



Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Case Report

Harnessing dehumanization theory, modern media, and an intervention tournament to reduce support for retributive war crimes[☆]

Alexander P. Landry^{a,*}, Katrina Fincher^b, Nathaniel Barr^c, Nicholas P. Brosowsky^d, John Protzko^e, Dan Ariely^f, Paul Seli^f

^a Department of Organizational Behavior, Stanford University, Stanford, CA 94305, United States of America

^b Department of Psychology, New School for Social Research, New York, NY 10011, United States of America

^c Department of Psychology, Sheridan College, Oakville, ON L6H2L1, Canada

^d Department of Psychology, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

^e Central Connecticut State University, New Britain, CT 06053, United States of America

^f Department of Psychology, Duke University, Durham, NC 27708, United States of America



ARTICLE INFO

Keywords:

Dehumanization
Retributive violence
Psychological interventions
Intervention tournament
Russo-Ukrainian war

ABSTRACT

We demonstrate how psychological scientists can curate rich-yet-accessible media to intervene on conflict-escalating attitudes during the earliest stages of violent conflicts. Although wartime atrocities all-too-often ignite destructive cycles of tit-for-tat war crimes, powerful third parties can de-escalate the bloodshed. Therefore, following Russia's illegal invasion of Ukraine, we aimed to reduce Americans' support for committing retributive war crimes against Russian soldiers. To intervene during the earliest stages of the invasion, we drew on theories of dehumanization and "parasocial" intergroup contact to curate publicly available media expected to humanize Russian soldiers. We then identified the most effective materials by simultaneously evaluating all of them with an *intervention tournament*. This allowed us to quickly implement a psychological intervention that reliably reduced support for war crimes during the first days of a momentous land war. Our work provides a practical, result-driven model for developing psychological interventions with the potential to de-escalate incipient conflicts.

During World War II, American servicemen extracted ears, teeth, and skulls from the corpses of Japanese troops as "trophies" (Weingartner, 1992). On the other side of the world, soldiers of the USSR were executing German POWs en masse (Edele, 2016). More recently, American soldiers mutilated al-Qaeda prisoners at Abu Ghraib (S. T. Fiske, Harris, & Cuddy, 2004) while Iraqi armed forces did the same to captured ISIS insurgents (Abdul-Ahad, 2017). What these far-flung atrocities—and innumerable others—have in common is that they were committed by soldiers who were retaliating against an unprovoked attack: Japan's bombing of Pearl Harbor, Germany's invasion of the Soviet Union, al-Qaeda's attack on the World Trade Center, and ISIS's occupation of Iraq all aroused powerful human motives for retribution (Carlsmith & Darley, 2008; Jackson, Choi, & Gelfand, 2019; Schumann & Ross, 2010).

Indeed, unjust military attacks all-too-often ignite destructive cycles of tit-for-tat war crimes, so it is crucial to intervene during the earliest

stages of these invasions. Thus, we aimed to reduce support for retributive war crimes in the immediate aftermath of an unjust military invasion. Although psychological interventions have promoted peace in contexts of protracted violence, they are relatively time-intensive and thus ill-suited for swiftly de-escalating incipient conflicts. Drawing on theories of dehumanization and "parasocial" intergroup contact, we surmount this limitation by harnessing rich, readily available online media intended to humanize the invading soldiers. We then identified the most effective material with an *intervention tournament*—a standardized procedure to evaluate several potential interventions at once. This allowed us to develop a humanizing intervention that dampened support for retributive war crimes during the first weeks of a momentous land war, providing a model for de-escalating incipient violence.

[☆] This paper has been recommended for acceptance by Dr Karina Schumann.

* Corresponding author.

E-mail address: alandry@stanford.edu (A.P. Landry).

<https://doi.org/10.1016/j.jesp.2023.104567>

Received 14 June 2023; Received in revised form 15 November 2023; Accepted 15 November 2023

Available online 25 November 2023

0022-1031/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Retributive violence and conflict escalation

Violent atrocities can ignite escalating cycles of reciprocal bloodshed (Collins, 2012). Over time, groups engaging in reciprocal violence can become locked into an intractable conflict with destructive consequences (Azar, 1990; Bar-Tal, 2007; Coleman, 2014). For instance, during World War II, American atrocities in the Pacific Theater inspired reciprocal brutality by Japanese troops such as the Bataan Death March (Dower, 2012). This only intensified American hatred of the Japanese and eventually contributed to the atomic bombing of nearly 300,000 Japanese civilians (Weingartner, 1992). In explaining this decision to the American public, then-President Harry Truman offered the compelling logic of retribution: “The Japanese began the war from the air at Pearl Harbor. They have been repaid manyfold” (quoted in Goldhagen, 2009). In other cases, retributive violence has erupted into civil war, ethnic cleansing, and outright genocide (Atran & Ginges, 2012; A. P. Fiske & Rai, 2014). Thus, when a conflict has the potential to ignite reciprocal atrocities, it is critical to intervene in its incipient stages.

2. Third-party support for retributive war crimes

People often want to punish moral transgressions—even those not directly harmed by the transgression (Fehr & Fischbacher, 2004; Henrich et al., 2006). When the transgression is particularly egregious, third parties may support retaliating with extreme violence (Foucault, 2023; Garland, 1993). For instance, unprovoked military attacks can incite third-party support for retributive war crimes against the aggressors (e.g., support for the defending army executing captured soldiers of the invading army; Sagan & Valentino, 2019; Watkins & Goodwin, 2020). Moreover, these retributive motives can spill over to target uninvolved members of the transgressor’s group (Lieberman & Skitka, 2019; Lickel, Miller, Stenstrom, Denson, & Schmader, 2006). Indeed, respondents not only supported retributive war crimes against soldiers, but also endorsed slaughtering civilians of the invading country (Sagan & Valentino, 2019; Watkins & Goodwin, 2020). Likewise, American observers who experienced moral outrage at Iraq’s unprovoked invasion of Kuwait supported retaliating by indiscriminately bombing Iraqi civilians (Lieberman, 2006; see also Liberman & Skitka, 2019, 2017).

Third-party attitudes toward war crimes are consequential because external actors can influence the course of violent conflicts (Dixon, 1996; Kydd & Straus, 2013; Rosenberg, Galis, & Zucker, 2016). Specifically, war crimes can be reduced if third-party nations, international bodies, NGOs, and the media raise the costs of committing them (e.g., through formal economic sanctions or informal “naming and shaming”; Dancy, 2016; Krain, 2012; Sikink, 2011). At the same time, belligerents may be emboldened to commit atrocities if they perceive a permissive international community (Goldhagen, 2009; Power, 2013; Rummel, 2017). Because the foreign policies of many third-party nations can be influenced by the sentiments of their citizens (Baum & Potter, 2015; Holsti, 2004), reducing these individuals’ support for retributive war crimes may help dampen incipient violence.

3. Dehumanization and retributive violence

Having good moral values is seen to indicate that one has transcended their selfish animal nature to develop “uniquely human” sensibilities (Haslam, 2006; Schwartz & Struch, 1989). Indeed, people consider having an intrinsically good moral compass essential to being human (Phillips, 2022; Strohminger & Nichols, 2014). Thus, people often overtly deny the humanity of those who commit egregious acts of harm (Bastian, Denson, & Haslam, 2013; Kelman, 1973; Phillips, 2022)—including groups who perpetrate unjust military attacks (Sagan & Valentino, 2017; Slovic, Mertz, Markowitz, Quist, & Västfjäll, 2020; Watkins & Laham, 2020). This explicit, blatant *dehumanization* imbues the perpetrators with the despicable qualities of “lower”

animals—signaling that they have a depraved nature falling so far beneath “civilized” humanity as to reveal a subhuman moral essence (Fincher, Kteily, & Bruneau, 2018; Kteily & Landry, 2022). Because of their depraved essence, the perpetrators are deemed incapable or unwilling of changing their evil ways—and thus threaten the ingroup’s fundamental values or even its very existence (Bar-Tal, 1989; Smith, 2021). Thus, engaging in retributive violence against them may appear necessary and even righteous (Kelman, 1973; Kteily & Landry, 2022).¹

Evidence supports the notion that dehumanization can incite retributive violence. For instance, many Americans reacted to an Iranian military massacre by dehumanizing Iranians, which predicted their support for bombing Iranian civilians (Sagan & Valentino, 2017; Slovic et al., 2020). Likewise, those who dehumanized Muslims supported retaliating against an Islamic terror strike by “wiping [every Muslim] off this world” (Kteily, Bruneau, Waytz, & Cotterill, 2015). Most relevant to the present research, third-party observers to war crimes committed during the Syrian civil war dehumanized the perpetrator group, which predicted their support for retributive war crimes (e.g., torturing captured fighters and killing civilian members of the group; Rousseau, Gorman, & Baranik, 2023; see also Bastian et al., 2013; Leidner, Castano, & Ginges, 2013).

3.1. Additional predictors of retributive violence: Dislike and ideological orientations

Despite this past research, some have questioned the link between dehumanization and retributive violence (Rai, Valdesolo, & Graham, 2017). Specifically, scholars have noted that the explicit, blatant denial of an outgroup’s humanity is often intertwined with intense negative affect toward them, and raised concerns that such “dehumanization” instead simply reflects *dislike* of the outgroup (Bloom, 2022; Over, 2021b; Rai et al., 2017). We maintain that dehumanization goes beyond “mere” dislike: by stripping its target of the moral elevation afforded to fellow human beings, dehumanization may render them particularly vulnerable to extreme harm (Fincher et al., 2018; Kteily & Landry, 2022). Nonetheless, the onus remains on dehumanization scholars to empirically demonstrate how these constructs are meaningfully distinct (Over, 2021a). Indeed, despite mounting evidence that dehumanization and dislike are separable (reviewed by Kteily & Landry, 2022), they generally converge in their relationships with outgroup hostility (e.g., desire for social distance or support for discriminatory policies; Kteily et al., 2015; Moore-Berg, Ankori-Karlinsky, Hameiri, & Bruneau, 2020). Therefore, along with dehumanization, in our studies we also accounted for dislike to examine their relative effects on retributive war crimes.

In examining predictors of support for retributive war crimes, we also considered three ideologies—beliefs about the proper social order and how it should be achieved (Jost, Federico, & Napier, 2009). *Social dominance orientation* (SDO) reflects an aversion to egalitarianism and motivation to enforce hierarchical differentiation between groups (Ho et al., 2015; Pratto, Sidanius, Stallworth, & Malle, 1994). *Right-wing authoritarianism* (RWA) reflects a tendency to perceive the world as dangerous, submit to authority, and rigidly adhere to social norms (Altemeyer, 1988). *Ideological conservatism* reflects a dispositional resistance to change and preference for the status-quo (Jost, Glaser, Kruglanski, & Sulloway, 2003). Despite their important differences, those higher in these ideologies tend to be more punitive, intolerant, and aggressive toward outgroups (Altemeyer, 1988; Nilsson & Jost, 2020; Pratto et al., 1994). Indeed, SDO, RWA, and conservatism have all been

¹ It is important to emphasize that here, we conceptualize dehumanization as the explicit, blatant denial of an outgroup’s humanity (see Kteily & Landry, 2022). There are also subtler forms of dehumanization (e.g., overlooking another’s mental capacities) but these appear to be of less relevance to extreme violence (Rai et al., 2017; see also Study 1 in the “Deviations From Preregistration” section of the Supplementary Information).

linked to retributive violence—including retaliating against military attacks by indiscriminately killing civilians (Kteily et al., 2015; Sagan & Valentino, 2017; Slovic et al., 2020).

4. The present research

On 24 February 2022, Russia invaded Ukraine in the largest attack on a European country since World War II. This violation of national sovereignty was widely considered a heinous act of aggression and ignited widespread moral outrage (Cohen, 2022). We expected that many observers would react to the invasion by dehumanizing Russians, which would motivate their support for the Ukrainians retaliating with war crimes of their own. Thus, we aimed to reduce third-party support for retributive war crimes by humanizing Russian soldiers. Moreover, we aimed to intervene during the earliest stages of the invasion—before it could ignite cycles of reciprocal atrocities. We focused on the United States—a powerful nation that was positioned to play a pivotal role in shaping the course of the war from its earliest days (U.S. Security Cooperation with Ukraine, 2022). Specifically, because US foreign policy is influenced to a large extent by public opinion (Holsti, 2004; Tomz, Weeks, & Yarhi-Milo, 2020), we focused on US citizens' support for retributive war crimes. As a first step, we conducted Study 1 in the week following the invasion to test our assumptions that (a) many US citizens would dehumanize Russian soldiers and (b) dehumanization would robustly predict their support for retributive war crimes.

5. Study 1

5.1. Method

5.1.1. Participants

Study 1 was conducted on March 3–4, 2022—the week following the Russian invasion. Assuming the relationships between our variables would stabilize around 250 observations (Schönbrodt & Perugini, 2013), and oversampling to account for potential data exclusions (see Procedure), we recruited 352 Amazon Mechanical Turk (MTurk) workers who were pre-screened for data quality by the participant recruitment platform CloudResearch (Litman & Robinson, 2020; Litman, Rosenzweig, & Moss, 2020).² Participants who reported that they were not US citizens in an opening demographics block (see Procedure) were immediately directed to the end of the survey. We also excluded participants who failed either of two attention checks embedded in the survey, leaving us with a final sample of 338 participants ($M_{\text{age}} = 38.91$, $SD_{\text{age}} = 11.50$; 72.2% White, 12.1% Black, 7.1% Asian, 7.1% Hispanic, 1.5% other; 60.7% male, 38.8% female, 0.6% other; 53.8% Democrat, 23.7% Republican, 21.6% Independent, 0.9% other). With this sample size, we were sensitive to detect an effect of $d = 0.39$ with 80% power in our primary regression model with five predictors (Faul, Erdfelder, Lang, & Buchner, 2007).

5.1.2. Measures

5.1.2.1. Conservatism. We measured conservatism by asking “In general, do you consider yourself liberal or conservative?” which participants responded to on a 1 (*Extremely Liberal*) to 7 (*Extremely Conservative*) Likert scale.

² Subsequent studies conducted with MTurk workers also recruited quality-approved participants via CloudResearch who had not participated in any prior studies.

5.1.2.2. Dehumanization. As is common in this line of work, we assessed dehumanization as the composite of two measures (e.g., Bruneau, Casas, Hameiri, & Kteily, 2022; Bruneau, Hameiri, Moore-Berg, & Kteily, 2021; Kteily et al., 2015; Kteily, Hodson, & Bruneau, 2016).³ One was the Ascent of Man scale (see Supplementary Fig. 1). Participants rated the humanity of Russian soldiers using a slider scale ranging from 0 to 100. Scores were reversed such that higher scores indicated greater dehumanization. The other measure asked participants to rate how well eight blatantly dehumanizing traits (“savage”, “like animals”, “lacking self-restraint”, “unsophisticated”, “barbaric”, “backward”, “primitive”, and “uncivilized”) described Russian soldiers on a Likert scale ranging from 1 (*Not at all*) to 7 (*Extremely*; $\alpha = 0.97$). We created the dehumanization composite by standardizing these two measures and averaging them together ($r(338) = 0.63$, $p < .001$).

5.1.2.3. Dislike. We used a standard feeling thermometer to measure Americans' dislike of Russian soldiers (Haddock, Zanna, & Esses, 1993; see Supplementary Fig. 2). Participants rated how they felt toward Russian soldiers using a slider scale ranging from 0 (*extremely cold*) to 10 (*extremely warm*). Scores were reversed such that higher scores indicated greater dislike.

5.1.2.4. Retributive war crimes. We measured support for retributive war crimes with the following six items: “Even though it goes against the rules of conventional warfare, Ukrainians should torture invading Russian soldiers for information about their military operations”, “The Ukrainians should not be bound by the rules of conventional warfare when fighting the invading Russian soldiers”, “Even though it goes against the rules of conventional warfare, the Ukrainians should use poisonous gases to kill invading Russian soldiers”, “Even though it goes against the rules of conventional warfare, the Ukrainians should use biological weapons to kill invading Russian soldiers”, “Even though it goes against the rules of conventional warfare, the Ukrainians should execute invading Russian soldiers who are unarmed”, and “The Ukrainians should try to annihilate the invading Russian soldiers by any means necessary” ($\alpha = 0.94$). These six items formed a single factor explaining 66.18% of the variance.

5.1.2.5. Social dominance orientation. We measured SDO with four items from Ho et al. (2015) on a 1 (*Strongly disagree*) to 6 (*Strongly agree*) Likert scale: “Group equality should not be our primary goal”, “It is unjust to try to make groups equal”, “An ideal society requires some groups to be on top and others to be on the bottom”, and “Some groups of people are simply inferior to other groups” ($\alpha = 0.92$).

5.1.2.6. Right-wing authoritarianism. We measured RWA with three items from Bizumic and Duckitt (2018) on a 1 (*Strongly disagree*) to 6 (*Strongly agree*) Likert scale: “What our country needs most is discipline, with everyone following our leaders in unity”, “Our society needs tougher government and stricter laws”, and “The facts on crime and the recent public disorders show we have to crack down harder on troublemakers if we are going to preserve law and order” ($\alpha = 0.90$).

5.1.3. Procedure

Participants first provided demographic information (including citizenship and conservatism). Those who reported not being US citizens were then immediately directed to the end of the survey. Remaining participants then reported their dehumanization and dislike of Russian soldiers (counterbalanced), and then their support for retributive war

³ Nonetheless, here and throughout, all patterns of results remain unchanged when considering the dehumanization measures separately (see Supplementary Information: “Analyses Considering Ascent and Animalistic Trait Measures Separately”).

crimes against Russian soldiers. Finally, they completed the SDO and RWA measures (counterbalanced).

5.2. Results

All analyses reported in this manuscript were performed with 1000 bootstrapped resamples, except for mediation analyses which were performed with 5000 bootstrapped resamples. Values in brackets indicate 95% CIs.

Table 1 presents descriptive statistics and intercorrelations among all variables, while Fig. 1 depicts the distributions of scores on both measures of dehumanization (i.e., Ascent of Man and Animalistic Traits separately, instead of the standardized composite) and war crime support.

Dehumanization was relatively high on both measures (Ascent of Man: $M = 22.98$, $SD = 28.80$; Animalistic Traits: $M = 3.52$, $SD = 1.83$). For instance, 64% of the sample ($n = 216$) explicitly rated Russian soldiers as less than human on the Ascent of Man scale (see Supplementary Fig. 1), while endorsement of the blatantly dehumanizing traits (e.g., “savage”, “like animals”) lay just below the scale midpoint. Moreover, one-third of the sample ($n = 112$) expressed at least slight support for unambiguous war crimes such as torturing and executing captured Russian soldiers ($M = 2.47$, $SD = 1.38$).

Consistent with expectations, dehumanization was strongly associated with war crime support, $r(338) = 0.50$, $p < .001$. To determine whether dehumanization explained unique variance in war crime support, over and above dislike and the ideological variables, we performed a three-step hierarchical linear regression.⁴ In the first step, we entered the ideological variables (SDO, RWA, and conservatism), $R^2 = 0.26$, $SE = 1.20$, $F(3, 334) = 38.54$, $p < .001$. In the second step, we added dislike to the model, $\Delta R^2 = 0.01$, $SE = 1.19$, $F(4, 333) = 30.94$, $p < .001$. In the third step, we added dehumanization to the model, $\Delta R^2 = 0.13$, $SE = 1.08$, $F(5, 332) = 44.08$, $p < .001$. In this full model, dehumanization emerged as the strongest predictor of war crime support, $b = 0.65$ [0.47, 0.85], $t(332) = 8.41$, $SE = 0.08$, $p < .001$ (see Table 2 for the full model output).

6. Study 2

Having observed normatively troubling levels of dehumanization and support for retributive war crimes—and a robust link between the two—we attempted to reduce war crime support by humanizing Russian soldiers. Given the potential for retributive violence to ossify into intractable cycles of bloodshed (Azar, 1990; Bar-Tal, 2007; Coleman, 2014), we aimed to rapidly develop a *humanization intervention* during the invasion’s earliest stages.

6.1. Developing a humanization intervention

Although perpetrators of egregious harm are often denied an essentially human moral compass (Kelman, 1973; Phillips, 2022), their expressions of moral agency or remorse can (re)humanize them (Borinca, Falomir-Pichastor, Andrighetto, & Halabi, 2021; Jin, Park, & Park, 2022; Khamitov, Rotman, & Piazza, 2016). Therefore, we aimed to expose Americans to Russian soldiers’ essentially human moral qualities following the invasion. Because of their vast geographical separation, we did so through immersive media content. Such *parasocial contact* with outgroups has proven effective because people process media similar to “direct” experiences (Horton & Wohl, 1956; Schiappa, Gregg,

⁴ As can be seen in Figure 1, the distributions of dehumanization and war crime support are skewed. Thus, we performed diagnostics of the regression model (e.g., plots of the residuals and tests of normality), which can be found in the Supplementary Information (“Regression Model Diagnostics”). We thank an anonymous reviewer for this suggestion.

& Hewes, 2005). Indeed, parasocial contact has been found to operate through similar processes as traditional intergroup contact to reduce dehumanization (e.g., by promoting empathy and correcting negative misperceptions; Bruneau et al., 2021, 2022; Moore-Berg, Hameiri, & Bruneau, 2022). Importantly, if the outgroup members are perceived to be representative of their group, the effects of (parasocial) contact with them can generalize to the entire outgroup (Brown & Hewstone, 2005; Rothbart & John, 1985). Thus, we focused on humanizing rank-and-file Russian soldiers in an effort to reduce support for retributive war crimes not only against them, but also “vicarious retribution” against uninvolved Russian civilians (Lieberman & Skitka, 2019; Lickel et al., 2006).

Although previous media interventions have targeted support for intergroup violence (Bilali, 2022; Hameiri, Porat, Bar-Tal, & Halperin, 2016), they were developed with relatively time-intensive methods that are ill-suited for rapidly intervening in incipient conflicts (e.g., recording months-long radio dramas; Bilali, 2022). Therefore, rather than developing our own intervention materials, we curated an array of existing media from a readily accessible online repository. We then identified the most effective content with an *intervention tournament*—a study that simultaneously tests several potential interventions against the same control condition on a set of standardized outcome measures (Hameiri & Moore-Berg, 2022). In doing so, we built on previous work that used intervention tournaments to test a range of curated media expected to reduce prejudice through various psychological processes (Bruneau, Kteily, & Falk, 2018; Moore-Berg, Hameiri, Falk, & Bruneau, 2022).

6.1.1. Curating humanizing media

We curated six short videos from an open-access online repository (YouTube), each of which depicted rank-and-file Russian soldiers expressing moral agency and/or remorse following the invasion. Videos 1 (2:18) and 2 (2:47) featured testimony from captured Russian soldiers discussing their opposition to the invasion and regret for participating in it. Likewise, Videos 3 (1:46) and 4 (0:53) each consisted of a Russian soldier displaying regret and sorrow while describing the invasion to their mother on the phone. Video 5 (1:22) featured a soldier expressing resentment at being treated like “cannon fodder” and opposition to how the Russian army was conducting the war. Video 6 (1:24) was a compilation of clips (including from Videos 1 and 3) that highlighted how many Russian soldiers opposed the invasion and were actively disobeying orders from Moscow.

However, the impact of testimony from the soldiers themselves may be limited because individuals often resist information from outgroup members if it contradicts their preexisting views about the group (Greenaway, Wright, Willingham, Reynolds, & Haslam, 2015). Therefore, we prefaced each video with an excerpt from a reputable news article that corroborated Russian soldiers’ moral agency (“[Russian soldiers] have abandoned their vehicles and have set off back towards the Russian border on foot”) and remorse (“there is an authentic sense that many Russian servicemen regret ever having come to Ukraine”; Harding, 2022).

The article excerpt and videos can be found at the OSF repository: https://osf.io/s83ye/?view_only=e2f7b2b1752a4de584700eb5e7d20b8c.

6.2. Method

6.2.1. Participants

On March 5, 2022, we recruited 430 MTurk workers via CloudResearch—aiming for approximately 60 participants per condition in order to detect small-to-medium main effects ($d = 0.35$) with 80% power in the one-way ANOVAs we conducted (Faul et al., 2007). After applying our exclusion criteria (see Procedure) we were left with a final sample of 406 participants ($M_{age} = 42.98$, $SD_{age} = 12.94$; 79.6% White, 6.2% Black, 8.4% Asian, 4.4% Hispanic, 1.4% other; 48.8% male, 50.7% female, 0.5% other; 50.5% Democrat, 24.6% Republican, 24.1% Independent, 0.7% other). This sample size enabled us to detect an effect of

Table 1
Descriptive Statistics and Intercorrelations.

Correlations	1	2	3	4	5
1. Dehumanization Composite	–				
2. Dislike	0.39***	–			
3. SDO	0.03	–0.29***	–		
4. RWA	0.34***	–0.10	0.44***	–	
5. Conservatism	0.04	–0.06	0.55***	0.39***	–
6. Retributive War Crimes	0.50***	0.03	0.29***	0.48***	0.12*
Descriptives	1	2	3	4	5
M	0.00	7.60	2.37	2.95	3.42
SD	0.90	2.42	1.41	1.36	1.83

Note. *** $p < .001$ * $p < .05$.

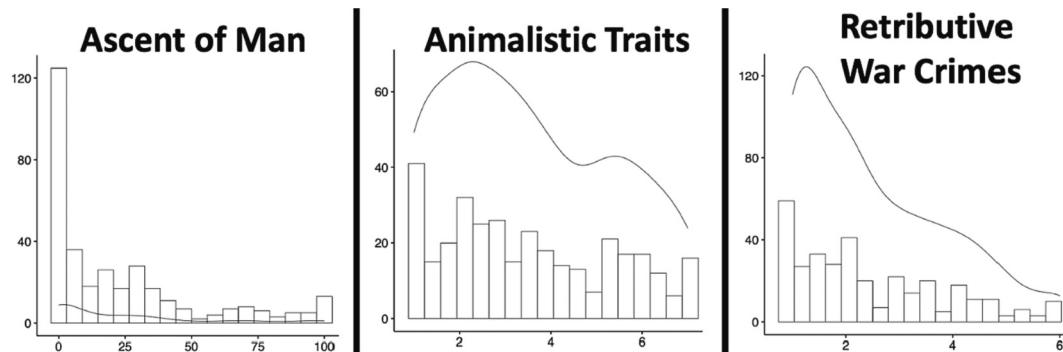


Fig. 1. Distributions of Scores on Both Dehumanization Measures (Ascent of Man and Animalistic Traits) and Support for Retributive War Crimes.

Table 2
Output for Regression Model Predicting Support for Retributive War Crimes.

	Unstandardized b	95% CI for b	Standard Error	t(332)	Standardized β	p	Variance Inflation Factor
Dehumanization	0.65	[0.47, 0.85]	0.10	8.41	0.43	< 0.001	1.42
Dislike	-0.03	[–0.08, 0.02]	0.03	–1.17	–0.06	0.188	1.37
SDO	0.20	[0.06, 0.33]	0.07	3.54	0.20	0.007	1.75
RWA	0.30	[0.19, 0.41]	0.06	5.64	0.30	< 0.001	1.52
Conservatism	-0.09	[–0.19, 0.00]	0.05	–2.35	–0.12	0.053	1.54

the planned contrasts between each of the intervention conditions and the control condition in the one-way ANOVAs we performed of $d = 0.28$ with 80% power (Faul et al., 2007).

6.2.2. Procedure

At the very beginning of the study, all participants were presented with a one-paragraph mock vignette to test their attention and reading comprehension (i.e., the content of the vignette was completely unrelated to the study). Those who failed a comprehension question about it were immediately directed to the end of the survey. Next, remaining participants completed a short block of demographics questions, including an item asking whether they were a US citizen. Those who did not report being a US citizen were then also directed to the end of the survey. The 406 remaining participants comprising our final sample were then randomly assigned to receive one of the six interventions or to an empty control condition. They then reported their dehumanization and dislike of Russian soldiers (counterbalanced), and then support for retributive war crimes against them (all measured as in Study 1; dehumanization: $r(406) = 0.66, p < .001$; retributive war crimes: $\alpha = 0.90$).

6.3. Results

Descriptive statistics for each condition are reported in Supplementary Table 7. Intercorrelations among measured variables are reported in Supplementary Table 8.

We tested for significant differences between each of the intervention conditions and the control condition. Specifically, we contrast-coded all video conditions against the control condition (i.e., the video condition was coded as +6 and each video condition was coded as –1). We then performed a series of one-way ANOVAs with dehumanization, dislike, and war crime support as the dependent variables. We observed significant effects on dehumanization, $t(399) = 4.85, d = 4.13, p < .001$, dislike $t(399) = 6.74, d = 5.73, p < .001$, and war crime support, $t(399) = 3.08, d = 2.62, p = .002$.⁵ Post-hoc tests with Tukey HSD corrections for multiple comparisons revealed that all interventions reduced dehumanization (Table 3) and dislike (Supplementary Table 9) relative to the control condition. Moreover, there were no significant differences between any of the intervention conditions on dehumanization or dislike. However, although all interventions led to a numeric decrease in war crime support relative to the control condition, only the effect of Video 1 approached significance in these post-hoc tests, $M_{Diff} = 0.53, [–0.02, 1.08], d = 0.52, SE = 0.19, p = .067$ (see Table 4). There were also no significant differences between any of the intervention conditions on support for war crimes.

Given our result-oriented focus, we selected the intervention (Video

⁵ Omnibus tests revealed main effects on dehumanization, $F(6, 399) = 4.05, \eta_p^2 = 0.06, p < .001$, and dislike, $F(6, 399) = 8.17, \eta_p^2 = 0.11, p < .001$, but not war crime support, $F(6, 399) = 1.76, \eta_p^2 = 0.03, p = .107$.

Table 3
Post-Hoc Comparisons of Interventions' Effect on Dehumanization Relative to the Control Condition.

	Mean Difference from Control	95% CI	d	Standard Error	p
Video 1 (n = 58)	0.63	[0.14, 1.12]	0.63	0.17	0.003
Video 2 (n = 58)	0.54	[0.05, 1.03]	0.50	0.17	0.022
Video 3 (n = 57)	0.63	[0.14, 1.12]	0.65	0.17	0.003
Video 4 (n = 59)	0.65	[0.17, 1.14]	0.65	0.16	0.002
Video 5 (n = 58)	0.65	[0.16, 1.14]	0.64	0.17	0.002
Video 6 (n = 58)	0.58	[0.09, 1.07]	0.58	0.17	0.009

Table 4
Post-Hoc Comparisons of Interventions' Effect on Support for Retributive War Crimes, Relative to the Control Condition.

	Mean Difference from Control	95% CI	d	Standard Error	p
Video 1 (n = 58)	0.53	[-0.02, 1.08]	0.52	0.19	0.067
Video 2 (n = 58)	0.35	[-0.20, 0.90]	0.32	0.19	0.502
Video 3 (n = 57)	0.45	[-0.11, 1.00]	0.45	0.19	0.204
Video 4 (n = 59)	0.40	[-0.15, 0.95]	0.37	0.19	0.308
Video 5 (n = 58)	0.43	[-0.12, 0.98]	0.41	0.19	0.238
Video 6 (n = 58)	0.46	[-0.09, 1.01]	0.41	0.19	0.173

1) that produced the largest decrease in war crime support relative to the control condition. The video from this “winning” intervention featured six captured Russian soldiers who described their experiences during the first days of the invasion, conveying clear moral opposition and remorse. Namely, they expressed bewilderment regarding the purpose of the invasion, resentment at being used as “cannon fodder” and “meat shields,” repugnance at their orders to attack “peaceful” Ukrainians, and guilt for having “acted like Nazis.” The efficacy of this video is consistent with theories of (im)morality’s central role in dehumanization (Brandt & Reyna, 2011; Phillips, 2022; Schwartz & Struch, 1989), and research that (re)humanized perpetrators by highlighting their moral capacities (Borinca et al., 2021; Jin et al., 2022; Khamitov et al., 2016). Moreover, this video was poignant, yet concise—totaling just over two minutes. Along with the short pre-video article excerpt (345 words), this represents a “light-touch” intervention: a brief treatment that is inexpensive and easy to implement (Paluck, Porat, Clark, & Green, 2021). The following day, we established the efficacy of this intervention in a proof-of-concept study (see Supplementary Information) and then moved to test it on a large, nationally representative sample.

7. Study 3

7.1. Method

7.1.1. Participants

From March 7–11, 2022, we recruited a nationally representative sample of Americans, matched to 2020 American National Election Survey demographics on age, ethnicity, and gender, through the ForthRight panel managed by Bovitz Inc. (see Supplementary Information for additional information about this panel). As noted above, we first tested this intervention in a proof-of-concept study (see

Supplementary Information). In that study, the lower bound of the 95% confidence interval for the effect of the intervention on retributive war crimes was $d = 0.17$. Therefore, we aimed to recruit at least 1500 participants to detect an effect of this magnitude with 95% power. We initially recruited 1994 participants, and after applying our preregistered exclusion criteria (see Procedure) we were left with a final sample of 1513 ($M_{age} = 46.45$, $SD_{age} = 16.32$; 66.6% White, 12% Black, 12% Hispanic, 5.8% Asian; 49.6% male, 48.9% female; 46.3% Democrat, 25.7% Republican, 24.2% Independent, 3.7% other; see Supplementary Table 10 for a comparison to 2020 ANES demographic benchmarks). This sample size enabled us to detect main effects of our intervention in the independent-samples t -tests we performed of $d = 0.14$ with 80% power (Faul et al., 2007).

7.1.2. Measures

Dehumanization, dislike, and support for retributive war crimes against Russian soldiers were measured as in previous studies (dehumanization: $r(1,513) = 0.64$, $p < .001$; war crimes against soldiers: $\alpha = 0.89$). Along with support for war crimes against Russian soldiers, we also tested whether, consistent with (parasocial) contact theory (Brown & Hewstone, 2005; Rothbart & John, 1985), humanizing these rank-and-file soldiers generalized to reduce support for committing war crimes against Russian civilians.

7.1.2.1. War crimes against civilians. We measured Americans’ support for war crimes against Russian civilians with five items on a 1–6 Likert scale (*Strongly disagree–Strongly agree*): “If it would end the war, the Ukrainians should target major Russian cities with airstrikes, even if that means many Russian civilians would die”, “If it would end the war, the Ukrainians should firebomb major Russian cities, even if that means many Russian civilians would die”, “If it would end the war, the Ukrainians should show no mercy toward Russian civilians”, “If it would end the war, the Ukrainians should make life unbearable for Russian civilians”, and “If it would end the war, countries allied with Ukraine should drop nuclear bombs on major Russian cities” ($\alpha = 0.88$).

7.1.3. Procedure

As in Study 2, all participants first read a mock vignette and were immediately directed to the end of the survey if they failed a comprehension question about it, as were those who did not report being a US citizen in a subsequent demographics block. The remaining 1513 participants comprising our final sample were then randomly assigned to receive the intervention ($n = 731$) or to an empty control condition ($n = 782$). They then reported their dehumanization and dislike of Russian soldiers (presented in counterbalanced order), and then their support for war crimes against Russian soldiers. Finally, they reported their support for war crimes against Russian civilians.

7.2. Results

Descriptive statistics for each condition are reported in Supplementary Table 11. Intercorrelations among measured variables are reported in Supplementary Table 12.

We first performed independent-samples t -tests. A manipulation check confirmed that the humanization intervention reduced Americans’ dehumanization of Russian soldiers, $M_{diff} = 0.48$ [0.40, 0.57], $t(1,511) = 10.79$, $d = 0.55$, $SE = 0.04$, $p < .001$. The intervention also reduced dislike of the soldiers, $M_{diff} = 1.92$ [1.58, 2.23], $t(1,511) = 11.84$, $d = 0.61$, $SE = 0.17$, $p < .001$. Importantly, an ANCOVA revealed that the intervention’s effect on reduced dehumanization held when adjusting for reductions in dislike, $M_{Diff(Adj)} = 0.35$ [0.26, 0.44], $t(1,511) = 7.74$, $d = 0.40$, $SE = 0.05$, $p < .001$, suggesting the intervention humanized Russian soldiers over and above making them appear more likeable. Additional t -tests revealed that the intervention reduced support for war crimes—both against Russian soldiers, $M_{Diff} =$

0.52 [0.40, 0.63], $t(1,511) = 8.71$, $d = 0.45$, $SE = 0.06$, $p < .001$, and against civilians, $M_{Diff} = 0.25$ [0.13, 0.38], $t(1,508) = 4.18$, $d = 0.22$, $SE = 0.06$, $p < .001$. See Fig. 2 for the distribution of support for war crimes against Russian soldiers and civilians between the conditions, and Fig. 3 for bar plots comparing mean values on these variables between the conditions.

Finally, we tested whether the effects of the intervention on reducing war crime support were mediated by reduced dehumanization and dislike. We performed two parallel mediations using the PROCESS macro for SPSS (Hayes, 2017). In the first, we entered a binary variable denoting experimental condition (0 = control, 1 = intervention) as the predictor, dehumanization and dislike as parallel mediators, and war crime support against Russian soldiers as the outcome. We observed an indirect effect of the intervention on reduced war crime support via reduced dehumanization, $\beta_{indirect} = -0.23$ [-0.28, -0.18], $SE = 0.03$. However, dislike appeared to act as a suppressor—there was a weak, positive indirect effect of the intervention on war crime support through dislike, $\beta_{indirect} = 0.07$ [0.04, 0.10], $SE = 0.02$. We then repeated this procedure with support for war crimes against civilians entered as the outcome variable. Again, reductions in dehumanization mediated the effect of the intervention on reduced war crime support, $\beta_{indirect} = -0.15$ [-0.19, -0.11], $SE = 0.02$, while dislike acted as a suppressor, $\beta_{indirect} = 0.04$ [0.01, 0.08], $SE = 0.02$ (see Fig. 4).⁶ For an elaboration on the divergent indirect effects of dehumanization and dislike, see the “Dehumanization: Beyond Dislike” section in the Supplementary Information.⁷

8. Discussion

Victims of unjust military attacks often retaliate with reciprocal atrocities that ignite cycles of intensifying bloodshed (Edele, 2016; Weingartner, 1992). Although powerful third parties can de-escalate incipient conflicts, their citizens may instead support committing retributive war crimes against the aggressors (Sagan & Valentino, 2019; Watkins & Goodwin, 2020). Indeed, following Russia’s blatantly unjust invasion of Ukraine, many Americans dehumanized Russian soldiers, which predicted their support for retaliating against these soldiers with unambiguous war crimes. Thus, we aimed to reduce Americans’ war crime support by exposing them to the essentially human moral qualities of rank-and-file Russian soldiers. To intervene during the war’s earliest stages, we curated online media and identified the most effective content with an intervention tournament. This enabled us to deploy an intervention which reduced support for war crimes against Russian soldiers and—consistent with (parasocial) contact theory (Brown & Hewstone, 2005; Rothbart & John, 1985)—extended to dampen support for war crimes against Russian civilians.

These results provide an important, ecologically valid advance on prior work pointing to the centrality of morality in perceptions of humanity (Brandt & Reyna, 2011; Phillips, 2022; Schwartz & Struch, 1989; Strohminger & Nichols, 2014) and demonstrating that highlighting perpetrators’ moral agency or remorse can (re)humanize them (Borinca et al., 2021; Jin et al., 2022; Khamitov et al., 2016). Whereas that

⁶ By “suppressor”, we mean a variable that increases the predictive validity of another variable by controlling for a third variable that obscures the relationship (Rucker, Preacher, Tormala, & Petty, 2011). Because dislike is negatively correlated with both the Humanization Intervention and support for war crimes, controlling for it in these mediation models served to amplify the relationship between the Humanization Intervention and reduced support for war crimes. Essentially, when dislike is high, it may obscure the true effect of the Humanization Intervention; when dislike is accounted for, the full effect of the Humanization Intervention on reducing support for war crimes is stronger.

⁷ Nonetheless, caution is warranted when inferring causality on the basis of cross-sectional mediation (Fiedler et al., 2011). We return to this point in the Discussion.

research was theoretical or used relatively contrived paradigms (e.g., hypothetical vignettes; Khamitov et al., 2016), we drew on it to develop an intervention with real media content that reduced support for war crimes during an emotionally charged military attack—demonstrating the promise of applying insights from basic dehumanization research to promote peace (see also Landry, Schooler, Willer, & Seli, 2023; Moore-Berg, Hameiri, & Bruneau, 2022).

We also demonstrate the importance of differentiating dehumanization from dislike. Although the two constructs are often tightly intertwined, dehumanization may go beyond “mere” dislike—by stripping its target of the moral elevation afforded to fellow human beings, it can render them particularly vulnerable to extreme harm (Fincher et al., 2018; Kteily & Landry, 2022). Our findings lend compelling support to this notion: it was specifically the intervention’s effect on reducing dehumanization—not dislike—that predicted downstream reductions in support for retributive war crimes. Combined with four correlational studies consistently linking dehumanization—but not dislike—to war crime support (see “Dehumanization: Beyond Dislike” in the Supplementary Information), these findings demonstrate dehumanization’s unique explanatory power. It appears that negative affect alone does not drive support for the extreme violence considered here. Rather, only those who considered Russians beneath the human pale supported such violence. We encourage further research to systematically investigate when and why dehumanization and dislike diverge in their effects on extreme violence (see also Landry, Druckman, & Willer, 2023).

Practically, our method can inform further efforts to de-escalate emergent social crises. Previous research has demonstrated the power of modern media to reduce support for intergroup violence (Bilali, 2022; Hameiri et al., 2016). However, that work intervened in long-established conflicts with time-intensive methods. Here, we rapidly intervened during the outbreak of an escalating war. To do so, we harnessed rich-yet-accessible online media in a three-step process. We first deduced a plausible psychological mechanism (dehumanization) contributing to our outcome of interest. We then searched for content targeting this mechanism, enabling us to efficiently parse the vast abundance of online media. Finally, we identified the most effective content with the intervention tournament method gaining prominence in the biological and psychological sciences (Hameiri & Moore-Berg, 2022; see also Bruneau et al., 2018; Moore-Berg, Hameiri, Falk, & Bruneau, 2022). This allowed us to deploy an effective intervention within days.

Nonetheless, we do not wish to downplay the necessity of approaches that precisely identify *why* interventions work with tightly controlled stimuli. Indeed, a major limitation of our method is that, in curating media developed for purposes other than experimental research, our intervention contained a range of content beyond that assumed to be driving the effects. For instance, although we aimed to humanize Russian soldiers by highlighting their moral agency and remorse, the descriptions of the hardships they suffered may have also humanized them by evoking empathy (Moore-Berg, Hameiri, & Bruneau, 2022).⁸

Relatedly, although our mediational results suggest that humanizing Russian soldiers drove reductions in war crime support, caution is warranted when interpreting cross-sectional indirect effects (Fiedler, Schott, & Meiser, 2011). In fact, the link between dehumanization and war crime support may be bidirectional, because dehumanization can also follow violence as perpetrators become inured to harm-doing or seek

⁸ Another plausible mechanism is *meta-dehumanization*, or the perception that an outgroup dehumanizes oneself or one’s ingroup (Kteily et al., 2016). Recent work has found that many Americans feel dehumanized by Russians, which causes them to respond with reciprocal dehumanization (Landry, Ihm, & Schooler, 2022). Thus, seeing these conciliatory young soldiers may have reduced Americans’ meta-dehumanization, leading to downstream reductions in their own dehumanization of Russians (see Kteily et al., 2016; Landry, Schooler, et al., 2023). We thank the editor for this insightful suggestion.

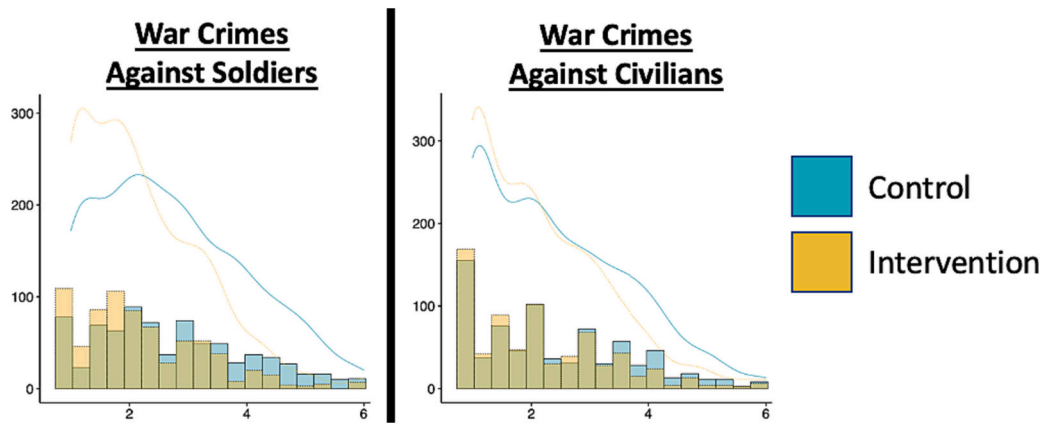


Fig. 2. Combined Histogram and Density Plots of the Distribution of War Crime Support for Each Condition.

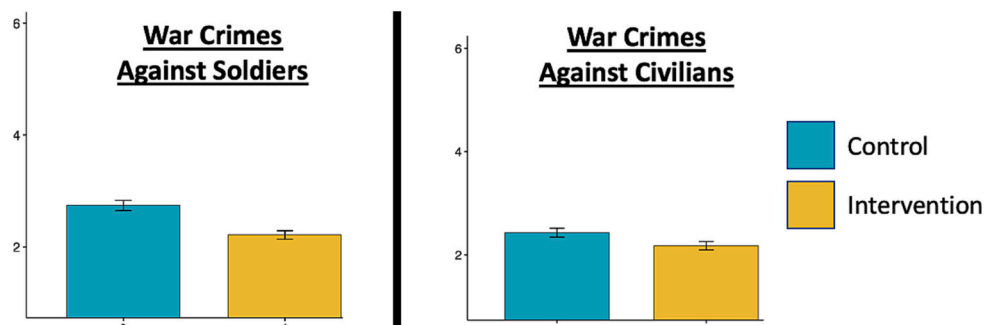


Fig. 3. Bar Plots Depicting Effects of the Humanization Intervention on War Crime Support.

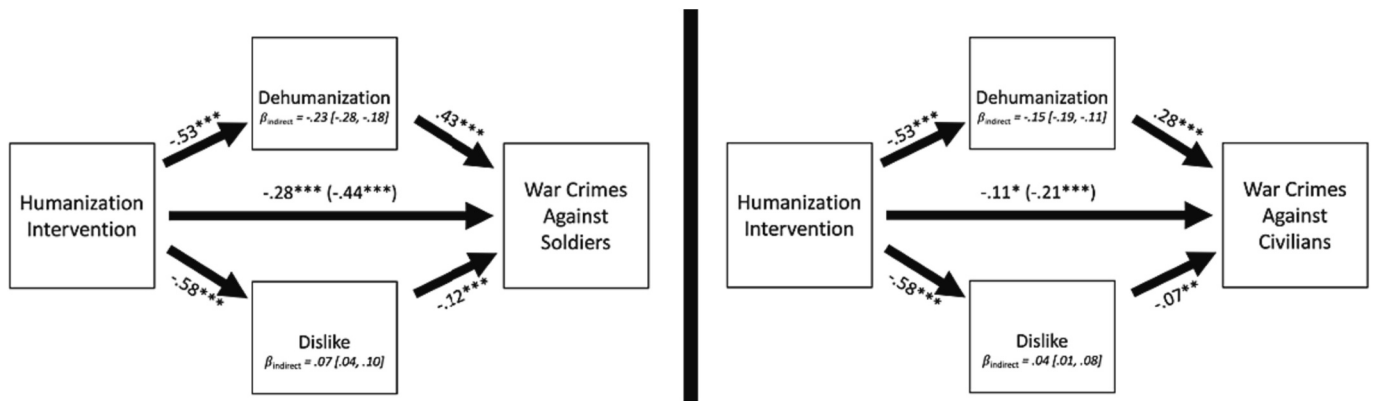


Fig. 4. Indirect Effects of the Intervention on War Crime Support, Mediated by Dehumanization and Dislike. Note. Values reflect standardized beta coefficients. Total effect is in parentheses. *** $p < .001$, ** $p = .005$; * $p = .044$.

to rationalize their behavior (Bandura, 1999; Luft, 2023). Encouragingly, because of this mutually reinforcing link, the modest reductions in war crime support we observed could perhaps be amplified through positive feedback. This possibility should be investigated with longitudinal designs. Longitudinal tests are also needed to determine whether our intervention’s effects endure for a meaningful length of time, as the durability of light-touch interventions is often questionable (Paluck et al., 2021).

Although powerful third parties can play a crucial role in stemming wartime atrocities (Krain, 2012; Kydd & Straus, 2013; Sikkink, 2011), it is also important to consider whether our intervention would have the same effects among those directly embroiled in the violent conflict. On the one hand, dehumanization may be more entrenched among this population, especially considering the tendency for such perceptions to

ossify in a conflict as protracted and brutal as that between Russia and Ukraine⁹ (see Bar-Tal & Halperin, 2011). Encouragingly, however, parasocial contact can have humanizing effects even among those more directly involved in violent conflicts (Bruneau et al., 2022; see also Hameiri et al., 2016). With the upmost care and ethical consideration, future research could test the efficacy of humanizing media content among those more proximal to the Russo-Ukrainian war.

A final concern is the intervention’s applied potential. We exposed individuals to the humanity of Russians by offering them monetary incentives to participate in an opt-in experiment. Given motivations to

⁹ For historical context regarding the protracted Russo-Ukrainian conflict, see Plokhly (2023)

maintain a consistent worldview (Kunda, 1990), it is questionable whether those who dehumanize Russians would willingly engage with this disconfirming information in the “real-world”, absent such extrinsic incentives. This limitation confronts most intervention studies that rely on extrinsic incentives, so researchers should place greater focus on adapting or developing more scalable methods of implementing their psychological interventions (see Landry & Halperin, 2023).

To address these limitations and outstanding questions, future work should (a) isolate the “active ingredients” of our intervention with tightly controlled stimuli, (b) confirm dehumanization’s causal impact on violence with experimental tests of mediation, (c) evaluate the intervention’s durability with longitudinal designs, (d) test the intervention’s generalizability in different populations, and (e) scale the intervention to naturalistic settings.

In sum, we demonstrate how psychological scientists can harness rich-yet-accessible media to craft timely and effective interventions. Highlighting the essentially human moral qualities of rank-and-file Russian soldiers reduced support for committing retributive war crimes against them and uninvolved Russian civilians. We thus advance theories of dehumanization and (parasocial) intergroup contact while providing a model for developing interventions with the potential to dampen incipient violence.

Open practices

This research was approved by Stanford University’s Institutional Review Board (IRB). All studies were performed in accordance with IRB guidelines and regulations and informed consent was obtained from all participants. All studies were preregistered. However, we deviated from our preregistrations in minor ways. We describe all deviations from these preregistrations and explain our rationale for doing so in the Supplementary Information. Preregistration forms, data, analysis code, and study materials can be accessed at https://osf.io/s83ye/?view_only=e2f7b2b1752a4de584700eb5e7d20b8c. For each study, all data was collected before any analyses were performed, and we report all manipulations and data exclusions. In Studies 1–2, participants completed two additional measures not reported in the main text (again, see the “Deviations from Preregistration” section in the Supplementary Information). Descriptions of these measures and data collected on them can be found in the Studies 1–2 preregistration forms and data files at the above OSF repository.

Declaration of Competing Interest

This work has not been published previously, is not under consideration for publication elsewhere, its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

Data availability

All data, analysis code, and study materials can be accessed at https://osf.io/s83ye/?view_only=e2f7b2b1752a4de584700eb5e7d20b8c

Acknowledgements

We thank Benoît Monin, Nir Halevy, Justin Berg, Ashley Martin, and Michal Kosinski for their feedback on an earlier draft of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jesp.2023.104567>.

References

- Abdul-Ahad, G. (2017, November 21). *After the liberation of Mosul, an orgy of killing*. The Guardian. <https://www.theguardian.com/world/2017/nov/21/after-the-liberation-of-mosul-an-orgy-of-killing>.
- Altemeyer, B. (1988). *Enemies of freedom: Understanding right-wing authoritarianism* (pp. xxix, 378). Jossey-Bass.
- Atran, S., & Ginges, J. (2012). Religious and sacred imperatives in human conflict. *Science*, 336(6083), 855–857. <https://doi.org/10.1126/science.1216902>
- Azar, E. E. (1990). *The Management of Protracted Social Conflict: Theory and cases*. Dartmouth.
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc.*, 3(3), 193–209. https://doi.org/10.1207/s15327957pspr0303_3
- Bar-Tal, D. (1989). Delegitimization: The extreme case of stereotyping and prejudice. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds.), *Stereotyping and prejudice: Changing conceptions* (pp. 169–182). Springer. https://doi.org/10.1007/978-1-4612-3582-8_8.
- Bar-Tal, D. (2007). Sociopsychological foundations of intractable conflicts. *American Behavioral Scientist*, 50(11), 1430–1453. <https://doi.org/10.1177/0002764207302462>
- Bar-Tal, D., & Halperin, E. (2011). Socio-psychological barriers to conflict resolution. In *Intergroup conflicts and their resolution: A social psychological perspective* (pp. 217–239). Psychology Press.
- Bastian, B., Denson, T. F., & Haslam, N. (2013). The roles of dehumanization and moral outrage in retributive justice. *PLoS One*, 8(4), Article e61842. <https://doi.org/10.1371/journal.pone.0061842>
- Baum, M. A., & Potter, P. B. K. (2015). War and democratic constraint: How the public influences foreign policy. In *In War and democratic constraint*. Princeton University Press. <https://doi.org/10.1515/9781400866472>.
- Bilali, R. (2022). Fighting violent extremism with narrative intervention: Evidence from a field experiment in West Africa. *Psychological Science*, 33(2), 184–195. <https://doi.org/10.1177/09567976211031895>
- Bizumic, B., & Duckitt, J. (2018). Investigating right wing authoritarianism with a very short authoritarianism scale. *Journal of Social and Political Psychology*, 6(1). <https://doi.org/10.5964/jssp.v6i1.835>. Article 1.
- Bloom, P. (2022). If everything is dehumanization, then nothing is. *Trends in Cognitive Sciences*, 26(7), 539. <https://doi.org/10.1016/j.tics.2022.03.001>
- Borinca, I., Falomir-Pichastor, J. M., Andrighetto, L., & Halabi, S. (2021). Overcoming negative reactions to prosocial intergroup behaviors in post-conflict societies: The power of intergroup apology. *Journal of Experimental Social Psychology*, 95, Article 104140. <https://www.sciencedirect.com/science/article/pii/S0022103121000408>.
- Brandt, M. J., & Reyna, C. (2011). The chain of being: A hierarchy of morality. *Perspectives on Psychological Science*, 6(5), 428–446. <https://doi.org/10.1177/1745691611414587>
- Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. In , Vol. 37. *Advances in experimental social psychology* (pp. 255–343). Academic Press. [https://doi.org/10.1016/S0065-2601\(05\)37005-5](https://doi.org/10.1016/S0065-2601(05)37005-5).
- Bruneau, E., Casas, A., Hameiri, B., & Kteily, N. (2022). Exposure to a media intervention helps promote support for peace in Colombia. *Nature Human Behaviour*, 1–11. <https://doi.org/10.1038/s41562-022-01330-w>
- Bruneau, E., Hameiri, B., Moore-Berg, S. L., & Kteily, N. (2021). Intergroup contact reduces dehumanization and meta-dehumanization: Cross-sectional, longitudinal, and quasi-experimental evidence from 16 samples in five countries. *Personality and Social Psychology Bulletin*, 47(6), 906–920. <https://doi.org/10.1177/0146167220949004>
- 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(3), 430–448. <https://doi.org/10.1177/0146167217744197>
- Carlsmith, K. M., & Darley, J. M. (2008). Psychological aspects of retributive justice. In , Vol. 40. *Advances in experimental social psychology* (pp. 193–236). Academic Press. [https://doi.org/10.1016/S0065-2601\(07\)00004-4](https://doi.org/10.1016/S0065-2601(07)00004-4).
- Cohen, R. (2022). *A surge of moral outrage over Russia’s war*. The New York Times. <https://www.nytimes.com/2022/03/01/world/europe/zelensky-ukraine-war-outrage.html>.
- Coleman, P. T. (2014). Intractable conflict. In *The handbook of conflict resolution: Theory and practice* (3rd ed., pp. 708–744). Jossey-Bass/Wiley.
- Collins, R. (2012). C-escalation and D-escalation: A theory of the time-dynamics of conflict. *American Sociological Review*, 77(1), 1–20. <https://doi.org/10.1177/0003122411428221>
- Dancy, G. (2016). Searching for deterrence at the international criminal court special issue: Strengthening the validity of international criminal tribunals. *International Criminal Law Review*, 17(4), 625–655. <https://heinonline.org/HOL/P?h=hein.journals/intcrimlr17&i=631>.
- Dixon, W. J. (1996). Third-party techniques for preventing conflict escalation and promoting peaceful settlement. *International Organization*, 50(4), 653–681. <https://doi.org/10.1017/S0020818300033543>
- Dower, J. (2012). *War without mercy: Race and Power in the Pacific war*. Knopf Doubleday Publishing Group.
- Edele, M. (2016). Take (no) prisoners! The red Army and German POWs, 1941–1943. *The Journal of Modern History*, 88(2), 342–379. <https://doi.org/10.1086/686155>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>

- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2), 63–87. [https://doi.org/10.1016/S1090-5138\(04\)00005-4](https://doi.org/10.1016/S1090-5138(04)00005-4)
- Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, 47(6), 1231–1236. <https://doi.org/10.1016/j.jesp.2011.05.007>
- Fincher, K. M., Kteily, N. S., & Bruneau, E. G. (2018). Our humanity contains multitudes: Dehumanization is more than overlooking mental capacities. *Proceedings of the National Academy of Sciences*, 115(15), E3329–E3330. <https://doi.org/10.1073/pnas.1800359115>
- Fiske, A. P., & Rai, T. S. (2014). *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge University Press.
- Fiske, S. T., Harris, L. T., & Cuddy, A. J. C. (2004). Why ordinary people torture enemy prisoners. *Science*, 306(5701), 1482–1483. <https://doi.org/10.1126/science.1103788>
- Foucault, M. (2023). Discipline and punish. In *Social theory re-wired* (3rd ed.). Routledge.
- Garland, D. (1993). *Punishment and modern society: A study in social theory*. University of Chicago Press.
- Goldhagen, D. J. (2009). *Worse than war: Genocide, eliminationism, and the ongoing assault on humanity*. Hachette UK.
- Greenaway, K. H., Wright, R. G., Willingham, J., Reynolds, K. J., & Haslam, S. A. (2015). Shared identity is key to effective communication. *Personality and Social Psychology Bulletin*, 41(2), 171–182. <https://doi.org/10.1177/0146167214559709>
- Haddock, G., Zanna, M. P., & Esses, V. M. (1993). Assessing the structure of prejudicial attitudes: The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology*, 65(6), 1105–1118. <https://doi.org/10.1037/0022-3514.65.6.1105>
- Hameiri, B., & Moore-Berg, S. (2022). Intervention tournaments: An overview of concept, design, and implementation. *Perspectives on Psychological Science*, 17456916211058090. <https://doi.org/10.1177/17456916211058090>
- Hameiri, B., Porat, R., Bar-Tal, D., & Halperin, E. (2016). Moderating attitudes in times of violence through paradoxical thinking intervention. *Proceedings of the National Academy of Sciences*, 113(43), 12105–12110. <https://doi.org/10.1073/pnas.1606182113>
- Harding, L. (2022, March 4). *Demoralised Russian soldiers tell of anger at being 'duped' into war*. The Guardian. <https://www.theguardian.com/world/2022/mar/04/russian-soldiers-ukraine-anger-duped-into-war>.
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis* (2nd ed.). A Regression-Based Approach: Guilford Publications.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., ... Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–1770. <https://doi.org/10.1126/science.1127333>
- Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., ... Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO_x scale. *Journal of Personality and Social Psychology*, 109(6), 1003–1028. <https://doi.org/10.1037/pspi0000033>
- Holsti, O. R. (2004). *Public opinion and American foreign policy* (Revised ed.). University of Michigan Press.
- Horton, D., & Wohl, R. (1956). Mass communication and Para-social interaction. *Psychiatry*, 19(3), 215–229. <https://doi.org/10.1080/00332747.1956.11023049>
- Jackson, J. C., Choi, V. K., & Gelfand, M. J. (2019). Revenge: A multilevel review and synthesis. *Annual Review of Psychology*, 70(1), 319–345. <https://doi.org/10.1146/annurev-psych-010418-103305>
- Jin, W. J., Park, S. H., & Park, J. (2022). Apology and its acceptance: Perceived reconciliatory attitudes reduce outgroup dehumanization. *Frontiers in Psychology*, 13, Article 809513. <https://doi.org/10.3389/fpsyg.2022.809513>
- Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, 60(1), 307–337. <https://doi.org/10.1146/annurev-psych.60.110707.163600>
- Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, 129(3), 339–375. <https://doi.org/10.1037/0033-2909.129.3.339>
- Kelman, H. G. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues*, 29(4), 25–61. <https://doi.org/10.1111/j.1540-4560.1973.tb00102.x>
- Khamitov, M., Rotman, J. D., & Piazza, J. (2016). Perceiving the agency of harmful agents: A test of dehumanization versus moral typecasting accounts. *Cognition*, 146, 33–47. <https://www.sciencedirect.com/science/article/pii/S0010027715300652?via%3Dihub>
- Krain, M. (2012). J'accuse! Does naming and shaming perpetrators reduce the severity of genocides or Politicides? *International Studies Quarterly*, 56(3), 574–589. <https://doi.org/10.1111/j.1468-2478.2012.00732.x>
- Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology*, 109(5), 901–931. <https://doi.org/10.1037/pspp0000048>
- Kteily, N., Hodson, G., & Bruneau, E. (2016). They see us as less than human: Metadehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology*, 110(3), 343–370. <https://doi.org/10.1037/pspa0000044>
- Kteily, N. S., & Landry, A. P. (2022). Dehumanization: Trends, insights, and challenges. *Trends in Cognitive Sciences*, 26(3), 222–240. <https://doi.org/10.1016/j.tics.2021.12.003>
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. <https://doi.org/10.1037/0033-2909.108.3.480>
- Kydd, A. H., & Straus, S. (2013). The road to hell? Third-party intervention to prevent atrocities. *American Journal of Political Science*, 57(3), 673–684. <https://doi.org/10.1111/ajps.12009>
- Landry, A., Druckman, J., & Willer, R. (2023). Need for chaos and dehumanization are robustly associated with support for partisan violence. *PsyArXiv*. <https://doi.org/10.31234/osf.io/pgksz>
- Landry, A., & Halperin, E. (2023). Intergroup psychological interventions: The motivational challenge. *PsyArXiv*. <https://doi.org/10.31234/osf.io/e5bcu>
- Landry, A., Ihm, E., & Schooler, J. W. (2022). Hated but still human: Metadehumanization leads to greater hostility than metaprejudice. *Group Processes & Intergroup Relations*, 25(2), 315–334. <https://doi.org/10.1177/1368430220979035>
- Landry, A., Schooler, J. W., Willer, R., & Seli, P. (2023). Reducing explicit blatant dehumanization by correcting exaggerated meta-perceptions. *Social Psychological and Personality Science*, 14(4), 407–418. <https://doi.org/10.1177/19485506221099146>
- Leidner, B., Castano, E., & Ginges, J. (2013). Dehumanization, retributive and restorative justice, and aggressive versus diplomatic intergroup conflict resolution strategies. *Personality and Social Psychology Bulletin*, 39(2), 181–192. <https://doi.org/10.1177/0146167212472208>
- Lieberman, P. (2006). An eye for an eye: Public support for war against evildoers. *International Organization*, 60(3), 687–722. <https://doi.org/10.1017/S002081830606022X>
- Lieberman, P., & Skitka, L. (2019). Vicarious retribution in US public support for war against Iraq. *Security Studies*, 28(2), 189–215. <https://doi.org/10.1080/09636412.2019.1551568>
- Lieberman, P., & Skitka, L. J. (2017). Revenge in US public support for war against Iraq. *Public Opinion Quarterly*, 81(3), 636–660. <https://doi.org/10.1093/poq/nfx005>
- Lickel, B., Miller, N., Stenstrom, D. M., Denson, T. F., & Schmader, T. (2006). Vicarious retribution: The role of collective blame in intergroup aggression. *Personality and Social Psychology Review*, 10(4), 372–390. https://doi.org/10.1207/s15327957pspr1004_6
- Litman, L., & Robinson, J. (2020). *Conducting online research on Amazon mechanical Turk and beyond*. SAGE Publications.
- Litman, L., Rosenzweig, C., & Moss, A. (2020, July 15). *New solutions dramatically improve research data quality on MTurk*. CloudResearch. <https://www.cloudresearch.com/research/blog/new-tools-improve-research-data-quality-mturk/>.
- Luft, A. (2023). The moral career of the genocide perpetrator: Cognition, emotions, and dehumanization as a consequence, not a cause, of violence. *Sociological Theory*, 07352751231203716. <https://doi.org/10.1177/07352751231203716>
- Moore-Berg, S. L., Ankori-Karlinsky, L.-O., Hameiri, B., & Bruneau, E. (2020). Exaggerated meta-perceptions predict intergroup hostility between American political partisans. *Proceedings of the National Academy of Sciences*, 117(26), 14864–14872. <https://doi.org/10.1073/pnas.2001263117>
- Moore-Berg, S. L., Hameiri, B., & Bruneau, E. G. (2022). Empathy, dehumanization, and misperceptions: A media intervention humanizes migrants and increases empathy for their plight but only if misinformation about migrants is also corrected. *Social Psychological and Personality Science*, 13(2), 645–655. <https://doi.org/10.1177/19485506211012793>
- Moore-Berg, S. L., Hameiri, B., Falk, E., & Bruneau, E. (2022). Reducing islamophobia: An assessment of psychological mechanisms that underlie anti-islamophobia media interventions. *Group Processes & Intergroup Relations*, 13684302221085832. <https://doi.org/10.1177/13684302221085832>
- Nilsson, A., & Jost, J. T. (2020). The authoritarian-conservatism nexus. *Current Opinion in Behavioral Sciences*, 34, 148–154. <https://doi.org/10.1016/j.cobeha.2020.03.003>
- Over, H. (2021a). Falsifying the dehumanization hypothesis. *Perspectives on Psychological Science*, 16(1), 33–38. <https://doi.org/10.1177/1745691620969657>
- Over, H. (2021b). Seven challenges for the dehumanization hypothesis. *Perspectives on Psychological Science*, 16(1), 3–13. <https://doi.org/10.1177/1745691620902133>
- Paluck, E. L., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, 72(1), 533–560. <https://doi.org/10.1146/annurev-psych-071620-030619>
- Phillips, B. (2022). “They’re not true humans:” beliefs about moral character drive denials of humanity. *Cognitive Science*, 46(2), Article e13089. <https://doi.org/10.1111/cogs.13089>
- Plokhly, S. (2023). *The Russo-Ukrainian war: The return of history*. W. W. Norton & Company.
- Power, S. (2013). *“A problem from hell”: America and the age of genocide*. Basic Books.
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67(4), 741–763. <https://doi.org/10.1037/0022-3514.67.4.741>
- Rai, T. S., Valdesolo, P., & Graham, J. (2017). Dehumanization increases instrumental violence, but not moral violence. *Proceedings of the National Academy of Sciences*, 114(32), 8511–8516. <https://doi.org/10.1073/pnas.1705238114>
- Rosenberg, S. P., Galis, T., & Zucker, A. (2016). *Reconstructing atrocity prevention*. Cambridge University Press.
- Rothbart, M., & John, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *Journal of Social Issues*, 41(3), 81–104. <https://doi.org/10.1111/j.1540-4560.1985.tb01130.x>
- Rousseau, D. L., Gorman, B., & Baranik, L. E. (2023). Crossing the line: Disgust, dehumanization, and human rights violations. *Socius*, 9. <https://doi.org/10.1177/23780231231157686>
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5(6), 359–371. <https://doi.org/10.1111/j.1751-9004.2011.00355.x>

- Rummel, R. J. (2017). *Death by government: Genocide and mass murder since 1900*. Routledge. <https://doi.org/10.4324/9780203793756>
- Sagan, S. D., & Valentino, B. A. (2017). Revisiting Hiroshima in Iran: What Americans really think about using nuclear weapons and killing noncombatants. *International Security*, 42(1), 41–79. https://doi.org/10.1162/ISEC_a_00284
- Sagan, S. D., & Valentino, B. A. (2019). Just war and unjust soldiers: American public opinion on the moral equality of combatants. *Ethics & International Affairs*, 33(4), 411–444. <https://doi.org/10.1017/S0892679419000431>
- Schiappa, E., Gregg, P. B., & Hewes, D. E. (2005). The parasocial contact hypothesis. *Communication Monographs*, 72(1), 92–115. <https://doi.org/10.1080/0363775052000342544>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Schumann, K., & Ross, M. (2010). The benefits, costs, and paradox of revenge. *Social and Personality Psychology Compass*, 4(12), 1193–1205. <https://doi.org/10.1111/j.1751-9004.2010.00322.x>
- Schwartz, S. H., & Struch, N. (1989). Values, stereotypes, and intergroup antagonism. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds.), *Stereotyping and prejudice: Changing conceptions* (pp. 151–167). Springer. https://doi.org/10.1007/978-1-4612-3582-8_7
- Sikkink, K. (2011). *The Justice Cascade: How human rights prosecutions are changing world politics (the Norton series in world politics)*. W. W: Norton & Company.
- Slovic, P., Mertz, C. K., Markowitz, D. M., Quist, A., & Västfjäll, D. (2020). Virtuous violence from the war room to death row. *Proceedings of the National Academy of Sciences*, 117(34), 20474–20482. <https://doi.org/10.1073/pnas.2001583117>
- Smith, D. L. (2021). *Making monsters: The uncanny Power of dehumanization*. Harvard University Press.
- Strohinger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159–171. <https://doi.org/10.1016/j.cognition.2013.12.005>
- Tomz, M., Weeks, J. L. P., & Yarhi-Milo, K. (2020). Public opinion and decisions about military force in democracies. *International Organization*, 74(1), 119–143. <https://doi.org/10.1017/S0020818319000341>
- U.S. Security Cooperation with Ukraine. (2022, May 6). United States Department of State. <https://www.state.gov/u-s-security-cooperation-with-ukraine/>.
- Watkins, H. M., & Goodwin, G. P. (2020). A fundamental asymmetry in judgments of soldiers at war. *Journal of Experimental Psychology: General*, 149(3), 419–444. <https://doi.org/10.1037/xge0000666>
- Watkins, H. M., & Laham, S. M. (2020). The principle of discrimination: Investigating perceptions of soldiers. *Group Processes & Intergroup Relations*, 23(1), 3–23. <https://doi.org/10.1177/1368430218796277>
- Weingartner, J. J. (1992). Trophies of war: U.S. troops and the mutilation of Japanese war dead, 1941–1945. *Pacific Historical Review*, 61(1), 53–67. <https://doi.org/10.2307/3640788>