim policies at Time 2: Study 1.



Note. Black bars = difference between the condition mean and control mean is significantly different, $p < .05$.

agree). Participants indicated their support for seven anti-Muslim policies (e.g., "We should preferentially admit Christian [vs. Muslim] refugees from the Middle East"). Pro-Muslim policies were reverse-coded, and all seven items were then averaged (Time 1: Cronbach's $\alpha = .83$; Time 2: Cronbach's $\alpha = .93$).

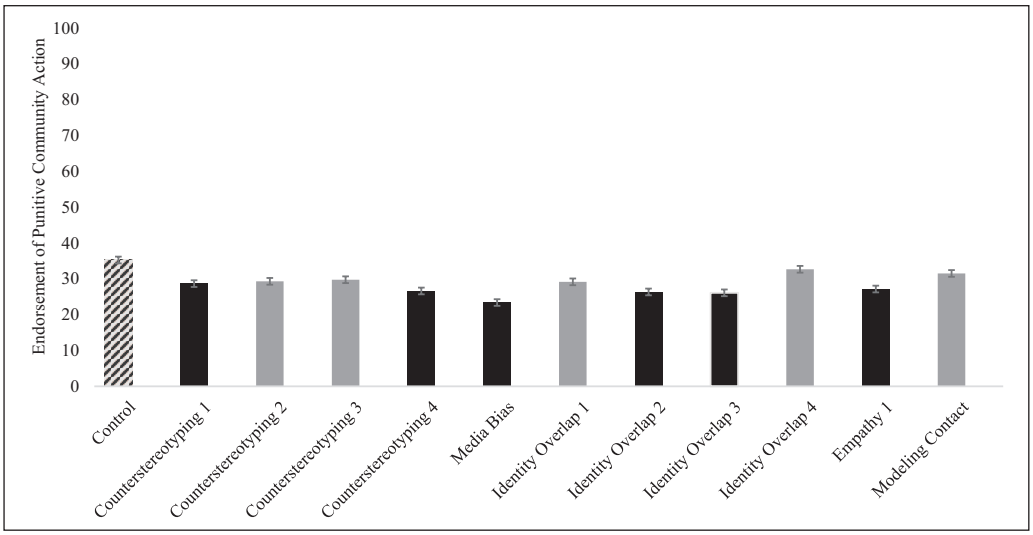
Participants then provided demographic information (e.g., age, gender, race, religion).

Results and Discussion

We conducted planned contrasts to identify significant differences between the control and video conditions. To account for multiple comparisons, we planned to replicate any significant effects with new samples to balance between Type I and Type II error risks. Mean results for each condition can be found in Figures 1 and 2. (See supplemental material for additional results, including descriptive statistics for all measures and conditions, planned comparisons between each video condition and the control, results controlling for religion, and interactions between time and condition.)[4]

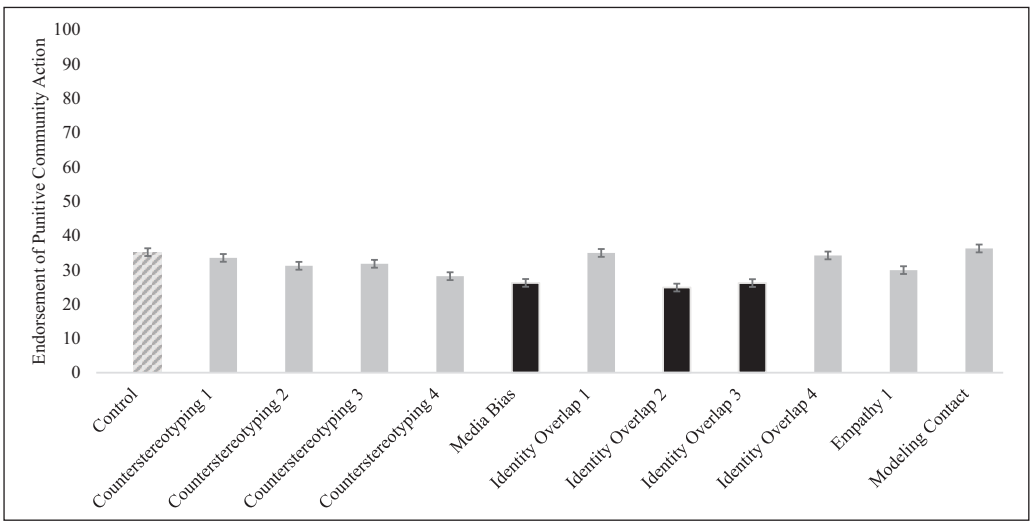
Participants' support for punitive community action was significantly lower than the control in five out of 11 conditions: Counterstereotyping 1 ($p = .047$, $d = 0.20$), Counterstereotyping 4 ($p = .008$, $d = 0.25$), Media Bias ($p < .001$, $d = 0.35$), Identity Overlap 2 ($p = .007$, $d = 0.26$), Identity

Figure 2a. Differences between condition and the control for punitive community action at Time 1: Study 1.



Note. Black bars = difference between the condition mean and control mean is significantly different, $p < .05$.

Figure 2b. Differences between condition and the control for punitive community action at Time 2: Study 1.



Note. Black bars = difference between the condition mean and control mean is significantly different, $p < .05$.

Overlap 3 ($p = .006$, $d = 0.27$), and Empathy 1 ($p = .014$, $d = 0.24$). One month after video exposure, participants continued to demonstrate decreased support for punitive community action in the Media Bias ($p = .016$, $d = 0.26$), Identity Overlap 2 ($p = .005$, $d = 0.30$), and Identity Overlap 3 ($p = .016$, $d = 0.27$) conditions, relative to control. The effect of the other three videos on endorsement of punitive community actions observed at Time 1 did not remain 1 month later ($ps \geqslant .059$, $ds \leqslant 0.20$), and the other null effects found at Time 1 remained at Time 2.

Support for anti-Muslim policies at Time 1 was significantly lower for participants in seven conditions, as compared to the control: Counterstereotyping 1 ($p = .014$, $d = 0.25$), Counterstereotyping 2 ($p = .013$, $d = 0.24$), Counterstereotyping 3 ($p = .037$, $d = 0.20$), Counterstereotyping 4 ($p = .028$, $d = 0.22$), Media Bias ($p < .001$, $d = 0.43$), Identity Overlap 2 ($p = .003$, $d = 0.30$), and Identity Overlap 3 ($p = .003$, $d = 0.30$). One month after video exposure, endorsement of anti-Muslim policies continued to be significantly lower for participants only in the Media Bias condition versus control ($p = .005$, $d = 0.31$). The effect of the other six videos on endorsement of anti-Muslim policies observed at Time 1 did not remain 1 month later ($ps \geqslant .068$, $ds \leqslant 0.20$), and the other null effects found at Time 1 remained at Time 2.

Study 2

Building on the promising results of the Identity Overlap 2, Identity Overlap 3, and media bias videos persistently reducing at least one key outcome measure in Study 1, in Studies 2–4, we turned our focus on identifying the mechanisms driving the effects of each of these videos.

In the Media Bias video, Suzanne Barakat, a Muslim American woman, describes the media aftermath of the murder of her family members by a White supremacist. She explains how the police and the media initially portrayed the murderer as motivated by a parking dispute, but viewers later find out that there was no parking dispute, and that the murderer was motivated by Islamophobia. As the video concludes, Barakat highlights differences in media portrayals of White versus Muslim extremism, and challenges viewers to think about biases in the ways the media covers stories about different groups of people. Thus, this video focuses on increasing knowledge about media biases through an emotional story, which could encourage viewers to reflect on their personal, media-driven biases. In the Identity Overlap 2 video, comedian Hasan Minhaj describes his experience post-9/11. He describes the anger he felt when his family's car was vandalized, and that he received threatening phone calls from fellow Americans because he is Muslim. He explains how his anger was fueled by others not realizing how much he loves America and how much he identifies as being an American. In the Identity Overlap 3 video, Suzanne Barakat (the same woman from the Media Bias video) provides detailed information about her family members' murder. She first describes how all-American her murdered family members were and then describes the murder in detail. Thus, in both Identity Overlap videos, the narrators describe how their Muslim identity complements their American identity as a central theme.

Given the content of the successful videos and the important role that recognition of media bias and identity overlap play in prejudice reduction (as noted in the introduction), in a preregistered study, we predicted that (a) recognition of media bias would indirectly affect the relationship between the Media Bias video and outcome measures, and (b) perceived identity overlap would indirectly affect the relationship between the Identity Overlap 2 and 3 videos and outcome measures, among other processes (see supplemental material). We also preregistered indirect effects of empathy and patriotism and a moderating effect of need for cognition, however, these hypotheses were not confirmed (see supplemental material for details).

Methods

Participants. We recruited 794 non-Muslim U.S. participants through MTurk. Thirty-five participants were excluded for failing the attention check question (~4%; "This is a check question. Please answer 5 for this question"), resulting in 759 participants ($M_{age} = 34.92$ years, $SD_{age} = 10.64$; 48% female; 73% White, 8% African American, 8% Asian, 5% Hispanic, 6% other; 54% Christian, 39% atheist/agnostic, 7% other). Participants were compensated US$1.00.

Measures and procedure. Participants were randomly assigned to watch either the Media Bias, Identity Overlap 2, or Identity Overlap 3 video, or were assigned to the no-video control (see supplemental

Table 2. Means, standard deviations, and planned contrast comparisons for each focal measure: Study 2.

	Condition	Recognition of media bias	Identity overlap	Punitive community action	Anti-Muslim policies
M (*SD*)	Control	62.55 (30.09)	3.27 (1.52)	35.09 (33.24)	3.14 (1.74)
	Media Bias	69.65 (28.72)	3.54 (1.66)	29.18 (30.92)	2.80 (1.64)
	Identity Overlap 2	59.69 (32.73)	3.66 (1.68)	30.84 (30.12)	3.01 (1.65)
	Identity Overlap 3	65.01 (31.16)	3.72 (1.72)	33.57 (31.54)	3.03 (1.67)
Planned contrasts comparing target videos and control: *t* value (Cohen's *d*)	Media Bias	2.23 (0.24)*	1.59 (0.17)	−1.82 (0.18)	−1.98 (0.20)*
	Identity Overlap 2	−0.90 (0.09)	2.33 (0.24)*	−1.32 (0.13)	−0.75 (0.08)
	Identity Overlap 3	0.78 (0.08)	2.67 (0.28)**	−0.47 (0.05)	−0.75 (0.06)

Note. For additional measures, see the supplemental material.
*$p < .05$. **$p < .01$. ***$p < .001$.

material for condition demographics). Participants then completed recognition of anti-Muslim media bias and identity overlap measures in randomized order. To assess recognition of anti-Muslim media bias, participants indicated how strongly the media is biased against Muslim Americans using a 0 (*not at all*) to 100 (*very much*) scale. To assess identity overlap, participants indicated their perceptions of identity overlap between Muslims and Americans with a series of seven pairs of circles with different degrees of overlap (ranging from complete separation to full overlap; see e.g., Swann et al., 2009). Then, participants completed the punitive community action and anti-Muslim policy support (Cronbach's $\alpha = .93$) outcome measures in randomized order, as in Study 1. Finally, participants completed the same demographic measures as in Study 1.

Results and Discussion

We conducted planned contrasts to compare each video condition to the control. Descriptive statistics for all measures and conditions, and comparisons between each video condition and the control, are shown in Table 2 (see supplemental material for additional results).

As predicted, participants indicated that the media was significantly more biased against Muslims after watching the Media Bias video ($p = .026$, $d = 0.24$) versus control. Yet recognition of media bias was not different between the control and Identity Overlap 2 ($p = .367$, $d = 0.09$) or Identity Overlap 3 ($p = .438$, $d = 0.08$) conditions.

Also as predicted, compared to control, perceived overlap in identity between Americans and Muslims was greater for participants in the Identity Overlap 2 ($p = .020$, $d = 0.24$) and Identity Overlap 3 ($p = .008$, $d = 0.28$) conditions, but not for participants in the Media Bias condition ($p = .111$, $d = 0.17$).

Despite the lasting effects of the three video conditions on the policy and punitive community action outcomes in Study 1, in Study 2 only participants who watched the Media Bias video reported significantly reduced support for anti-Muslim policies (Media Bias: $p = .049$, $d = 0.20$; Identity Overlap 2: $p = .455$, $d = 0.08$; Identity Overlap 3: $p = .523$, $d = 0.06$), relative to the control. None of the videos reduced support for punitive community action (Media Bias: $p = .070$, $d = 0.18$; Identity Overlap 2: $p = .188$, $d = 0.13$; Identity Overlap 3: $p = .638$, $d = 0.05$), relative to the control.

We then examined the indirect relationships between the video conditions and the key outcome measures using Process Version 3.3 Model 4 (Hayes, 2018). For all studies, we dummy-coded condition to contrast each video condition with the control, and then included recognition of media bias and identity overlap as parallel mediators. As preregistered, there was an indirect effect of recognition of media bias, but not identity overlap, on the relationship

Table 3. Standardized indirect effects: Study 2.

Intervention video vs. control → Recognition of media bias → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Media Bias	−0.06	0.03	[−0.12, −0.01]
Identity Overlap 2	0.03	0.03	[−0.03, 0.08]
Identity Overlap 3	−0.02	0.03	[−0.08, 0.03]
Intervention video vs. control → Recognition of media bias → Punitive community action			
	Indirect effect	Boot *SE*	95% CI
Media Bias	−0.06	0.03	[−0.12, −0.01]
Identity Overlap 2	0.02	0.03	[−0.03, 0.08]
Identity Overlap 3	−0.02	0.03	[−0.07, 0.03]
Intervention video vs. control → Identity overlap → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Media Bias	−0.04	0.03	[−0.09, 0.01]
Identity Overlap 2	−0.06	0.03	[−0.11, −0.01]
Identity Overlap 3	−0.06	0.03	[−0.12, −0.01]
Intervention video vs. control → Identity overlap → Punitive community action			
	Indirect effect	Boot *SE*	95% CI
Media Bias	−0.04	0.03	[−0.10, 0.01]
Identity Overlap 2	−0.07	0.03	[−0.13, −0.01]
Identity Overlap 3	−0.08	0.03	[−0.14, −0.02]

between the Media Bias video versus the control and anti-Muslim policies and punitive community action. As preregistered, we found an indirect effect of identity overlap for the Identity Overlap 2 and Identity Overlap 3 videos versus the control condition, but not for the Media Bias video versus the control, on the outcomes (see Table 3).

We did not find support for other hypothesized indirect effects, including empathy and dehumanization, or for the moderation of need for cognition (see supplemental material).

Overall, Study 2 provided evidence linking the successful videos identified in Study 1 to the predicted mechanisms, which, in turn, related to policy-relevant outcomes. In Studies 3 and 4a–c, we sought to obtain additional evidence relevant to the robustness of the significant indirect effects observed in Study 2.

Study 3

We conducted Study 3 to clarify the robustness of the mechanism effects of the three key videos. In a preregistered study, we hypothesized that there would be an indirect effect of recognition of media bias on the relationship between the Media Bias video and outcome measures (i.e., anti-Muslim policy support and punitive community action measures), and an indirect effect of identity overlap on the relationship between the Identity Overlap 2 and Identity Overlap 3 videos and outcomes. We also preregistered that there might be an indirect effect of collective

Table 4. Means, standard deviations, and planned contrast comparisons for each focal measure: Study 3.

	Condition	Recognition of media bias	Identity overlap	Punitive community action	Anti-Muslim policies
M (*SD*)	Control	54.78 (29.02)	3.72 (1.66)	39.31 (31.23)	3.47 (1.72)
	Media Bias	63.61 (29.10)	3.43 (1.66)	40.42 (34.40)	3.48 (1.83)
	Identity Overlap 2	57.72 (29.76)	3.83 (1.51)	33.65 (30.65)	3.22 (1.66)
	Identity Overlap 3	59.48 (28.89)	3.81 (1.74)	38.20 (31.36)	3.43 (1.78)
Planned contrasts comparing target videos and control: *t* value (Cohen's *d*)	Media Bias	2.97 (0.30)**	−1.73 (0.17)	0.34 (0.03)	0.08 (0.01)
	Identity Overlap 2	1.00 (0.10)	0.64 (0.07)	−1.76 (0.18)	−1.41 (0.15)
	Identity Overlap 3	1.60 (0.16)	0.50 (0.05)	−0.35 (0.04)	−0.18 (0.02)

Note. For additional measures, see the supplemental material.
*$p < .05$. **$p < .01$. ***$p < .001$.

blame, however, this hypothesis was not confirmed (see supplemental material for details).

Methods

Participants. We recruited 862 non-Muslim U.S. participants for this study through MTurk. Sixty-five participants were excluded from this study for failing the attention check question (~8%; "This is a check question. Please answer 5 for this question"), resulting in 797 participants (M_{age} = 34.37 years, SD_{age} = 11.13; 44% female; 60% White, 10% African American, 18% Asian, 6% Hispanic, 6% other; 52% Christian, 36% atheist/agnostic, 12% other). Participants were compensated US$1.00.

Measures and procedure. Participants were first assigned to either a video condition or the no-video control. In the video conditions, participants either watched the Media Bias, Identity Overlap 2, or Identity Overlap 3 video from Studies 1 and 2 (see supplemental material for condition demographics). Participants then completed the same recognition of media bias, identity overlap, punitive community action, anti-Muslim policies (Cronbach's α = .93), and demographic measures as in our previous studies (see supplemental material for collective blame).

Results and Discussion

We conducted the same analyses as in Study 2. Descriptive statistics for all measures, conditions, and comparisons can be found in Table 4 (see supplemental material for additional results).

As preregistered, relative to the control, participants indicated that the media is significantly more biased against Muslims after watching the Media Bias video (p = .003, d = 0.30), but not the Identity Overlap 2 (p = .318, d = 0.10) or Identity Overlap 3 (p = .111, d = 0.16) video. However, as in Study 2, there was no direct effect of any of the video conditions, versus the control, on anti-Muslim policies (Media Bias: p = .934, d = 0.01; Identity Overlap 2: p = .160, d = 0.15; Identity Overlap 3: p = .855, d = 0.02) or support for punitive community action (Media Bias: p = .733, d = 0.03; Identity Overlap 2: p = .080, d = 0.18; Identity Overlap 3: p = .730, d = 0.04). As preregistered, there was a significant indirect effect of recognition of media bias on the relationship between the Media Bias video versus the control on the outcomes, but not for the Identity Overlap 2 or Identity Overlap 3 videos versus the control on the outcomes (see Table 5).

For identity overlap, there was no difference between the control and any video condition (Media Bias: p = .084, d = 0.17; Identity Overlap 2: p = .523, d = 0.07; Identity Overlap 3: p = .619, d = 0.04). There was not an indirect effect of identity overlap for any video versus control on anti-Muslim policies and punitive community action. The lack of an indirect effect here is contrary to our preregistered hypotheses; we preregistered that there would be an indirect effect of identity overlap on the relationship between the Identity Overlap 2 and 3 videos and the outcome measures (see Table 5). Finally,

Table 5. Standardized indirect effects: Study 3.

Intervention video vs. control → Recognition of media bias → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Media Bias	−0.04	0.02	[−0.09, −0.01]
Identity Overlap 2	−0.01	0.02	[−0.05, 0.01]
Identity Overlap 3	−0.02	0.02	[−0.06, 0.01]
Intervention video vs. control → Recognition of media bias → Punitive community action			
	Indirect effect	Boot *SE*	95% CI
Media Bias	−0.07	0.03	[−0.13, −0.02]
Identity Overlap 2	−0.02	0.03	[−0.08, 0.02]
Identity Overlap 3	−0.04	0.03	[−0.09, 0.01]
Intervention video vs. control → Identity overlap → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Media Bias	0.01	0.01	[−0.003, 0.04]
Identity Overlap 2	−0.01	0.01	[−0.03, 0.01]
Identity Overlap 3	−0.004	0.01	[−0.03, 0.01]
Intervention video vs. control → Identity overlap → Punitive community action			
	Indirect effect	Boot *SE*	95% CI
Media Bias	0.04	0.02	[−0.01, 0.08]
Identity Overlap 2	−0.01	0.02	[−0.06, 0.03]
Identity Overlap 3	−0.01	0.02	[−0.06, 0.03]

we did not observe an effect of the intervention videos on collective blame (see supplemental material).

Overall, the goal of Study 3 was to replicate and confirm the recognition of media bias and identity overlap mechanism effects for the key videos of interest (Media Bias, Identity Overlap 2, and Identity Overlap 3). As expected, the robust indirect effect of the Media Bias video on outcomes through recognition of media bias was replicated here. These results confirmed the effectiveness of recognition of media bias for reducing Islamophobia. However, the less robust indirect effects of idenity overlap on the relationship between the Identity Overlap 2 and Identity Overlap 3 videos and the punitive community action and policy outcomes were not replicated. Although we found consistent evidence that perceived identity overlap is related to policy and punitive community action preferences, we did not find robust support in this study for which videos can elicit perceptions of identity overlap.

Studies 4a–c

In Studies 4a–c, we aimed to replicate the effects found in Study 2 with three additional well-powered samples.

Methods

Participants. We recruited a total of 1,964 non-Muslim U.S. participants for these studies through MTurk; however, 110 participants failed the attention check question (~6%; "This is a check question. Please answer 5 for this question"), resulting in 1,854 participants in the final sample (Study 4a: $N = 766$; Study 4b: $N = 557$;

Table 6. Means, standard deviations, and planned contrasts for each measure: Study 4a.

	Condition	Recognition of media bias	Anti-Muslim policies
M (*SD*)	Control	0.59 (5.31)	2.93 (1.66)
	Media Bias	2.31 (5.75)	2.79 (1.66)
	Empathy 1	1.77 (5.12)	2.82 (1.62)
	Counterstereotypical	1.10 (5.59)	2.74 (1.64)
Planned contrasts:	Media Bias	3.09 (0.31)**	−0.79 (0.08)
Target video vs control:	Empathy 1	2.15 (0.23)*	−0.65 (0.07)
t value (Cohen's *d*)	Counterstereotypical	0.92 (0.09)	−1.11 (0.12)

*$p < .05$. **$p < .01$. ***$p < .001$.

Study 4c: $N = 531$; see supplemental material for complete demographic information).

Measures and procedures. Participants were randomly assigned either to watch the video of interest (Study 4a: Media Bias; Study 4b: Identity Overlap 2; Study 4c: Identity Overlap 3), to a video control (Studies 4a–c: Empathy 1; Study 4a: Counterstereotypical), or to a no-video control condition (all studies). Participants then completed the same recognition of media bias (Studies 4a and 4c), identity overlap (Studies 4b and 4c), anti-Muslim policies (Studies 4a–c), and punitive community action (Study 4b) measures as in Studies 2 and 3.

Results and Discussion

We conducted the same analyses as in Study 2 (see supplemental material for complete reporting of results as well as additional exploratory variables).[5]

Study 4a. As predicted, and consistent with Study 2, participants perceived that the media was significantly more biased against Muslims after watching the Media Bias video ($p = .002$, $d = 0.31$) versus the no-video control. Those in the Empathy 1 condition also reported greater recognition of media bias relative to the control ($p = .032$, $d = 0.23$). There was not a significant difference in recognition of media bias between the Counterstereotypical video and the no-video control ($p = .361$, $d = 0.09$). There was also no difference between any video condition versus the control on anti-Muslim policies (Media Bias: $p = .427$, $d = 0.08$; Empathy 1: $p = .518$, $d = 0.07$; Counterstereotypical 3: $p = .269$, $d = 0.12$), which we suspect is due to lower levels of support for anti-Muslim policies in the control than in Study 1. As in Study 2, we found a significant indirect effect of recognition of media bias for the Media Bias and Empathy 1 videos versus the no-video control, but not for the Counterstereotypical video versus the no-video control, on anti-Muslim policies (see Tables 6 and 7).

Study 4b. As in Study 2, relative to the controls, participants in the Identity Overlap 2 video condition indicated greater overlap between Muslims and Americans ($p = .049$, $d = 0.20$), but there was no difference in identity overlap scores between those who watched the Empathy 1 video and the control ($p = .992$, $d < 0.01$). However, there was no difference between any video condition versus the control for punitive community action (Identity Overlap 2: $p = .393$, $d = 0.08$; Empathy 1: $p = .148$, $d = 0.15$) or anti-Muslim policies (Identity Overlap 2: $p = .477$, $d = 0.07$; Empathy 1: $p = .979$, $d = 0.01$). Again, this lack of an effect was likely due to the control group's lower mean levels of punitive community action and anti-Muslim policy support in this study than in Study 1. We did not find an indirect effect of identity overlap for the Identity Overlap 2 and Empathy 1 videos versus the no-video control condition on anti-Muslim policies or punitive community action (see Tables 8 and 9).

Table 7. Standardized indirect effects: Study 4a.

Intervention video vs. control → Recognition of media bias → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Media Bias	−0.06	0.02	[−0.11, −0.02]
Empathy 1	−0.04	0.02	[−0.09, −0.005]
Counterstereotypical	−0.02	0.02	[−0.06, 0.02]

Table 8. Means, standard deviations, and planned contrasts for each measure: Study 4b.

	Condition	Identity overlap	Punitive community action	Anti-Muslim policies
M (*SD*)	Control	3.42 (1.77)	32.78 (32.98)	3.43 (1.55)
	Identity Overlap 2	3.77 (1.74)	30.06 (31.37)	3.33 (1.37)
	Empathy 1	3.42 (1.63)	28.22 (27.16)	3.44 (1.42)
Planned contrasts comparing target videos and control: *t* value (Cohen's *d*)	Identity Overlap 2	1.98 (0.20)*	−0.86 (0.08)	−0.71 (0.07)
	Empathy 1	−0.01 (< 0.01)	−1.45 (0.15)	0.03 (0.01)

*$p < .05$. **$p < .01$. ***$p < .001$.

Study 4c. There was no difference between any video condition versus the control on identity overlap (Identity Overlap 3: $p = .342$, $d = 0.10$; Empathy 1: $p = .192$, $d = 0.14$), recognition of media bias (Identity Overlap 3: $p = .210$, $d = 0.13$; Empathy 1: $p = .807$, $d = 0.03$), or anti-Muslim policies (Identity Overlap 3: $p = .270$, $d = 0.11$; Empathy 1: $p = .431$, $d = 0.08$). We also did not find an indirect effect of idenity overlap or recognition of media bias for the Identity Overlap 3 video or the Empathy 1 video (vs. no-video control) on anti-Muslim policies (see Tables 10 and 11).

Mini Meta-Analysis

To gauge the overall effect of the Media Bias, Identity Overlap 2, and Identity Overlap 3 interventions, versus the control, on anti-Muslim policy support and punitive community action across studies, we ran a series of mini meta-analyses using data from all studies in which at least one of the key dependent variables was measured (Goh et al., 2016). We used fixed effects by weighting the mean effect size by sample size. The meta-analyses showed that the effects across studies were relatively small, but significant, for the Media Bias, Identity Overlap 2, and Identity Overlap 3 interventions on both anti-Muslim policy support[6] and punitive community action (see Table 12). Thus, across studies, all videos reduced anti-Muslim policy support and punitive community action as compared to the control.

General Discussion

To counter the rise in hostility towards Muslims around the world, members of Muslim rights organizations have developed media interventions aimed at reducing Islamophobia. Here, we evaluated the effectiveness of 11 of these videos in reducing Islamophobia immediately after exposure and 1 month later. We identified three videos that had lasting effects on our two key measures of Islamophobia and tested the mechanism(s) through which these three "winning" videos (Media Bias, Identity Overlap 2, and Identity Overlap 3) exerted their effects.

The clearest and most robust finding from the current work was that a video (Media Bias) that

Table 9. Standardized indirect effects: Study 4b.

Intervention video vs. control → Identity overlap → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Identity Overlap 2	−0.05	0.03	[−0.11, 0.002]
Empathy 1	0.001	0.03	[−0.05, 0.05]
Intervention video vs. control → Identity overlap → Punitive community action			
	Indirect effect	Boot *SE*	95% CI
Identity Overlap 2	−0.06	0.03	[−0.13, 0.002]
Empathy 1	0.001	0.03	[−0.06, 0.06]

Table 10. Means, standard deviations, and planned contrasts for each measure: Study 4c.

	Condition	Identity overlap	Anti-Muslim policies	Recognition of media bias
M (*SD*)	Control	3.77 (1.51)	3.39 (1.71)	3.10 (4.57)
	Identity Overlap 3	3.92 (1.64)	3.20 (1.61)	2.46 (5.08)
	Empathy 1	3.56 (1.47)	3.26 (1.60)	2.97 (4.52)
Planned contrasts comparing target videos and control: *t* value (Cohen's *d*)	Identity Overlap 3	0.95 (0.10)	−1.10 (0.11)	−1.26 (0.13)
	Empathy 1	−1.31 (0.14)	−0.79 (0.08)	−0.25 (0.03)

*$p < .05$. **$p < .01$. ***$p < .001$.

Table 11. Standardized indirect effects: Study 4c.

Intervention video vs. control → Recognition of media bias → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Identity Overlap 3	0.02	0.02	[−0.01, 0.05]
Empathy 1	0.003	0.01	[−0.02, 0.03]
Intervention video vs. control → Identity overlap → Anti-Muslim policies			
	Indirect effect	Boot *SE*	95% CI
Identity Overlap 3	0.001	0.01	[−0.02, 0.01]
Empathy 1	−0.002	0.01	[−0.02, 0.02]

focused on media bias against Muslims reduced support for anti-Muslim policies both immediately and 1 month later, and that this effect works, in part, through the hypothesized increase in awareness of media bias against Muslims. Past research suggests that biases in the media influence attitudes towards marginalized groups in harmful ways; for instance, media bias can lead to generalizations of individual attributes to the population as a whole and can result in endorsement of punitive punishment for benign offenses (e.g., Dixon, 2006; Gilliam & Iyengar, 2000). The

Table 12. Mini meta-analysis.

Anti-Muslim policies				
	Mean *d*	95% CI	*Z*	*p*
Media Bias	0.18	[0.08, 0.28]	3.42	< .001
Identity Overlap 2	0.15	[0.05, 0.25]	2.88	.004
Identity Overlap 3	0.12	[0.02, 0.23]	2.40	.016
Punitive community action				
	Mean *d*	95% CI	*Z*	*p*
Media Bias	0.17	[0.06, 0.29]	2.94	.003
Identity Overlap 2	0.17	[0.07, 0.27]	3.24	.001
Identity Overlap 3	0.12	[0.001, 0.23]	1.98	.048

Note. For these analyses, we included data from all studies that included at least one of the key dependent variables. Anti-Muslim policies: Media Bias = Studies 1–3 and 4a; Identity Overlap 2 = Studies 1–3 and 4b; Identity Overlap 3 = Studies 1–3 and 4c. Punitive community action: Media Bias = Studies 1–3; Identity Overlap 2 = Studies 1–3 and 4b; Identity Overlap 3 = Studies 1–3.

deleterious effects of consuming biased news media can occur even after brief exposure: a mere 12 minutes of media content linking Islam to terrorism can lead to a significant increase in Islamophobia (Ogan et al., 2014). But why does media bias have such strong, negative effects on attitudes? According to the reinforcing spirals of media selectivity and effects theory (Slater, 2007), media use influences people's attitudes, which in turn determines the media to which people choose to be exposed. This leads to a reciprocal cycle in which these processes mutually reinforce each other. If people are continually exposed to media that perpetuates anti-Muslim tropes, then these views will likely be perpetuated and reinforced in the media people seek.

The current research demonstrated that a 2-minute video that draws people's attention to media bias durably reduces support for anti-Muslim policies by increasing recognition of anti-Muslim media bias. What makes this finding particularly striking is that the video itself did not explicitly call out media bias. Instead, Suzanne Barakat, a Muslim American woman, spoke about the horrific murder of her family members and the media coverage of the event, and only at the end did she ask the audience to reflect on the hypocrisy and bias in the media, "If roles were reversed and an Arab, Muslim, or Muslim appearing person had killed three White American college students, execution style, in their home, what would we have called it?" As past research suggests, messages might be more persuasive when the moral is implicit and viewers are required to develop their own conclusions (Perloff, 2010). The fact that Islamophobia was reduced immediately after watching the video suggests that people may have responded to it by correcting the negative views about Muslims they had previously acquired, through the media and other sources, by now acknowledging that their former views were biased and hypocritical (a finding that conceptually replicates Bruneau et al., 2018). Ideally, a media bias intervention's effects would go further than this—if media bias is truly countered, then people would be less likely affected by anti-Muslim media content they may encounter in the future and might take steps to demand fairer coverage.

The fact that the effects of a single exposure to the Media Bias video in the current research remained 1 month later is encouraging, particularly since the enduring effects of this video occurred in the context of a still biased American media environment (e.g., Dixon, 2006; Gilliam & Iyengar, 2000; Gilliam et al., 2002). This research provides initial evidence for a potentially promising approach to anti-Islamophobia efforts. Yet,

repeated exposure to messages can have cumulative effects over time (Hornik, 2002), and the promising effects of the Media Bias video on helping people recognize bias in media coverage and adjust their attitudes highlight the need for more research to determine whether more intensive or repeated interventions of this kind could produce even stronger and longer lasting effects.

Our second main finding was a robust correlation between perceived Muslim and American identity overlap and reduced support for anti-Muslim policies. This is consistent with self-categorization theory (Turner et al., 1987) and the common ingroup identity model (Dovidio et al., 1998), which suggest that people who place value on similar identities are more likely to see others as similar, which then improves intergroup relationships. For example, minority groups who emphasize their shared identity with majority group members are evaluated more positively and evoke less anger than those who emphasize their minority group membership (Wirtz & Doosje, 2013). In two of the videos (Identity Overlap 2 and Identity Overlap 3), the Muslim American protagonists emphasized their American identity as complementary to their Muslim identity. Although the videos increased perceived identity overlap in some instances, this effect was not consistent across studies. One possibility is that the inconsistencies in identity overlap between studies could be due to prior attitudes. For instance, participants in Study 2's control condition had lower baseline levels of identity overlap ($M = 3.27$) than control participants from Studies 3 and 4c ($M = 3.72$ and 3.77, respectively), suggesting that there was more room for movement in Study 2's intervention condition than in Studies 3 and 4c. Thus, these findings provide some tentative support for perceived identity overlap as an effective technique for causally reducing prejudice (e.g., Crisp & Beck, 2005; Wirtz & Doosje, 2013), in parallel with more robust evidence that perceiving more identity overlap is associated with reduced support for anti-Muslim policies.

In addition to highlighting the mechanisms through which the more successful videos worked, we also identified eight intervention videos that were either not initially successful at reducing support for anti-Muslim policies and/or endorsement of punitive community action methods or only had short-term effects on the outcome measures of interest. At the same time, while the additional videos did not affect our key outcomes of interest, some of them were successful at changing other important outcomes not focused on in this paper—such as collective blame of Muslims (see supplemental material). Understanding not only whether an intervention works but also which types of interventions work through which mechanisms provides both practical and theoretical benefits for intervention research. Specifically, identifying interventions that exert effects through different mechanisms could help others design interventions that have synergistic effects and/or that may work when used in parallel or serially. Likewise, understanding why some of these intervention videos did not impact our chosen outcomes is an important avenue for future research as it will provide key insights into how to frame future intervention content. For example, while two of our identity overlap videos successfully reduced Islamophobia, the other two videos labeled as tapping into identity overlap did not affect the outcomes (perhaps due to unappealing content that made viewers resistant to the messages), and in Studies 3 and 4a–c, we also found less consistent support for the effects of our target videos on perceived identity overlap. This might suggest person by message interactions that were not captured in the current data; thus, we encourage future researchers to utilize intervention tournament experiments to understand both why an intervention works and does not work in particular contexts (see Bar-Tal & Hameiri, 2020).

Relatedly, although we did not find consistent effects of our key intervention videos on the key outcome measures across studies, the mini meta-analysis suggests that these effects are present albeit small. We are encouraged by the fact that these short videos were able to change entrenched prejudiced attitudes that are typically challenging to change. Likewise, our data show persistence of

the video effects over time—an assessment that is often overlooked in prejudice reduction research (Paluck et al., 2021). Thus, although these effects are small, they are not trivial. However, we contend that longer interventions and/or more intervention exposure might increase the effect size, and that the interventions might vary in their effectiveness according to participant identity, baseline attitudes, and other variables that could be fruitfully explored in future studies designed to amplify the effects observed here.

Beyond utilizing intervention tournaments to determine which type of message works to correct perceptions reinforced by biased portrayals in the media, it is also important to work to prevent biased portrayals in the first place. Mainstream media is typically rooted in power imbalances (Ramasubramanian & Banjo, 2020); who owns and operates the media largely determines what groups are portrayed and how they are portrayed. If mainstream media is owned and operated primarily by White people from Western, educated, industrialized, rich, and democratic (WEIRD) nations, then the voices of non-White and non-WEIRD people are less likely to be heard (Ramasubramanian & Banjo, 2020). However, by recognizing the power imbalances in the media and working to disrupt them, it is possible to amplify the voices of the marginalized and transcend the perpetuation of harmful stereotypes (Ramasubramanian & Banjo, 2020). Thus, more attention should be devoted to understanding and reducing these power imbalances to create more equitable media.

Despite the strengths of this research, it should also be interpreted in the context of certain limitations. First, although we identified recognition of media bias and identity overlap as two key mechanisms for our successful intervention videos, it is likely that they are not the only mechanisms at play. For example, our Media Bias intervention video also highlights the senseless murder of Barakat's family, which could have elicited other emotions than the ones tested (e.g., anger). Although we were able to tease out additional potential mechanisms in each follow-up study (such as collective blame, patriotism, empathy; see supplemental material), this does not mean that we exhausted the list of all possible mechanisms for each video. Future research should focus on identifying additional mechanisms that work independently from and simultaneously with our discovered ones. Second, experimenter demand effects could have affected this research. Although we did not explicitly state what these studies were about at the beginning of each one, participants could have intuited that they were about Islamophobia and therefore altered their responses accordingly. But given that most of the videos did not reduce Islamophobia, demand characteristics are less likely an explanation for our significant effects. In other words, if demand were responsible for the effects observed, we would expect to see this across conditions. Third, although we partnered with Muslim rights organizations to curate the intervention videos, we did not assess how these videos affect Muslim populations. Although research on Muslims' reactions to media content has been limited, research suggests that Muslim Americans' exposure to negative media content about Muslims leads to greater desire to distance themselves from non-Muslim Americans (Saleem & Ramasubramanian, 2017). Thus, more research is needed to understand how Muslims respond to messages similar to those utilized in this research.

Fourth, one key component missing from our analyses is a treatment of intersectionality (for a review on the importance of intersectionality, see Ramasubramanian & Banjo, 2020; see also Moore-Berg & Karpinski, 2019). Here, we examined how non-Muslim Americans think about Muslims as a whole, yet Muslims (broadly speaking) are highly heterogenous. Therefore, future research should incorporate intersectional approaches into their intervention design. Finally, although MTurk samples can mirror other online (Chandler et al., 2019) and face-to-face samples (Casler et al., 2013), and have been used in other similar intervention tournament experiments (Bruneau et al., 2018; Gallardo et al., 2021), these

samples can be subject to data quality issues (Kennedy et al., 2020). Using various techniques, including preventing MTurk workers from participating in more than one study and including attention check questions, we attempted to maintain the high quality of data. Still, we encourage future researchers to consider replicating these results with nationally representative samples.

Conclusion

We employed an intervention tournament to identify media interventions that most successfully reduced Islamophobia, and then identified the mechanisms through which the successful videos worked. Of the 11 videos assessed, we found that a Muslim woman calling out media bias against Muslims served as the most successful intervention in changing anti-Muslim policy support and support for punitive community action measures, with effects that endured for at least 1 month. As predicted, even though only a small portion of the video addressed media bias, acknowledgment of media bias indirectly affected the relationship between the video on the outcomes. This research illustrates the importance of recognizing the structural inequalities against Muslims that exist in the media and highlights how drawing attention to such biases can effectively reduce Islamophobia.

Acknowledgements

The authors would like to thank Rachel Brown for her assistance collecting videos for the video tournament. We also would like to thank Lee-Or Ankori-Karlinsky for helpful conversations and feedback on this research.

Data accessibility

All data, syntax, and preregistrations can be viewed at the Open Science Framework (OSF) website (https://osf.io/5jmn3/?view_only=9d1b104a68d642d780823ccecc9a3bbd).

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Samantha L. Moore-Berg https://orcid.org/0000-0003-2972-2288
Boaz Hameiri https://orcid.org/0000-0002-0241-9839

Supplemental material

Supplemental material for this article is available online.

Notes

1. We also collected data for symbolic threat, realistic threat, parochial empathy, feeling thermometer, dehumanization, meta-dehumanization, common Muslim tropes, social dominance orientation, and right-wing authoritarianism; however, these are beyond the scope of this research and are not reported in this manuscript.
2. To increase data quality, we included an attention check and filter on MTurk that prevented participants from taking part in more than one study. We also set a minimum hit approval rate of $> 95\%$ and set the number of HITs approved to over 100.
3. The attrition rate between Time 1 and Time 2 was 16% ($n = 336$). Neither condition nor demographic variables—gender, ethnicity, religion, socioeconomic status, education, or country of origin—were significant predictors of attrition (all $ps > .140$).
4. All effects remained when controlling for religion in all studies. There was also no Condition x Religion interaction on any outcome variable for any study (see supplemental material).
5. We assessed additional exploratory measures that are not the focus of this research: Study 4a: collective blame, feeling thermometer, two measures of dehumanization, and empathy; Study 4b: empathy toward the main character, video-elicited emotions, feeling thermometer, and two measures of dehumanization; Study 4c: parochial empathy, feeling thermometer, and two measures of dehumanization.
6. When we examined both mini meta-analyses with Time 2 means from Study 1, the effects of the Media Bias and Identity Overlap 2 interventions on both anti-Muslim policy support and punitive community action held, while the effects of the Identity Overlap 3 intervention on both outcome measures became nonsignificant.

References

Allen, M., D'Alessio, D., Emmers, T. M., & Gebhardt, L. (1996). The role of educational briefings in

mitigating effects of experimental exposure to violent sexually explicit material: A meta-analysis. *The Journal of Sex Research*, *33*, 135–141. https://doi.org/10.1080/00224499609551825

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

American Civil Liberties Union. (2018). *Timeline of the Muslim ban*. https://www.aclu-wa.org/pages/timeline-muslim-ban

Bar-Tal, D., & Hameiri, B. (2020). Interventions for changing well-anchored attitudes in the context of intergroup relations. *Social and Personality Psychology Compass*, *14*, Article e12534. https://doi.org/10.1111/spc3.12534

Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations. *Social Issues and Policy Review*, 3(1), 141–177. https://doi.org/10.1111/j.1751-2409.2009.01013.x

Bayrakli, E., & Hafez, F. (Eds.). (2017). *European Islamophobia report*. Seta Foundation for Political, Economic and Social Research. https://www.islamophobiaeurope.com/wp-content/uploads/2018/04/EIR_2017.pdf

Bergsma, L. J., & Carney, M. E. (2008). Effectiveness of health-promoting media literacy education: A systematic review. *Health Education Research*, *23*, 522–542. https://doi.org/10.1093/her/cym084

Bleich, E., & van der Veen, A. M. (2021). Media portrayals of Muslims: A comparative sentiment analysis of American newspapers, 1996–2015. *Politics, Groups, and Identities*, *9*, 20–39. https://doi.org/10.1080/21565503.2018.1531770

Bruneau, E. G., Cikara, M., & Saxe, R. (2017). Parochial empathy predicts reduced altruism and the endorsement of passive harm. *Social Psychological & Personality Science*, 8(8), 934–942. https://doi.org/10.1177/1948550617693064

Bruneau, E., Kteily, N., & Falk, E. (2018). Interventions highlighting hypocrisy reduce collective blame of Muslims for individual acts of violence and assuage anti-Muslim hostility. *Personality and Social Psychology Bulletin*, *44*, 430–448. https://doi.org/10.1177/0146167217744197

Cantor, J., & Wilson, B. J. (2003). Media and violence: Intervention strategies for reducing aggression. *Media Psychology*, *5*, 363–403. https://doi.org/10.1207/S1532785XMEP0504_03

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*, 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009

Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond Mechanical Turk. *Behavior Research Methods*, *51*, 2022–2038. https://doi.org/10.3758/s13428-019-01273-7

Crisp, R. J., & Beck, S. R. (2005). Reducing intergroup bias: The moderating role of ingroup identification. *Group Processes & Intergroup Relations*, *8*, 173–185. https://doi.org/10.1177/1368430205051066

Davidson, B. (2017). Storytelling and evidence-based policy: Lessons from the grey literature. *Palgrave Communications*, *3*, Article 17093. https://doi.org/10.1057/palcomms.2017.93

Dixon, T. L. (2006). Psychological reactions to crime news portrayals of Black criminals: Understanding the moderating roles of prior news viewing and stereotype endorsement. *Communication Monographs*, *73*, 162–187. https://doi.org/10.1080/03637750600690643

Dixon, T. L., & Linz, D. (2000). Overrepresentation and underrepresentation of African Americans and Latinos as lawbreakers on television news. *Journal of Communication*, *50*, 131–154. https://doi.org/10.1111/j.1460-2466.2000.tb02845

Dovidio, J. F., Gaertner, S. L., & Validzic, A. (1998). Intergroup bias: Status, differentiation, and a common in-group identity. *Journal of Personality and Social Psychology*, *75*, 109–120. https://doi.org/10.1037/0022-3514.75.1.109

Dovidio, J. F., Love, A., Schellhaas, F. M. H., & Hewstone, M. (2017). Reducing intergroup bias through intergroup contact: Twenty years of progress and future directions. *Group Processes & Intergroup Relations*, *20*, 606–620. https://doi.org/10.1177/1368430217712052

Federal Bureau of Investigation (FBI). (2019). *2017 hate crime statistics*. https://ucr.fbi.gov/hate-crime/2017

Figueroa-Caballero, A., & Mastro, D. (2019). Examining the effects of news coverage linking undocumented immigrants with criminality: Policy and punitive implications. *Communication Monographs*, *86*, 46–67. https://doi.org/10.1080/03637751.2018.1505049

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model*. Psychology Press.

Gallardo, R., Hameiri, B., Moore-Berg, S. L., & Bruneau, E. G. (2021). The collective praise intervention: A

brief intervention highlighting prosocial behavior reduces hostility towards Muslims. *Group Processes & Intergroup Relations*. Advance online publication. https://doi.org/10.1177/13684302211021665

Gilliam, F. D., & Iyengar, S. (2000). Prime suspects: The influence of local television news on the viewing public. *American Journal of Political Science*, *44*, 560–573. https://doi.org/10.2307/2669264

Gilliam, F. D., Valentino, N. A., & Beckmann, M. N. (2002). Where you live and what you watch: The impact of racial proximity and local television news on attitudes about race and crime. *Political Research Quarterly*, *55*, 755–780. https://doi.org/10.1177/106591290205500402

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, *10*, 535–549. https://doi.org/10.1111/spc3.12267

Graham-McLay, C. (2019, May 2). Death toll in New Zealand mosque shootings rises to 51. *The New York Times*. https://www.nytimes.com/2019/05/02/world/asia/new-zealand-attack-death-toll.html

Hameiri, B., & Moore-Berg, S. L. (in press). Intervention tournaments: An overview of concept, design, and implementation. *Perspectives on Psychological Science*.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Hornik, R. C. (2002). *Exposure: Theory and evidence about all the ways it matters. Social Marketing Quarterly,* 8(3), 31-37. https://doi.org/10.1080/15245000214135.

Jeong, S., Cho, H., & Hwang, Y. (2012). Media literacy interventions: A meta-analytic review. *Journal of Communication*, *62*, 454–472. https://doi.org/10.1111/j.1460-2466.2012.01643.x

Kearns, E. M., Betus, A. E., & Lemieux, A. F. (2019). Why do some terrorist attacks receive more media attention than others? *Justice Quarterly*, *36*, 985–1022. https://doi.org/10.1080/07418825.2018.1524507

Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. G. (2020). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods*, *8*, 614–629. https://doi.org/10.1017/psrm.2020.6

Kishi, K. (2017). *Assaults against Muslims in U.S. surpass 2001 level*. Pew Research Center. https://www.pewresearch.org/fact-tank/2017/11/15/assaults-against-muslims-in-u-s-surpass-2001-level/

Kteily, N., & Bruneau, E. (2017). Darker demons of our nature: The need to (re)focus attention on blatant forms of dehumanization. *Current Directions in Psychological Science*, *26*, 487–494. https://doi.org/10.1177/0963721417708230

Kunst, J., Thomsen, L., Sam, D., & Berry, J. (2015). "We are in this together": Common group identity predicts majority members' active acculturation efforts to integrate immigrants. *Personality and Social Psychology Bulletin*, *41*, 1438–1453. https://doi.org/10.1177/0146167215599349

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., . . . Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*, 1765–1785. https://doi.org/10.1037/a0036260

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*, 1001–1016. https://doi.org/10.1037/xge0000179

Lipka, M. (2014). *How many people of different faiths do you know?* Pew Research Center. https://www.pewresearch.org/fact-tank/2014/07/17/how-many-people-of-different-faiths-do-you-know/

Mastro, D., Lapinski, M., Kopacz, M., & Behm-Morawitz, E. (2009). The influence of exposure to depictions of race and crime in TV news on viewer's social judgments. *Journal of Broadcasting & Electronic Media*, *53*, 615–635. https://doi.org/10.1080/08838150903310534

Matthes, J., Schmuck, D., & von Sikorski, C. (2019). Terror, terror everywhere? How terrorism news shape support for Anti-Muslim policies as a function of perceived threat severity and controllability. *Political Psychology*, 40(5), 935–951. https://doi.org/10.1111/pops.12576

Media Tenor. (2013). *U.S. TV primetime news prefer stereotypes: Muslims framed mostly as criminals*. http://

us.mediatenor.com/en/library/speeches/259/us-tv-primetimenews-prefer-stereotypes

Miller, E. (2017). *Ideological motivations of terrorism in the United States, 1970–2016*. START Nacional Consortium for the Study of Terrorism and Responses to Terrorism; University of Maryland. https://www.start.umd.edu/pubs/START_IdeologicalMotivationsOfTerrorismInUS_Nov2017.pdf

Moore-Berg, S. L., & Karpinski, A. (2019). An intersectional approach to understanding how race and social class affect intergroup processes. *Social Psychology and Personality Compass*, 13, e12426. https://doi.org/10.1111/spc3.12426

New America. (2018). *Survey examines perceptions of Muslim Americans*. https://www.newamerica.org/muslim-diaspora-initiative/press-releases/survey-examines-perceptions-muslim-americans/

Ogan, C., Willnat, L., Pennington, R., & Bashir, M. (2014). The rise of anti-Muslim prejudice: Media and Islamophobia in Europe and the United States. *The International Communication Gazette*, *76*, 27–46. https://doi.org/10.1177/1748048513504048

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, *60*, 339–367. https://doi.org/10.1146/annurev.psych.60.110707.163607

Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, *72*, 533–560. https://doi.org/10.1146/annurev-psych-071620-030619

Perloff, R. M. (2010). *The dynamics of persuasion: Communication and attitudes in the 21st century* (4th ed.). Routledge.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pilkington, E. (2017, December 30). Anti-sharia laws proliferate as Trump strikes hostile tone toward Muslims. *The Guardian*. https://www.theguardian.com/us-news/2017/dec/30/anti-sharia-laws-trump-muslims

Prati, F., Vasiljevic, M., Crisp, R. J., & Rubini, M. (2015). Some extended psychological benefits of challenging social stereotypes: Decreased dehumanization and a reduced reliance on heuristic thinking. *Group Processes & Intergroup Relations,* 18(6), 801–816. https://doi.org/10.1177/1368430214567762

Ramasubramanian, S., & Banjo, O. O. (2020). Critical media effects framework: Bridging critical cultural communication and media effects through power, intersectionality, context, and agency. *Journal of Communication*, *70*, 379–400. https://doi.org/10.1093/joc/jqaa014

Riek, B., Mania, E., Gaertner, S., McDonald, S., & Lamoreaux, M. (2010). Does a common ingroup identity reduce intergroup threat? *Group Processes & Intergroup Relations*, *13*, 403–423. https://doi.org/10.1177/1368430209346701

Saleem, M., Prot, S., Anderson, C. A., & Lemieux, A. F. (2017). Exposure to Muslims in media and support for public policies harming Muslims. *Communication Research*, *44*, 841–869. https://doi.org/10.1177/0093650215619214

Saleem, M., & Ramasubramanian, S. (2017). Muslim Americans' responses to social identity threats: Effects of media representations and experiences of discrimination. *Media Psychology*, *22*, 373–393. https://doi.org/10.1080/15213269.2017.1302345

Scharrer, E., & Ramasubramanian, S. (2015). Intervening in the media's influence on stereotypes of race and ethnicity: The role of media literacy education. *Journal of Social Issues*, *71*, 171–185. https://doi.org/10.1111/josi.12103

Shaver, J. H., Sibley, C. G., Osborne, D., & Bulbulia, J. (2017). News exposure predicts anti-Muslim prejudice. *PLoS One*, *12*, Article e0174606. https://doi.org/10.1371/journal.pone.0174606

Sides, J., & Gross, K. (2013). Stereotypes of Muslims and support for the war on terror. *The Journal of Politics*, *75*, 583–598. https://doi.org/10.1017/S0022381613000388

Slater, M. D. (2007). Reinforcing spirals: The mutual influence of media selectivity and media effects and their impact on individual behavior and social identity. *Communication Theory*, *17*, 281–303. https://doi.org/10.1111/j.1468-2885.2007.00296.x

Swann, W. B., Jr., Gómez, Á., Seyle, D., Morales, J. F., & Huici, C. (2009). Identity fusion: The interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology*, *96*, 995–1011. https://doi.org/10.1037/a0013668

Transue, J. E. (2007). Identity salience, identity acceptance, and racial policy attitudes: American national identity as a uniting force. *American Journal of Political Science*, *51*, 78–91. https://doi.org/10.1111/j.1540-5907.2007.00238.x

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Re-discovering the social group: A self-categorization theory*. Blackwell.

Valentino, N. A. (1999). Crime news and the priming of racial attitudes during evaluations of the president. *Public Opinion Quarterly*, *63*, 293–320. https://doi.org/10.1086/297722

Valentino, N. A. (2001). The mass media and group priming in American elections. In R. P. Hart & D. R. Shaw (Eds.), *Communication in U.S. elections: New agendas* (pp. 145–166). Rowman & Littlefield Publishers.

Wirtz, C., & Doosje, B. (2013). Reactions to threatening critical messages from minority group members with shared or distinct group identities. *European Journal of Social Psychology*, *43*, 50–61. https://doi.org/10.1002/ejsp.1910